BEHAVIORAL AND BRAIN SCIENCES (2000) **23,** 513–571
Printed in the United States of America

Précis of *Neural organization: Structure, function, and dynamics*

Michael A. Arbib
USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520
arbib@pollux.usc.edu **www-hbp.usc.edu**

Péter Érdi
Department of Biophysics, KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, H-1525 Budapest, Hungary
erdi@rmki.kfki.hu **www.rmki.kfki.hu/biofiz/biophysics.html**

Abstract: *Neural organization: Structure, function, and dynamics* shows how theory and experiment can supplement each other in an integrated, evolving account of the brain's structure, function, and dynamics. (1) Structure: Studies of brain function and dynamics build on and contribute to an understanding of many brain regions, the neural circuits that constitute them, and their spatial relations. We emphasize Szentágothai's modular architectonics principle, but also stress the importance of the microcomplexes of cerebellar circuitry and the lamellae of hippocampus. (2) Function: Control of eye movements, reaching and grasping, cognitive maps, and the roles of vision receive a functional decomposition in terms of schemas. Hypotheses as to how each schema is implemented through the interaction of specific brain regions provide the basis for modeling the overall function by neural networks constrained by neural data. Synthetic PET integrates modeling of primate circuitry with data from human brain imaging. (3) Dynamics: Dynamic system theory analyzes spatiotemporal neural phenomena, such as oscillatory and chaotic activity in both single neurons and (often synchronized) neural networks, the self-organizing development and plasticity of ordered neural structures, and learning and memory phenomena associated with synaptic modification. Rhythm generation involves multiple levels of analysis, from intrinsic cellular processes to loops involving multiple brain regions. A variety of rhythms are related to memory functions. The Précis presents a multifaceted case study of the hippocampus. We conclude with the claim that language and other cognitive processes can be fruitfully studied within the framework of neural organization that the authors have charted with John Szentágothai.

Keywords: cognitive maps; computational neuroscience; dynamics; hippocampus; memory; modular architectonics; neural modeling; neural organization; neural plasticity; rhythmogenesis; Szentágothai

P1. In memory of John Szentágothai

The present Précis and *BBS* multiple book review are dedicated, with respect and affection, to the memory of János (John) Szentágothai. The Précis is based on the book *Neural organization: Structure, function, and dynamics* (Arbib, Érdi, & Szentágothai, 1997; we refer to the book as *Organization* in what follows).

The idea of writing *Organization* arose when the three authors took part in the first week of a school organized by Francesco Ventriglia on "Neural modeling and neural networks," which was held on the Isle of Capri in October of 1992, a week that included the celebration of John Szentágothai's 80th birthday. Szentágothai and Arbib had previously coauthored *Conceptual models of neural organization* (1975) and Szentágothai and Érdi (1989) had written papers together on the self-organization of the nervous system. These collaborations provided the basis

MICHAEL A. ARBIB is Professor of Computer Science (of which he is also Chair), Neuroscience, Biomedical Engineering, Electrical Engineering, and Psychology at the University of Southern California (USC). Born in England, Arbib grew up in Australia and received his Ph.D. in mathematics from MIT. After five years at Stanford, he became chairman of Computer and Information Science at the University of Massachusetts, Amherst in 1970 and remained in the department until moving to USC in 1986. The author or editor of more than 30 books, Arbib recently edited *The Handbook of Brain Theory and Neural Networks.* His current research focuses on brain mechanisms of visuomotor behavior, on neuroinformatics, and on the evolution of language.

PÉTER ÉRDI is Head of the Department of Biophysics of the KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences. He is a professor at the Technical University of Budapest, and Kossuth University Debrecen, and has a Szechenyi Professor Fellowship (1999–2002) to the Eötvös University, Budapest. He has been working on the applications of the theory of nonlinear dynamic systems to chemical and biological phenomena. The activity of his computational neuroscience research group has focused on the theoretical and computational approach to the olfactory system and of the hippocampus. Érdi had been a close coworker of John Szentágothai between 1981 and 1994.

 0140-525X/00 $12.50

for the new book. A lengthy draft had been completed at the time of John Szentágothai's death in September of 1994. In fact, John was working on the book that very morning. We thus start with a few words of appreciation for John's career.

John Szentágothai (1912–1994) is known for his many pioneering contributions to neuroanatomy. His scientific career begain in the mid-1930s, when he helped verify the neuron doctrine against the reticular theory. (His early papers appeared under his original family name of Schimert.) In the late 1930s and early 1940s he elaborated his secondary generation method as a technique for detecting pathways between brain regions. Szentágothai served in the Chair of Anatomy at Pécs University Medical School in Hungary from 1946 to 1963. Combining anatomical and physiological methods, he made pioneering studies on the vestibulo-ocular reflex arc, then worked on the functional anatomy of the spinal cord, brainstem, and cerebellum. He was also involved in neuroembryological and neuroendocrinological research. In 1963, Szentágothai moved to the First Department of Anatomy of the Semmelweis University Medical School, Budapest, where he worked until the penultimate day of his life.

Szentágothai's anatomical discoveries in the cerebellum, together with the physiological findings of John Eccles and Masao Ito, led to a fruitful cooperation and an epoch-making monograph, *The cerebellum as a neuronal machine* (Eccles et al. 1967). From the late 1960s his research concentrated on the functional organization of the cerebral cortex. He formulated (and refined in the light of new data) the modular architectonic principle of the cerebral cortex as the anatomical basis for physiologically defined cortical modules. He searched for "the essence of the neural" and hoped to find it in the self-organization of spontaneous (random) activity into biologically significant spatiotemporal activity. A characteristic autobiography entitled "Too 'much' and too 'soon' " was written for a Festschrift dedicated to him on his 70th birthday (Szentágothai 1982). For a brief summary of his activity written for his 80th birthday, see Zaborszky et al. (1992).

Szentágothai's final reflections on neuroscience are preserved in *Organization,* especially in Chapter 2 (save for the last section), Chapter 7 (save for the last section), section 8.1, section 9.1, and much of section 9.2. We thank Tamas Freund, Attila Gulyás, Miklos Réthelyi, George Székely, and especially Jozsef Takács for their help in editing portions of this material. In some sense, this material – as Szentágothai's last scientific writings – became a "sacred text" which we could not change beyond minor editing, much though we would have enjoyed debating some of his ideas with John had he still been alive. As we completed the other sections, our continuing "conversations" with John, based on many earlier interactions, strongly influenced our work.

P2. Introduction

Organization provides a comprehensive view of neural organization in the spirit of the cooperative development of theory and experiment. A "good" model is responsive to available data; an "interesting" set of data will test hypotheses that are theory-laden, whether the theory be formal or not. Our task is not to provide "final models" or "a complete unified theory of the brain." Rather, we seek to show how theory and experiment can supplement each other in an integrated, evolving account of structure, function, and dynamics. Much of modern neuroscience seems to us excessively reductionist, focusing on the study of ever smaller microsystems without appreciating their contribution to the behaving organism. We do not reject the data gained in this way, but are concerned to restore some equilibrium between systems neuroscience, cellular neuroscience, and molecular neuroscience.

For example, one of *Organization*'s recurrent themes is an attempt to bridge different levels of organization by linking the learning rules that structure a variety of brain regions both to the functional roles of those regions and to the emerging understanding of the neurochemistry of synaptic plasticity and its variation from region to region. This is but one of many ways in which we show how theory and experiment can be intricately intertwined in a continuing cycle of analysis and synthesis. We now turn to a brief characterization of the three approaches – structural, functional, and dynamical – which inform our account of neural organization. The obvious identification of authors is Szentágothai ↔ structure, Arbib ↔ function, and Érdi ↔ dynamics, but this is only a first approximation, for a functional model may involve more or less dynamics, and vice versa, and the anatomical data necessary to ground a model may be more or less those which attracted Szentágothai's attention.

P2.1. Structural approach

Studies of brain function and dynamics build on, and contribute to, an understanding of many brain regions and of the neural circuits which constitute them. *Organization* thus reviews anatomical data that integrate the overall spatial relations between a variety of brain regions with a selection of critical details of neural morphology and synaptic connectivity. This analysis of neural structure is guided by a developmental view that approaches the complexity of the adult nervous system through an understanding of the way in which that complexity emerges during embryogenesis, thus linking the structural approach to dynamical models of self-organization. The developing nervous system can generate movement before it becomes responsive to sensory stimuli, consonant with the emphasis on action-oriented perception in our functional studies, analyzing the ways in which sensory systems are specialized to serve a variety of behaviors. As a basis for our functional and dynamical analysis of a variety of systems, later chapters progress through regions of the brain which, singly or in combination, underlie these systems: the segmented part of the neuraxis (discussed as a case study in the Structural Overview of Ch. 2), the olfactory system, the hippocampus, the thalamus, the cerebral cortex, the cerebellum, and, finally, the basal ganglia.

P2.2. Functional approach

Organization first approaches complex functions such as the control of eye movements, reaching and grasping, the use of a cognitive map for navigation, and the roles of vision in these behaviors, by the use of *schemas* in the sense of units that provide a functional decomposition of the overall skill or behavior. A schema account becomes a brain model when we offer hypotheses as to how each schema is implemented through the interaction of specific brain regions. A

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

brain-based schema model may be tested by analysis of the behavior of animals with localized lesions or reversible inactivation of specific brain regions or by human brain imaging. Such a model provides the basis for modeling the overall function by neural networks that plausibly implement (usually in a distributed fashion) the schemas in the brain. Further analysis may then proceed bottom up (as the neural data drive further research) as well as top down (as we refine our schema-theoretic formulations).

The models in *Organization* use neural networks in the sense of computational neuroscience in which the structure of the network and the function of the neuron are constrained, at some appropriate level of detail, by the data of neuroanatomy and neurophysiology. This is in contrast to neural networks in the sense of connectionism in which the structure of the network is generic (e.g., a multilayered feedforward network, or a fully connected network) and the connections are determined by some "learning rule" that may be nonbiological, as in the case of backpropagation, rather than constrained by anatomical data.

P2.3. Dynamical approach

Dynamic system theory offers a conceptual and mathematical framework to analyze spatiotemporal neural phenomena occurring at different levels of organization, such as oscillatory and chaotic activity both in single neurons and in (often synchronized) neural networks, the self-organizing development and plasticity of ordered neural structures, and learning and memory phenomena associated with synaptic modification. We discuss a variety of rhythms (and arhythmia) found in the olfactory bulb and olfactory cortex, in the hippocampus, and in the thalamocortical system. In most cases we relate these rhythms to memory functions. We also study learning rules in both developmental processes (self-organization) and in the acquisition of a variety of behaviors. In this way, *Organization* grounds our functional analysis of neural organization in a dynamic systems analysis of the neural networks that implement the basic schemas.

P3. An outline of the book and this Précis

Part I of *Organization*, "Overviews," opens with a chapter, The many themes of neural organization, which expands on the above discussion of structure, dynamics, and function to introduce a variety of themes that weave in and out of subsequent chapters, binding the book into a moderately coherent whole. We then devote three chapters to detailed overviews of our three methods for understanding neural organization: a structural overview, a functional overview, and a dynamical overview.

Part II, "Interacting systems of the brain," uses a structural organization to order our integrated approach to structure, function, and dynamics. Almost all of Chapters 5 through 10 begin with a structural analysis of a specific brain region as the prelude to our account of the dynamics of the neural circuits and of the function of the region. Learning, memory, and plasticity are discussed in this functional and dynamic context. Chapter 5 looks at the role of rhythm generation and chaotic patterns in both olfactory bulb and olfactory cortex. Chapter 6 analyzes rhythm generation in, and memory functions of the hippocampus, as well as providing an extensive account of cognitive maps in the rat and declarative memory in humans. Chapter 7 offers a primarily structural account of the thalamus which emphasizes that, far from being simply a set of relay structures, it binds the cerebral cortex in a variety of subtle loops to sensory systems, cerebellum, and the basal ganglia. Chapter 8 first studies the modular structure and self-organization of visual cortex, then the role of different thalamocortical oscillatory rhythms in the transition between sleeping and waking; and it models the interaction of multiple cortical regions in vision, saccade control, and cortically guided reaching and grasping. Later chapters then extend our understanding of the cerebral cortex by showing how its function can be fully understood only through analysis of its "cooperative computation" with the cerebellum and basal ganglia. Chapter 9 analyzes the role of the cerebellum in both motor control and classical conditioning. Finally, Chapter 10 provides an account of the role of the basal ganglia in motor coordination and learning which contrasts their role with that of cerebellum and emphasizes the important role of the dopamine system in its functioning.

In this way, *Organization* provides a structural analysis of many important brain regions integrated with models of a number of the functions these regions serve, both singly and in concert, and of the dynamics of their neural networks. We study a variety of systems involved in sensory analysis (especially for olfaction and vision), rhythm generation, sensorymotor integration (with special attention to visual guidance of eye, arm, and hand movements), and for learning and memory, as well as offering an account of the self-organization of several components of the nervous system.

Part II concludes with Chapter 11, Prospects for a neuroscience of cognition, which both summarizes the progress exhibited in the preceding chapters and points the way for the broader use of our methodology in the future development of a cognitive neuroscience. Section 11.2, Multiple levels, multiple methodologies, and the need for their integration, analyzes the themes of this book from a more philosophical, methodological perspective. We start with a discussion of the implications of a nonmonolithic approach to the brain involving multiple views and multiple theories. We then examine general issues of brain theory, stressing the transition from the Cartesian reflex paradigm to the self-directed, self-organizing brain, and listing a number of principles of neural organization. The concluding section, 11.3, Implications and outlook for cognitive neuroscience, suggests ways in which the ideas developed in this volume may contribute to future work in cognitive neuroscience.

In summarizing the book's contributions, the next three sections of this Précis will examine "threads" associated with our three overarching themes of structure, function, and dynamics. As the reader will see, a number of sections of the book are cited under each of these headings, exemplifying our point that these three themes are not separate areas of study, but must be interwoven to provide coherent models that integrate and challenge a sufficiently broad range of data. In the same way, the emphasis of Chapters 5 to 10 on one brain region at a time – Chapter 5, olfactory bulb and olfactory cortex; Chapter 6, hippocampus; Chapter 7, thalamus; Chapter 8, cerebral cortex; Chapter 9, cerebellum; Chapter 10, basal ganglia – should not hide our essential claim that to understand the brain we must understand how many such regions work together in the integrated production of behavior. In any case, we now sum-

marize the next three sections of this Précis, each providing a different perspective from which to integrate themes developed throughout the entire volume. In each of these sections one subsection is expanded in a more detailed way to better illustrate our methodology. (To avoid confusion in what follows, sect. P*x* refers to sect. *x* of the Précis, whereas sect. *x simpliciter* refers to sect. *x* of *Organization.*)

Section P4 (based on sect. 11.1.1) examines the following threads for *structure:* (1) phylogeny and ontogeny; (2) the architectonic basis for analysis of function and dynamics of circuitry; and (3) the embedding of the region in loops and pathways integrating it with other regions. We note that the transition from structure to function rests crucially (but not only) on "filling in the signs" in the structure as a basis for the interplay of excitation and inhibition. We will discuss thread (2) more extensively.

Section P5 (based on sect. 11.1.2) studies the threads of *behavior:* (1) the place of schema theory in neuroscience; (2) high-level constraints on system modeling in general afforded by the study of behavior; and then the use of (3) the saccadic system and (4) reaching and grasping to challenge the creation of models of function and learning at the level of both schemas and neural networks. We will present thread (4) in a more detailed way below.

Section P6 (based on sect. 11.1.3) reviews the threads woven into our study of *dynamics:* (1) the basic dynamic concepts of fixed points, rhythmogenesis and synchronization, and chaos; (2) self-organization; (3) plasticity and the modeling of learning; and, finally, dynamics considered at the finer and finer levels of (4) compartmental modeling and (5) neurochemistry. In section P6 we emphasize thread (1).

After these three overview sections, section P7 (based on Ch. 6) presents a case study, illustrating the application of our pluralistic strategy to understanding the functional organization and performance of the hippocampus. We show how computational models help to integrate the results of studies at different levels of neural organization (from the subcellular level to the systems level incorporating the hippocampal formation) and of distinct subdisciplines of the neurosciences (from micro- and macroanatomy, and intracellular and EEG-level physiology through animal behavioral and psychological studies).

Finally, section P8 (based on sect. 11.3) samples our views on the prospects for a neuroscience of cognition. In *Organization,* we testify to the vitality of a computational neuroscience open to the data of empirical neuroscience by presenting not only our own models but also those of many other researchers. We thus particularly welcome commentary that presents examples of modeling and data collection that carry forward *Organization*'s program of developing a neuroscience in which the computational and empirical study of neural organization pay attention to all three dimensions of structure, function, and dynamics. While we will read with interest commentary that touches on Szentágothai's views, we will not in general be able to reply at any length in the Response.

P4. Structure

P4.1. Phylogeny and ontogeny

Section 2.1 introduced the idea that the embryology of a structure may make clear crucial relationships that may be obscured in the adult form. We used the segmented part of the neuraxis, that is, the spinal cord and lower brainstem, to ground discussion of the progressive loss of segmentation in the upper brainstem, including the diencephalon. Moreover, we showed that the embryonic nervous system is able to generate movement before it is able to respond to sensory stimuli, supporting the action-oriented view of brain function stressed in Chapter 3. Turning from ontogeny to phylogeny, we studied the evolution of the mammalian brain, seeing how, as the brain evolves, basic structures become overlaid with more and more complex structures that can both inhibit and coordinate what has evolved before. This evolutionary theme was taken up from a functional point of view in section 3.2, where we showed how brain function can be analyzed in a process of evolutionary refinement of models in which basic systems serve as the substrate for the designed "evolution" of more refined systems – new schemas often arise as "modulators" of existing schemas, rather than as new systems with independent functional roles.

Section 7.1 studied an intermediate embryonic stage to illuminate the relationship between the diencephalon and the telencephalic vesicles. The two developmentally different structures are welded together by the progressive emergence of fiber tracts (white matter) formed by the axonal processes of neurons – both ascending and descending. These processes, as in the formation of the spinal cord and the lower brainstem, have the tendency to gather in and to occupy the outer surface areas of the original neural tube. The structural overview of cerebral cortex began with a view of its development (sect. 8.1), then section 8.2 analyzed the development of two striking examples of modular architectonics in the primary visual cortex: ocular dominance columns and orientation columns. Finally, our structural view of the cerebellum was initiated (sect. 9.1) by a review of its phylogeny and ontogeny, showing how the relative size of different parts of cerebellum may vary from species to species, and tracing the delicate mechanisms of cell growth, migration, and interaction which yield the quasi-crystalline structure of the mature cerebellar cortex.

P4.2. The architectonic basis for the analysis of function and dynamics

Section 2.2 approached a hierarchy of levels of structural analysis – neurons; networks; integrated system – by presenting data that support Szentágothai's modular architectonic principle, namely that neuronal connectivity in a typical neural center is sufficiently specific to permit disassembly of the whole network into neuronal modules of characteristic internal connectivity; and the larger structure can be reconstituted by repetition of these modules.

Szentágothai analyzed the spinal gray matter and the lower brainstem in these terms, and then showed that the upper diencephalic and telencephalic parts of the brainstem do not retain the quasi-segmental arrangement of the lower neuraxis but that elements of the basic architectural principle of the neuraxis are preserved. Turning from modular structure, we noted the importance in many structures of local circuit neurons, and of the complex synaptic arrangements called glomeruli with their synaptic triads in, for example, the cerebellar cortex, the olfactory bulb, and some of the anterior thalamic nuclei.

Szentágothai applied the modular architectonics princi-

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

ple to the cerebral cortex, linking observations of anatomical regularities to the observations of Mountcastle (1957) on physiological "columns" in somatosensory cortex and of Hubel and Wiesel (1959) on visual cortex. However, we (Arbib & Érdi) coined the term "multiple models of modularity" (sect. 2.3) to stress that the search for a hierarchy of levels of analysis of neurons, networks, and integrated systems is not confined to the "column-like" structures of cerebral sensory cortex. Other important anatomical regularities discussed in *Organization* are the quasi-crystalline structure of the cerebellar cortex and the basic lamellar structure of the hippocampus.

Section 5.1 reviewed the layers, cell types, and synaptic organization of the olfactory bulb and then continued the story into the olfactory cortex, stressing how its laminar structure relates to the afferent fibers from the olfactory bulb, reexcitatory connections that may be the anatomical substrate for olfactory associative memory, commissural fiber systems connecting the two halves of the cortex, and neurochemically varied centrifugal inputs from different brain areas. Section 6.1 then provided a structural view of the hippocampus, treating in turn the intrinsic organization – cells and circuits – of the hippocampus, the hippocampal afferents and efferents, both cortical and subcortical, and basic quantitative data on cell numbers and on the convergence and divergence of connections. Special attention was paid to the "synaptic matrices" revealed in lamellae orthogonal to the long axis of the hippocampus – thus establishing a "basic circuit" for the analysis of hippocampal function.

Section 7.1 advanced the theme of modular architectonics by noting that frontally oriented successive discs of the cortex from front to aft have a mutual, although more clearly thalamocortical, relationship with close to sagittal discs of the thalamus.

Section 8.1 analyzed the synaptic connectivity of cortical neurons, noting the interplay of excitation and inhibition, and the role of neuron chains in cortical function and then returned to the modular architectonics principle with Szentágothai arguing for the modular structural organization of the cortex, but questioning the functional role of such units. Section 8.2 offered two striking examples of modular architectonics seen in the primary visual cortex: ocular dominance columns and orientation columns.

With this, we turn to a more detailed review of Szentágothai's contributions to *Organization,* mostly those where he explained his views in historical context, and then we will make some remarks on them in the light of recent results. Since we obviously do not have the competence in neuroanatomy that Szentágothai had, we note once again that although we welcome commentary on this material, we will not in general be able to respond at any length.

We start with the concept of *neuron chains of the cortex.* The principal question to be considered here concerns the main neuron chain for routing afferent impulse patterns through the particular piece (or region) of cortex under study. Even without exact knowledge of the details of synaptic connectivity, the early diagrams of Ramón y Cajal did convey the essence of the main neuron chains involved. The next crucial step was Lorente de Nó's magnificent abstraction in his chapter on the cortex written for the 1938 edition of J. F. Fulton's *Physiology of the central nervous system.* Although the abstraction went a bit too far in substituting a single bouton for the entire interneuronal connection, two very basic facts were recognized, but only one of them correctly interpreted. One fact recognized by Lorente de Nó was the essentially vertical (up and down) orientation of the neuron chains. The neuronal chains of the cortex are indeed preferentially vertically oriented. Even if some connections for intermediate (1–10 mm) distances are tangential, the final synaptically active parts of the axons are vertically oriented. The other fundamental principle of coupling was called reciprocity and multiplicity of synaptic connections. Although the observation was essentially correct, even if perhaps a bit exaggerated, both Ramón y Cajal and Lorente de Nó had a curious "blind spot" in their field of scientific vision. This is all the more astonishing, because nervous inhibition was postulated already by René Descartes, and was experimentally first observed by the Russian physiologist Sechenov (1863), and thoroughly elaborated upon in the works of C. S. Sherrington (1906). However the word *inhibition* does not occur according to our knowledge in any of the works of Ramón y Cajal and was explained by Lorente de Nó (1938) as effected by some kind of Wedensky effect (an outdated concept – ingenious, but basically wrong).

The revival of Golgi studies, especially in the USA by scientists like M. A. Arbib and A. E. Scheibel, C. Fox, the Lunds, and many others) – the Russian neurohistologists (Poljakov, Skolnik-Jaross, Leontovitch) had kept Golgi studies continuously on very high standards – was a prelude to efforts trying to understand neuron chains by a combination of Golgi stains, experimental axon and synapse degeneration, and chronically isolated but vascularized neural tissue block techniques, introduced by Szentágothai. However, this was only a transitory stage toward the more sophisticated approach introduced in the late 1970s and early 1980s by Peter Somogyi (1977; Somogyi et al. 1979; 1981). Given Szentágothai's attention to the work of Somogyi, we briefly turn here to Somogyi et al.'s (1998) discussion of "salient features of synaptic organization in the central cortex." Here the attempt is made to define basic cortical circuits on the basis of quantitative studies of the synaptic connectivity of identified cortical neurons and their molecular dissection. By studying the precise location of postsynaptic GABA and glutamate receptors at cortical synapses, Somogyi et al. are able to argue that, due to the exclusion of G protein-coupled receptors from the postsynaptic density, the presence of extrasynaptic receptors, and the molecular compartmentalization of the postsynaptic membrane, the synapse should include membrane areas beyond the membrane specialization. Subsequently, they examine five organizational principles:

1. The cerebral cortex consists of: (a) a large population of principal neurons reciprocally connected to the thalamus and to each other via axon collaterals releasing excitatory amino acids, and (b) a smaller population of mainly local circuit GABAergic neurons.

2. Differential reciprocal connections are also formed among GABAergic neurons.

3. All extrinsic and intracortical glutamatergic pathways terminate on both the principal and the GABAergic neurons, differentially weighted according to the pathway.

4. Synapses of multiple sets of glutamatergic and GABAergic afferents subdivide the surface of cortical neurons and are often co-aligned on the dendritic domain.

5. A unique feature of the cortex is the GABAergic axo-axonic cell, influencing principal cells through $GABA_A$ re-

ceptors at synapses located exclusively on the axon initial segment.

Having said all this, they find that the basic circuit proves highly adaptable when comparisons are made between cortical areas or layers:

> The basic circuit is most obvious in the hippocampus where a relatively homogeneous set of spatially aligned principal cells allows an easy visualization of the organizational rules. Those principles which have been examined in the isocortex proved to be identical or very similar. In the isocortex, the basic circuit, scaled to specific requirements, is repeated in each layer. As multiple sets of output neurons evolved, requiring subtly different needs for their inputs, the basic circuit may be superimposed several times in the same layer. Tangential intralaminar connections in both the hippocampus and isocortex also connect output neurons with similar properties, as best seen in the patchy connections in the isocortex. The additional radial superposition of several laminae of distinct sets of output neurons, each representing and supported by its basic circuit, requires a co-ordination of their activity that is mediated by highly selective interlaminar connections, involving both the GABAergic and the excitatory amino acid releasing neurons. The remarkable specificity in the geometry of cells and the selectivity in placement of neurotransmitter receptors and synapses on their surface, strongly suggest a predominant role for time in the coding of information, but this does not exclude an important role also for the rate of action potential discharge in cortical circuitry. (Somogyi et al. 1998)

It is interesting to see that Somogyi et al. (1998) emphasized the common organization principles of the basic circuits of the hippocampus and of a single layer of the isocortex. With this we return to Szentágothai's words: "The analyses possible in the hippocampal formations need further development to lead to an improved insight into the intricacies of the complex network of the many types of inhibitory neurons in the neocortex."

Now we continue with Szentágothai's discussion of the modular architectonics principle. The concept of a modular architectonics principle arose from two entirely independent sources: (1) the observation by the Scheibels (1958) of certain spatial regularities in the arborization both of dendrites and of axonal ramification in the lower brainstem; and (2) some kind of "columnar" organizations of the somatosensory cortex (Mountcastle 1957) and the even more convincing observation by Hubel and Wiesel (1959) of so-called "orientation columns" in the visual cortex. While the observations of the Scheibels had their main follow-up in our own (Szentágothai) studies in the spinal cord, the observations of Mountcastle and Hubel and Wiesel prompted efforts to adapt new findings in cortical histology to the new insight gained from emergent cortical physiology. Actually, a very first attempt by Szentágothai (1967) anticipated the concept of the "modular architectonics principle" under a different name: "elementary integrative unit." This expression was later abandoned because it was misleading and in fact led to certain misinterpretations (very unfortunately in some otherwise fundamental writings of Sir John Eccles). The first explicit modular neuron connectivity model of the cortex was proposed (Szentágothai 1969) for the somatosensory cortex in a diagram reproduced as Figure 8.12 in *Organization.* This model, as well as the earlier 1967 integrative unit model, was still under the influence – and mimicking – Szentágothai's earlier cerebellar cortex model. It took a further 10 years before, still using the earlier "pre-Somogyi" guesswork strategy, obtaining the more realistic model of Szentágothai (1983).

Though we cannot review the newer data and their interpretations in full detail, we (Arbib and Érdi) will briefly review some recent developments related to modular architectonics. The whole issue seems to be controversial.

Lev and White (1997) studied the organization of pyramidal cell apical dendrites in the primary motor cortex of the mouse. According to their data, apical dendrites of callosal neurons aggregate to form clusters that are composed exclusively of dendrites belonging to this type of projection cell. Thus the cellular composition of cortical modules seems to be much more specific than was thought earlier. They hinted that the sensory cortices may also have such a kind of modular organization.

Based on studies on the macaque monkey cortex, Britten (1998) reported that columnar organization exists beyond the "early" stages (e.g., area 17) of cortical processing. By analyzing response selectivity, a clustered organization of neurons sharing response properties was found in the medial superior temporal area of extrastriate cortex, which is most likely the highest visual area on this pathway.

Results of recent investigations of the cytochrome oxidase (CO) blobs can by and large be considered as arguments for the existence of columnar organization. Murphy et al. (1998) suggested that in visual cortex of normal and strabismic monkeys the fundamental periodicity of this particular columnar architecture was formed prenatally, and is not modifiable by experience. Elston et al. (1999) found a correspondence between the dendritic fields of layer V pyramidal cells and the CO bands, but they also found many examples where the dendrites crossed the boundary between bands.

Vernon Mountcastle (1997), one of the pioneers of the columnar organization of the neocortex, has recently stated that modules may vary in cell type and number, in internal and external connectivity, and in mode of neuronal processing between different large entities, but within any single large entity they have a basic similarity of internal design and operation. A cortical area defined by the rules of classical cytoarchitectonics may belong to different systems. Therefore distributed structures may serve as the anatomical bases of distributed function.

In *Organization,* we mentioned that Swindale (1990) explicitly criticized the modular architectonic principle and some debate between him and Szentágothai ensued. Swindale (1998) reiterated his interpretation and stated that instead of using the concept of modular organization of the cortex, recent studies reveal "a more fluid arrangement in which several separate maps are superimposed, with relatively weak geometric linkages and no common modular submit." In fact, Szentágothai considered cortical organization to be the result of the interplay between determinism and chance and never talked about rigid modules.

These few notes complete our brief tour of the architectonic basis for the analysis of function and dynamics. Section 9.2 studied the quasi-crystalline structure and the "space economy" of the cerebellar cortex in some quantitative detail. The only output cells of the cerebellar cortex are Purkinje cells, and these inhibit nuclear cells. Thus we stressed the integration of the circuitry of the cerebellar cortex into a "microcomplex" which unites a microzone of cortex with the region of nucleus to which it projects, and with which it shares afferents – thus establishing a "basic circuit" for the analysis of cerebellar function. Section 10.1 examined the structure of the striatum, the region of the

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

basal ganglia which receives most of its input pathways: by looking at the two pathways, direct and indirect, whereby it acts upon the output pathways of the basal ganglia; and by looking at the division of the striatum into "patches" embedded in a "matrix."

P4.3. The embedding of regions in loops and pathways

Section 2.3 emphasized that many regions of the brain are best thought of as embedded within even larger systems integrated by loops traversing many brain regions. Consider, for example, the links of the cerebellar system – "upstream" with the cerebral cortex and "downstream" with the spinal cord – which are closed in the cerebellar nuclei to which the output cells of the cerebellar cortex project. The fact that the output of cerebellar cortex is purely inhibitory ties into the theme that the passage from structure to function is often based on understanding the patterns of interplay of excitation and inhibition. In this case, the inhibition from the cerebellar cortex serves to modulate the activity in the cerebellar nuclei, which serves in turn to tune and coordinate motor pattern generators located elsewhere in the nervous system.

The role of the thalamus as the chief relay for sensory input to cerebral cortex is only a small fraction of the crucial role of the thalamus in all kinds of pathways. We thus devoted section 7.3 to thalamocortical loops and cooperative computation, reviewing the descending control of sensory systems in the lateral geniculate nucleus of mammals as well as thalamocortical oscillations. We then briefly reviewed the thalamocortical loops involving the basal ganglia and cerebellum. Section 9.3 then stressed that the microcomplexes mentioned above integrate cerebellar cortex into a cerebellar system that modulates motor pattern generators (MPGs). Section 10.1 showed the basal ganglia to be embedded in four disjoint loops of the form "cortex → striatum → SNr (substantia nigra pars reticulata) (thalamus → prefrontal cortex": the oculomotor circuit, the motor circuit, the "cognitive" (dorsolateral prefrontal) circuit, and the limbic (lateral orbitofrontal) circuit, providing our concluding example of the embedding of regions in loops and pathways.

P5. Function

P5.1. The place of schema theory in neuroscience

Chapter 3 presented schema theory as a framework for the rigorous analysis of behavior that requires no prior commitment to hypotheses on the localization of each schema (unit of functional analysis), but which can be linked to a structural analysis as and when this becomes appropriate. Section 3.1 introduced an approach to schema theory which emphasizes action-oriented perception, with the paradigm of the action-perception cycle replacing the stimulus-response paradigm. However, a number of later sections have made clear that schema theory and the action-perception cycle – and our approach to functional neuroscience in general – are not limited to those forms of, for example, sensorimotor coordination for which extensive neural data are available, and which we treated at length in Part II. These less-constrained topics include section 3.1.5 on Visual scene interpretation, section 6.5 on Hippocampal function and human memory, and section 8.6, From action-oriented perception to cognition, and the discussion in sections 11.3.4, Schema theory and the construction of reality, and 11.3.5, Language.

Section 3.1 explored the constraints imposed by linking schema theory to functional neuroscience and provided a quasi-formal introduction to perceptual and motor schemas, coordinated control programs (illustrated with an introduction to the visual control of reaching and grasping), cooperative computation, and schema assemblages (the basis for a schema-based model of visual perception which provided a perspective on short-term and long-term memory). A simple account of approach and avoidance behavior in frogs illustrated the use of lesion data in making a schema-based account of a function into a brain model.

P5.2. Behavioral constraints

Section 3.2 presented *Rana computatrix,* a set of models of visuomotor coordination in frog and toad, studying approach, avoidance, and detour behavior to show how perception may demand the mutual refinement of one perceptual schema by another, how multiple motor schemas may act together to yield complex motor behaviors, and how brain function can be analyzed in a process of evolutionary refinement. We then studied neural mechanisms of avoidance behavior to provide our first example of how neural modeling can be used to replace schemas with neural networks of equivalent functionality.

Sections 3.3 and 3.4 introduced schemas for looking, reaching, and grasping to demonstrate that much is to be learned at the level of schema analysis prior to, or in concert with, the analysis of neural circuitry. These schemas were shown, in Chapters 8 and 10, to be distributed across cerebral cortex and basal ganglia, and Chapter 9 showed the role of cerebellum in their adaptation and coordination.

Section 6.4 combined a functional view of the hippocampus – its role in the cognitive maps underlying navigation and spatial behavior in rats – with a dynamic view of how synaptic plasticity may enable hippocampal cells to learn to encode different "places" in a cognitive map. We offered a general framework for the study of spatial representation and cognitive maps in rats, including the general idea of World Graphs as cognitive maps for motivated behavior, and then reviewed the neurophysiology of spatial representation, with special emphasis on the "place cells" of the regions CA3 and CA1 of the hippocampus. We offered two contrasting systems views of the role of the hippocampus in navigation, in each case emphasizing that the representation of current place in CA3 and CA1 is insufficient for a cognitive map that underlies navigation. Section 6.5 viewed the role of hippocampal function in human memory, introducing the crucial dichotomies of procedural versus declarative memory and of skill versus episodic learning. The data suggest that the hippocampus is involved in declarative rather than procedural memory and in episodic rather than skill learning. We closed the chapter by discussing the very much open question: "Is there a commonality of mechanism between the two main functions attributed to the hippocampus: cognitive mapping in rats and declarative memory in humans?" (Because learning is a crucial aspect of adaptive behavior, a number of the issues discussed here under the heading of Function – Behavioral Constraints overlap those discussed below in sect. P6.3 Dynamics – Plasticity: Modeling learning.)

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

Section 8.2 offered a very brief look at how primary visual cortex provides input to a variety of visual processes, introducing psychophysical and neurophysiological data on spatial visual perception, and studying the role of long-range horizontal connections in the integration of information. We had little more to say about low-level visual processing, but section 8.4 did discuss cortical mechanisms for using vision in the control of movement. That section focuses on frontal-parietal interactions in cortex, but many other brain regions are involved in the integration of vision with action. Section 8.5 turned to the theme of learning of coordinated behaviors, providing both a schema-level analysis of motor set and the neuralization of coordinated control programs, and a specific neural network model of visual-motor conditional learning. This theme, foreshadowed in the study of rat spatial learning (Ch. 6), was further developed in our study of cerebellum (Ch. 9) and basal ganglia (Ch. 10). Section 8.6 then charted basic processes underlying cognitive functions from a high-level schema-theoretic viewpoint.

Section 9.3 focused on the issue of how skills are acquired, viewing the cerebellum as a learning machine, stressing the idea that cerebellar nuclei modulate motor pattern generators (MPGs) while the cerebellar cortex learns how best to modulate the cerebellar nuclei: modulating the modulator. We first studied the role of the cerebellum in the vestibulo-ocular reflex (VOR) where data strongly support the notion of a functional role of an adaptive microcomplex in modulating the gain of eye movements that compensate for head movements, and in the classical conditioning of the rabbit eye-blink response.

Section 9.4 then presented models of how the cerebellum adapts the metrics of movement to changing circumstances. We showed how the detailed circuitry of the cerebellar cortex and of various nuclei with which it interacts could modulate activity in MPG-related loops on a short-term basis, that is, one appropriate to the current circumstances as in adjusting to step height in climbing a flight of stairs. The remaining subsections reviewed approaches to modeling adaptation of motor control where the adaptation persists on a long-term basis, involving synaptic plasticity: we presented feedback-error learning whereby the cerebellum could "take over" motor control from other parts of the brain, but argued that the cerebellum "works" by modulating and coordinating multiple motor pattern generators (MPGs), rather than by replacing them.

In addition to behavior in normal subjects, we can learn about neural function and dynamics by studying their fate in subjects with a variety of diseases. Section 9.3 noted clinical data on the role of the cerebellum in skilled movements as a basis for our study of how these skills are acquired, viewing the cerebellum as a learning machine. Section 10.2 focuses on diseases of the basal ganglia, showing that the distinct movement disorders seen in Huntington's disease (hyperkinetic and hypotonic) and Parkinson's disease (hypokinetic and hypertonic) are associated with decreased basal ganglia output in Huntington's disease and a marked increase in Parkinson's disease.

P5.3. The saccadic system

Section 3.3 introduced schemas for controlling the rapid eye movements called saccades, and focused on the homology between the tectum in frog and toad and the superior colliculus in primates – the whole body movement of the frog toward its prey corresponding to the orienting of gaze toward a visual target in the monkey. It showed how schemas for working memory and for dynamic remapping may extend the monkey's saccadic repertoire to include saccades to remembered targets or to two targets in succession.

Later sections then showed how circuitry in various regions of the brain may contribute to these and other schemas. Section 8.4 presented a detailed model of cortico-thalamic systems for saccade control, with particular attention to the roles of posterior parietal cortex, frontal eye fields, and thalamus in providing mechanisms for dynamic remapping and working memory. Section 9.4 modeled the cerebellar role in saccade adaptation, refining the cortico-thalamic model by adjusting the metrics of saccades, a feat of learning that is impossible in animals or humans lacking certain portions of the cerebellum. Section 10.4 then completed our model of saccade control by focusing on the "oculomotor loop" of the basal ganglia, presenting a model of the role of the entire loop in generating saccades, and modeling the interactions between the basal ganglia and the "working memory" systems of prefrontal cortex. Interruption of SNr (substantia nigra pars reticulata) inhibition allows reciprocal connections between frontal eye fields and thalamus to generate a spatial "memory" cycle or loop. Once a saccade has been made to a remembered target, the memory trace must be erased to prevent generation of further saccades of equal magnitude and direction. We posit that the activity of this spatial working memory could be regulated by the inhibitory topographic projection from SNr to thalamus.

Section 10.6 suggested that if the frontal eye field and direct visual input do not yield the encoding of a unique target in the deep layers of SC (superior colliculus), then a winner-take-all (WTA) mechanism will "choose" one of the targets. We then argued that experience based on inferotemporal or prefrontal information may provide contextual, learned information to bias activity in the basal ganglia to "tip the balance" to one "winner" or another, presenting models of visual-motor conditioning, including spatial generalization and sequential behavior based on the strong hypothesis that this learning is mediated by cortico-striatal plasticity.

P5.4. Reaching and grasping

In studying the role of perception in mediating behavior, we stress that there is in general no complete and objective "percept" of an object, but rather a set of partial characterizations (including parameters that we may not be able to represent symbolically in any explicit fashion) related to the current set of goals and motivations of the observer – which may keep unfolding as interaction with, or contemplation of, the object continues. Section 3.4 illustrated this, using the schemas involved in reaching and grasping. We presented the concepts of virtual fingers and opposition space to offer a precise but compact description of the degrees of freedom involved in a number of grasping movements, and then turned to a series of experiments which motivated the design of a new coordinated control program that explicitly involves a coordinating schema as well as perceptual and motor schemas.

Section 8.4 turned to cortical systems for reaching and grasping, and we shall discuss this work in some detail. The work of Ungerleider and Mishkin (1982; Mishkin et al.

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

1983) distinguished two visual systems in extrastriate visual processing: the ventral system, V1 → V2 → V4 → IT (inferotemporal cortex), is characterized as the cortical *what* (pattern recognition) system; and the dorsal system, extending from V1 to PP (posterior parietal cortex) is characterized as the cortical *where* (object location) system. Goodale and Milner (1992) reviewed a variety of data including those on the ability of a patient with a ventral lesion to carry out a variety of object manipulations even though unable to verbally report on the object parameters used to guide these actions. They concluded that the dorsal system mediates the required sensorimotor transformations for visually guided actions directed at such objects and so extended the Mishkin-Ungerleider dichotomy to view the dorsal system as the *how* system since location (where) is only one of many properties needed to determine how to interact with an object.

P5.4.1. The FARS model. We have developed a detailed model, the FARS (Fagg-Arbib-Rizzolatti-Sakata) model, based on the interactions between the AIP (anterior intraparietal sulcus) area of PP, and the F5 area of premotor cortex in monkeys trained to grasp objects. About half of the neurons related to hand movements in AIP fired almost exclusively during one type of grip, with precision grip being the most represented grip type (Sakata et al. 1992; Taira et al. 1990). Some cells demonstrate specificity toward the size of the object to be grasped; and some cells demonstrated independence from the size of the object. A few cells show modulation based on the object's position and/or orientation in space. The visual responses of these cells thus provide a distributed code for *affordances* for grasping, that is, the various ways in which an object may be grasped (as distinct from recognizing *what* the object is). Most neurons in AIP also show phasic activity related to the motor behavior. Five identifiable phases occur in the paradigm used by Sakata to study these cells: set (*key phase*), preshape, enclose, hold (*object phase*), and ungrasp. Cells participate in varying degrees during different phases of the movement, but are usually most highly active during the preshape and enclose phases of movement. Very important: once an AIP cell becomes active, it typically remains active until the object is released.

The main anatomical connections of F5 are with AIP and the hand field of the precentral motor area (Matelli et al. 1985; Muakkassa & Strick 1979). Rizzolatti et al. (1988) described various classes of F5 neurons which discharge during specific hand movements (e.g., grasping, holding, tearing, manipulating). The largest class is related to grasping. The temporal relations between neuron discharge and grasping movements vary among neurons.

We now outline the FARS model, implemented in terms of simplified but biologically plausible neural networks (Fagg & Arbib 1998). Given visual input from an object, the model AIP computes its affordances. The corresponding set of grasps is passed to F5. As a function of task or other information supplied by prefrontal cortex, F5 selects one of the specified grasps, and is responsible for unfolding the grasp in time. F5 activity is broadcast back to AIP, strengthening the affordance that corresponds to the selected grasp. Motor responses in AIP are explained as corollary discharges from F5, and AIP provides an *active memory* for the grasp which is continuously updated. This is similar to the *dynamic remapping* mechanism in our study of saccades (*Organization,* sect. 8.4.2), in which motor afference updated a map of targets of potential eye movements.

The location of target objects is passed to F4, which represents the arm goal position. Since grasp programming affects arm movements, the model modulates F4 with information from AIP specific to the affordance/grasp pair selected by the AIP/F5 system. A neighboring region, the *posterior intraparietal area* (PIP), codes object-centered information (Sakata, personal communication) concerning different shapes presented to the monkey. In the model, PIP codes the shape and size of the object to be grasped. An affordance derived from PIP maps an object configuration to one possible grasp for that object. Castiello et al. (1991) studied impaired grasping in a patient (AT) with a lesion impairing the pathway V1 → PP, and found evidence for a mapping from object identity to affordances which is effective whenever the nature of the object supports such a mapping. The model thus includes a corresponding path PIP → IT → AIP.

The FARS model analyzes the interaction between AIP and F5 populations during execution of the Sakata paradigm. AIP units include visual-related cells which recognize objects that require a specific grasp and motor-related cells which are active for specific grasps. Each F5 unit fires during a different phase of the program. At each program phase, the state is reported back to the AIP motor-type population. The full model also includes the role of SII in creating and monitoring haptic expectations, the role of dorsal premotor cortex (F2) in the association of arbitrary stimuli with motor program preparation, and the role of area 46 as a working memory in tasks requiring information to be held during a delay period. However, these details (and the presentation of simulation results) were beyond the scope of *Organization* (see Fagg & Arbib 1998 for the details). We do stress, however, that the circuitry controlling F5 programs in the model is not intrinsic to F5: the effective connections between program states are not coded within F5 but are managed by the combined action of pre-SMA (F6) and the basal ganglia (BG).

One simulation study showed how the model performs when a delayed instruction stimulus is used to inform a subject how to grasp an object. The model is presented with a single object (a small cylinder), and asked to perform one of three tasks. The three different tasks are:

1. Grasp the cylinder using a precision pinch;
2. Grasp the cylinder using a side opposition; and
3. As a function of an instruction stimulus (e.g., the color of a light), grasp the cylinder using either a precision pinch or a side opposition.

(We speak rather loosely here. The model is not a robot. It transforms visual codes for objects to neural codes for movements via neural network models of diverse brain regions.) In the model, area F2 (dorsal premotor cortex) has a high level of activity in the conditional task as this region is involved only when the model must map an arbitrary stimulus to a motor program (in this case, a grasp); the region does not receive IS (instructional stimulus) inputs in the nonconditional task. F5 receives inputs from F2, causing an increase in the region's activity level which is passed on through excitatory connections to AIP.

P5.4.2. Synthetic PET. In further work, we sought to understand how our study of the monkey could be related to the results of human brain imaging. In the human, PET and

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

fMRI techniques allow us to achieve a global view of the systems involved in performing a task, but at the expense of a very coarse spatial and temporal resolution. In the monkey, on the other hand, we are able to examine individual cells and resolve single spikes, but have tremendous difficulty in examining entire circuits. It is thus important to develop techniques that allow experimental results at both levels to be brought together as we attempt to understand the different systems. Arbib et al. (1995) proposed *synthetic PET imaging* as a way to draw conclusions in one domain from experimental results in the other. The synaptic activity of a region A during a task is computed as the measure of instantaneous synaptic activity in region A integrated over the time required to perform the task (which might involve multiple trials). The simulated synaptic activity of a region can then be compared over several conditions.

With this background, we summarize one prediction made by applying synthetic PET to the FARS model. The conditional task is to grasp a cylinder using either a precision pinch or a side opposition, the choice being determined by an instruction stimulus (the color of a light). We tabulated regions in the model that demonstrate a change in synaptic activity in the conditional task above and beyond those involved when the subject knows a priori which of the two grasps to perform. The most significant change predicted by the model is the level of activity exhibited by area F2 (dorsal premotor cortex). Its high level of activity in the conditional task is due to the fact that this region is only involved when the model must map an arbitrary stimulus to a motor program (in this case, a grasp). In the nonconditional task, the region does not receive instruction stimulus (IS) inputs, and thus its synaptic activity is dominated by the general background activity in the region. The additional IS inputs in the conditional task have a second-order effect on the network, as reflected in small changes in activity in F5, BG, and AIP. The increased synaptic activity in F5 is due to the additional inputs from F2 (into the supporting inputs of some columns in F5). These inputs also cause an increase in the region's *activity level*, which is passed on through excitatory connections to AIP.

The above synthetic PET experiments raised some important questions about how instruction stimuli are mapped to arbitrary motor programs, and about the relative representation of different grasps. These predictions were tested in a human PET experiment (Grafton et al. 1998).

The model predicts that the conditional task should yield much higher activation in F2 (dorsal premotor cortex), some activation of F5, and a slight activation of AIP. The human experiment confirmed the F2 result, but failed to confirm the predictions for F5 and AIP. In fact, in humans there is an activation of the inferior parietal cortex (AIP), but no significant activation of ventral premotor cortex. The model involves reciprocal connections between regions F5 and AIP, and a projection from F2 to F5 – but the strength of the projection from F2 to F5 is essentially a free parameter of the model: there is a wide range of values over which the model will correctly perform the conditional and nonconditional tasks. The implication is that, by tuning this parameter, we can control this projection's contribution to the synaptic activity measure in F5. However, the original FARS model is such that difference in AIP synaptic activity from the nonconditional to the conditional task will always be less than the difference observed in F5. One possibility for repairing this problem in the model is to reroute the F2 information so that it enters the grasp decision circuitry through AIP, rather than F5.

The low-level details of the FARS model were derived primarily from neurophysiological results obtained in the monkey. The synthetic PET approach extracts measures of regional synaptic activity as the model performs a variety of tasks. These measures are then compared to rCBF (regional cerebral blood flow) observed during human PET experiments as the subjects perform tasks similar to those simulated in the model. In some cases, the human results provide confirmation of the model behavior. In other cases, where there is a mismatch between model prediction and human results, it is possible (as we have shown) to use these negative results to further refine and constrain the model and, on this basis, design new experiments for both primate neurophysiology and human brain imaging. Our point here is not to highlight (or hide!) the flaws in the present FARS model, but rather to suggest that it provides a useful platform for further modeling, and that synthetic PET provides a technique (itself open to fruitful modification) to ensure that the future modeling is responsive, and contributory, to future developments in both monkey neurophysiology and human neurology.

P5.4.3. Back to the "guided tour." Section 9.4 modeled the role of cerebellum in adapting a particular class of arm movements – the adaptation of dart throwing to wearing prisms. Section 10.5 examined the still somewhat discordant views on the roles of the basal ganglia in motor coordination and learning, emphasizing the view that the basal ganglia receives rich contextual information and then "releases" components of motor programs through disinhibition. We suggested that basal ganglia neuronal responses are task-dependent, with the current responsiveness of a neuron possibly determined by the state of behavioral experience or learning, and temporarily maintained until the relevant memory is formed in premotor cortices. We emphasized "channels" involved in working memory and in initiation of movement. The basal ganglia were seen, in part, as controlling a kind of "working memory" for coordinated control programs of motor schemas.

Roles for basal ganglia inhibition may include focusing (suppression of inappropriate movements), sequencing (suppression of forthcoming movement during preparation), and "simulation" in the sense of suppressing motor areas to allow activity in association cortex to be "disconnected," yielding covert "mental simulation" without immediate overt movement. We sought to distinguish the roles of basal ganglia and cerebellum, suggesting that basal ganglia are involved in explicitly combining the "pieces" that make up a skilled behavior, while the cerebellum serves to turn a procedure into a skill, adjusting parameters to adapt and coordinate components of the movement to yield a seamless whole.

P6. Dynamics

P6.1. Fixed points, rhythmogenesis and synchronization, and chaos

Chapter 4 stressed that neural systems can be studied at different levels, such as the molecular, membrane, cellular, synaptic, network, and system levels. Moreover, we noted

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

two main neurodynamical problems: study of the dynamics of activity spreading through a network with fixed wiring; and study of the dynamics of the connectivity of networks with modifiable synapses – both in normal ontogenetic development, and in learning as a network is tuned by experience. We introduced the key dynamical concept of an attractor, a pattern of activity which "captures" nearby states of an autonomous system. An attractor may be an equilibrium point, a limit cycle (oscillation), or a strange attractor (chaotic behavior). We also looked at the structure-function problem: for what overall patterns of connectivity will a network exhibit a particular temporal pattern of dynamic behavior? The results given were suggestive rather than directly applicable to biologically realistic models of neural networks. Section 4.5 introduced Hopfield networks to show work on neural networks motivated by statistical mechanics, including ideas of "energy," "temperature," and the statistical distribution of patterns in relation to an attractor-based model of pattern recognition – and then gave a critique of "computation with attractors."

Section 4.2 introduced the topic of oscillatory behavior in neural systems: single cell oscillations resulting from the interplay of a few currents; and central pattern generators (CPGs) in which a network of neurons can produce rhythmic behavior in the absence of sensory input. Bifurcation analyses were used to show a transition from equilibrium point to small amplitude oscillation, or from oscillation to chaos, as some control parameter passes through a critical value. We studied phase lags in chains of oscillators (mimicking data on the spinal cord of the lamprey), the importance of long-range coupling in the synchronization of more fully coupled networks (as in models of cortical structures), and bifurcation analysis of gait transitions in locomotion.

Section 4.3 studied chaotic behavior in the nervous system. Chaotic systems are characterized by sensitivity to initial conditions. We showed that the structural conditions of chaos occur at different hierarchical levels of neural organization. Neurochemical synaptic transmission is often characterized as a random process, but the "dripping faucet" model may be adapted to explain this apparent randomness as a case of deterministic chaos. We found that global cortical dynamics, as seen at the global level in the electroencephalogram (EEG), may also exhibit chaotic behavior. We discussed "dynamical diseases," introducing the diagnosis and also the control of chaos associated with normal and pathological brain functions. We also discussed the possible – controversial but intriguing – functional roles of chaos in normal brain activity, including perception and memory formation.

Section 5.2 focused on the dynamics of activity in the olfactory system, which we shall reiterate here in a fairly detailed way. The demonstration of oscillatory activity in olfactory systems (Adrian 1942; 1950) was one of the first experiments to illustrate stimulus-induced activity in the mammalian central nervous system. Phenomenologically, there are two main types of rhythmic activity in the olfactory system: slow and fast oscillations. Slow oscillations around 5 Hz may be imposed on the olfactory bulb by the respiratory nuclei (Freeman 1991), or may be induced in the olfactory cortex by cholinergic antagonists (Biedenbach 1966; for further details see sect. 5.2.3). The "respiratory wave" in the olfactory bulb is generated by the granule cells in response to input from the receptors through the periglomerular and mitral and tufted (M/T) cells. It can be detected in the mucosa by volume conduction. The slow background activity is phase locked with the respiratory wave and it is identified with the sniff cycle. A sniff cycle is composed of an inhalation and an exhalation, and its duration is 200–500 msec for rabbits. A slow potential evoked by odorants (Ottoson 1959) appeared also in the electro-olfactogram, a receptor potential recorded in the nasal mucosa, which can spread through the brain by volume conduction.

Electroencephalogram (EEG) patterns also show fast oscillations with frequency 35–90 Hz in different parts of the olfactory bulb (Freeman 1978, Freeman & Schneider 1982). (The terminology here is: slow ~1 Hz, intermediate ~10 Hz, and fast ~40 Hz.) It was suggested by the pioneering work of Freeman (1975) that the spatiotemporal activity patterns of the olfactory bulb can be interpreted within the framework of dynamical system theory. Later it was suggested (e.g., Skarda & Freeman 1987) that sensory information was encoded in spatiotemporal periodic and chaotic patterns. Generally, field potentials do not provide sufficient information about the underlying neural mechanisms. Still, one important message of Freeman's experimental and theoretical work is that rhythmic-arhythmic bulbar activity is the result of the interactions between excitatory (mitral) and inhibitory (granule) cell populations. The situation, however, might be more complex. Odor-induced (mitral cell) activity is under GABAergic control (Duchamp-Viret et al. 1993). Even in the case of blockade of the $GABA_A$ mediated inhibitory effect of the granule cells, oscillatory bulbar activity may occur as a consequence of recurrent excitatory connections.

Both network-level and detailed single neuron modeling techniques have been used to probe the structural bases of the generation of different rhythms and spatiotemporal patterns. First, a number of network studies are based on single-compartment models in which single cell activity is characterized by an internal state defined by the intracellular membrane potential, and by an output expressed as a firing frequency. Modeling illustrates how the interactions among excitatory mitral cell and inhibitory granule cell populations may generate oscillatory and more complex temporal patterns (Aradi et al. 1995; Érdi et al. 1993; Li & Hopfield 1989). More specifically, given the model anatomical structure and the related set of ordinary differential equations, the first question to be answered is what qualitative dynamical behavior emerges in the parameter space. The nature of attractors, the parameter windows belonging to them, and the bifurcation sequences are determined through systematic (numerical) studies. These results showed that chaos can only be found when there is sufficient lateral excitation. It has also been observed that all neurons oscillate in phase in each periodic region (and also during damping oscillation to a stable focus) and no wave phenomena have been detected in the parameter range studied here.

In other studies, mitral and granule cells were the subject of detailed single neuron modeling (Aradi & Érdi 1996; Bhalla & Bower 1993). Four types of problems of signal generation and propagation have been studied:

1. The effects of the individual currents and their role in the generation and suppression of action potentials, and in the control of firing frequencies (intracompartmental studies).

2. Signal propagation through the compartments of both the mitral and granule cells have been simulated. The effects of both orthodromic and antidromic stimulation have been demonstrated.

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

3. The excitatory-inhibitory coupling between the mitral and granule cells through dendro-dendritic synapses and the effects of the (partial) blockade of the GABAergic inhibition have been shown.

4. Dynamic behavior of a skeleton network of the bulbar circuitry taking into account even the periglomerular cells has been studied.

What kind of explanations can be obtained by using such techniques? One example of the dynamic effect of the self-excitation between mitral cells was that GABA antagonists produce prolonged depolarization in the mitral cells (Nicoll & Jahr 1982; Nowycky et al. 1981), and the reentrant excitation in the mitral layer may be associated to a particular mode of bulbar rhythmogenesis. In the case of two mitral cells connected by mutual excitatory couplings, synchronized burst activity may appear (see Fig. 5.10B in *Organization*). These simulation results are in accordance with the physiological findings and the suggestion that the blockage of GABAergic inhibition controls the odor-induced activity (Duchamp-Viret et al. 1993). According to the simulation experiments, both the mitral-granule feedback loop and the self-excitation in the mitral layer may provide the anatomical substrate of bulbar rhythmogenesis.

We also modeled rhythmic activity in the olfactory cortex. Similarly, section 6.2 gave a dynamical analysis of electrical activity patterns in the hippocampus, addressing data on the normal electrical activity patterns known as theta rhythms and sharp waves, and the abnormal electrical activity exhibited in epileptic seizures. We also compared the hippocampus with the olfactory system, with special attention to the neural mechanisms of rhythm generation and synchronization.

Thalamus and cortex are highly interconnected by reciprocal projections, giving rise to characteristic dynamic patterns. High frequency rhythms are associated with the waking state, low-frequency rhythms are associated with sleeping. Section 8.3 analyzed the balance between oscillations intrinsic to single neurons and network properties in the generation of thalamocortical oscillations. We analyzed the intrinsic electrophysiological properties of thalamic neurons, thalamocortical neurons, and reticular thalamic neurons, and then studied the dynamics of spindle oscillations, and of delta and slow sleep oscillations. We analyzed the role of brain stem control and cellular mechanisms in thalamocortical activation, and closed by using a number of models to explore the role of single cell dynamics versus emergent network properties.

P6.2. Self-organization: Modeling development

Both ontogenetic development of neural structures and their plastic behavior are often considered as dynamic processes in the state space of synaptic connections. The "self-organization" of the nervous system is in general a broader process, including addition as well as removal of synapses, and the modification of synaptic strengths. Self-organizing mechanisms are related to normal ontogenetic development (this subsection) and learning (see the next subsection).

Section 4.4 focused on retinotectal connections, discussing the following issues: specificity versus plasticity; genetically prespecified versus environmentally controlled wiring; marker theories versus activity-dependent mechanisms; decrease of synaptic strength by normalization rule only or by selective mechanisms; deterministic versus stochastic models; sets of discrete nerve cells versus continuous neural fields; and positional information.

Section 8.2 showed that modular architectonics may be seen as a pattern of organization resulting from the dynamics of self-organization rather than being completely laid down in the genome. In particular, the section provided models of the development of two examples of modular architectonics in primary visual cortex: ocular dominance columns and orientation columns. Section 10.3 discussed self-reorganization of the striatum, both the self-organizing character of the pattern formation for striatal compartments and the relationship between the modular remapping architecture, the tonic firing of certain striatal neurons, and their role in coordinated motor behavior.

P6.3. Plasticity: Modeling learning

Whatever the model of the individual neuron, neural tissues may be modeled as networks of intricately connected neurons in which strengths w_{ij} of the synaptic connections may themselves be described by differential (or difference) equations. These "learning rules" were introduced in section 4.5. These included Hebbian learning and its variations, which include means to avoid saturation of synaptic strengths, ways to accommodate various time delays, differential learning mechanisms, as well as "anti-Hebbian" rules to describe features of dissociations of patterns. We studied synaptic matrix models of associative memory, but also saw how invariant pattern recognition may be modeled using the dynamic link architecture in which Hebbian plasticity is invoked on a fast time scale.

Section 5.3 reviewed learning and plasticity in the olfactory bulb and olfactory cortex. Building on the relation of different attractor regions to different lateral connection strengths, we showed how synaptic modification can induce transitions between these regions. Another study of the olfactory bulb models associative memory, and shows that incomplete input patterns due to lower odor concentrations can also be identified as proper stimuli if a suitable learning rule is used to modify the lateral connections between mitral cells. We presented two "scenarios" for learning and memory in the olfactory cortex. One was based on the observation that the "sniffing rhythm" of 5 Hz may be optimal for inducing long-term potentiation (LTP) in olfactory cortex, and described a hierarchical clustering of input stimuli. The other was based on the argument that the mechanism of object recognition in the olfactory cortex is close to those offered by abstract associative memory models, emphasizing that the incoming (bulbar) information has a complex, distributed representation while the intrinsic excitatory connections between pyramidal cells are spatially extensive, overlapping, and modifiable.

Section 6.4 used the role of the hippocampus in the cognitive maps underlying navigation and spatial behavior in rats to ground a dynamic view of synaptic plasticity, showing how Hebbian-like plasticity may enable hippocampal cells to learn to encode different "places" in a cognitive map. We reviewed various neural network models of place cell training, allocentric location, and navigation, one of which pays special attention to data relating place cell activity to the theta rhythm, that is, relating the dynamics of rhythmogenesis to the synaptic dynamics of learning.

Section 8.5 studied the learning of coordinated behaviors, providing both a schema-level analysis of motor set and

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

the neuralization of coordinated control programs, and a specific neural network model of visual-motor conditional learning.

Section 9.3 focused on models of the cerebellum as a machine for learning motor skills, starting with the Marr-Albus model which views the Purkinje cell (the output cell of cerebellar cortex) as a perceptron, noting that data on long-term depression (LTD) support the Albus version of the model, namely that "coincidence" of climbing fiber and parallel fiber activity on a Purkinje cell depresses the efficacy of the synapses of parallel fibers active during the conjunction. We stressed the idea that cerebellar nuclei modulate motor pattern generators (MPGs) while the cerebellar cortex learns how best to modulate the cerebellar nuclei: modulating the modulator. Section 9.4 reviewed approaches to modeling adaptation of motor control where the adaptation persists on a long-term basis, involving synaptic plasticity (with the emphasis remaining on LTD of parallel fiber Purkinje cell synapses).

Section 10.6 argued that experience based on inferotemporal or prefrontal information may provide contextual, learned information to bias activity in the basal ganglia to "tip the balance" to one course of action or another. We presented models of visual-motor conditioning, including spatial generalization and sequential behavior based on the strong hypothesis that this learning is mediated by cortico-striatal plasticity which mediates a form of reinforcement learning in which dopamine released by the SNc acts as the reinforcement signal to toggle between Hebbian and anti-Hebbian learning.

P6.4. Compartmental modeling

Section 4.2 introduced some of the specific formalisms used to treat neurons and neural networks as dynamic systems. The framework for the detailed treatment of the dynamics of the membrane potential of a patch of neuron is provided by the neuronal cable equation and the Hodgkin-Huxley equation, and its relatives.

A whole neuron may either be modeled by a multicompartment model with compartments chosen to take into account the location of the entering synaptic currents or the geometry of dendritic branching, say, or as a single-compartment model characterized by a single membrane potential. The leaky integrator neuron is a popular model for the single-compartment case.

Section 8.4.1 provided the formalism for large-scale models of the nervous system used in many of our models which are based on simple (single-compartment) models of neurons. We now review cases where compartmental modeling has already yielded additional insights. Both periodic and chaotic temporal patterns can be generated at the single neuron level (sect. 4.3). Basic phenomena can be modeled with membrane equations involving two functionally distinct currents, the slow and fast currents, in which a series of complex patterned activities (simple slow oscillation, bursting, bursting-chaos, beating-chaos, and beating) can be generated by changing the time constant of inactivation of the slow current.

In section 5.2.2, in addition to modeling a network built from "integrate-and-fire" elements, multicompartmental models were given for the mitral and granule cells: six and four compartments were taken into account, respectively. This demonstrated specific effects of the individual ionic conductances on the overall performance of the compartment, signal propagation through the compartments, and synchronization in small networks. Section 5.2.3 presented the Wilson-Bower-Hasselmo model of temporal patterns in the piriform cortex which uses a five-compartment model for each pyramidal cell and explicit delays for transmission and axonal activity to clarify the assumptions leading to near 40 Hz cortical oscillations. By contrast, the Liljenstrom-Hasselmo model is designed to simulate modulatory cholinergic effects. Their network is built from relatively simple units whose output depends on a factor Q designed to represent the level of acetylcholine. Depending on the values of Q, the system may exhibit convergence to a fixed point, limit cycle oscillation, or (at least transient) chaotic behavior. Moreover, the strengths of the synaptic connections can also drastically influence the dynamic behavior.

Section 6.2 presented multicompartmental neuron models of pyramidal cells and interneurons of the CA3 region of hippocampus as a basis for the study of large networks of CA3 neurons in which we can see how variations in key parameters can switch the network between normal and epileptiform activity. Section 8.3 used a number of models to explore the role of single cell dynamics versus emergent network properties in thalamocortical oscillations. In section 9.2 we studied the simulation of a single Purkinje cell as a very detailed compartmental model with realistic ion conductances and synaptic currents in each compartment. Because this takes massive computing resources to simulate a single cell, the models of cerebellar function in sections 9.3 and 9.4 used simpler, single-compartmental models: but we pointed the way to future multilevel modeling which will relate system behavior to the fine details of neuronal function.

P6.5. Neurochemistry

Finally, we recall material assessing the biological grounding of learning rules used in the section on Plasticity: Modeling learning. Section 6.3 started by looking at one of the best-studied forms of dynamics at the synaptic level, namely long-term potentiation (LTP), showing its implication in experimental studies of Hebbian synaptic modification, and analyzing models of potentiation based on AMPA and NMDA receptors. We linked this back to dynamics at the activity level by studying the role of NMDA receptors in the generation of oscillations at the cellular level. We also discussed the need for long-term depression (LTD) in Hebbian synapses.

Section 9.4 modeled the cerebellar role in saccade adaptation, extending our view of LTD by stressing the notion of a "window of eligibility" to constrain the timing relation between the parallel fiber "context" and the climbing fiber "training signal." Section 10.6 studied cortico-striatal plasticity, positing a form of reinforcement learning in which dopamine released by the SNc acts as the reinforcement signal to toggle between Hebbian learning (positive reinforcement, LTP) and anti-Hebbian learning (negative reinforcement, LTD). We also suggested that shaping of the eligibility signal may be task dependent, setting an important goal for neuroscience to bridge from this systems level of neural analysis to that of synaptic neurochemistry.

P7. The hippocampus: A case study

It is generally agreed that the hippocampal formation has a crucial role in learning and memory processes. The hip-

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

pocampus is reciprocally connected to many neural centers and it is thought to prepare information for long-term storage. Moreover, the hippocampus has an important role in neurological diseases. Alzheimer's disease, epilepsy, and ischemia are associated with learning and memory impairment, and are accompanied by selective neuronal death or characteristic changes in the hippocampal circuitry. In this section, we review the discussion of the hippocampus provided in Chapter 6 of *Organization,* and make some additional remarks in the light of recent results.

P7.1. Levels, methods, problems

The hippocampus has been studied on different levels and by different methods:

1. The anatomical organization of the hippocampus including its afferent and efferent systems and the local circuitry of its components;
2. The electrical activity patterns related to global brain states and the underlying single-cell activities;
3. The cellular synaptic plasticity that occurs during long-term potentiation (LTP);
4. The role of the hippocampus of rats learning a spatial environment; and
5. The function of hippocampus in human memory.

Computational theories try to understand how the hippocampal neural circuitry and the whole cortico-hippocampal loop, supplemented with specific subcortical inputs, can implement different types of dynamic activity ("brain states") such as theta rhythms and sharp waves, and how these activity patterns elicit long-term potentiation (LTP). LTP is assumed to be the cellular basis for memory formation (Bliss & Lømo 1973). The relationship between the brain states and the enhancement of synaptic modifiability (LTP) can also be established by computational methods.

The functional view of the hippocampus related to navigation and memory phenomena should and could be unified with the structural approach by using dynamic computational models.

P7.2. Anatomical organization

P7.2.1. Global organization. The hippocampal formation is a cortical structure located in the temporal lobe. It is called archicortex for its evolutionary precedence over neocortex, and is relatively simple compared to neocortical structures. It has an elongated C-shaped form, and looks like a tube oriented perpendicular to the corpus callosum. Structurally the hippocampus is the simplest form of cortex, but this simplicity is in stark contrast to its role in processing information from the external world through the sensory systems, and from the "internal world" conveyed by subcortical inputs. Whereas, for example, primary visual cortex is specialized for processing a single modality, the hippocampus is functionally one of the most complex supramodal association areas, with many routes to many cortical areas. Polymodal association areas converge directly or indirectly on the entorhinal cortex which in turn forms the principal source of afferents to hippocampus. The hippocampus receives refined information from virtually all sensory modalities, both exteroceptive and interoceptive, via entorhinal cortex, and is thought to prepare information for long-term storage elsewhere in the cortex with the return projections from hippocampus possibly providing cells in polymodal cortex with a "condensed sketch" of the overall context in which their unimodal input occurred.

P7.2.2. Cell types. The *principal cells* of the dentate gyrus, the *granule cells,* generally do not have basal dendrites, but only have spiny apical dendrites. Their axons form the mossy fibers, which pass through the hilus (the area contained within the C formed by DG) before terminating on the dendrites of the CA3 pyramidal cells. The granule cell axons are considered to form excitatory synapses; the most likely neurotransmitter is glutamate. The hilus itself contains "polymorphic" cells, that is, cells of varied morphology. The principal cells of the hippocampus proper, the *pyramidal cells,* have thick apical dendrites extending through the stratum radiatum up to the stratum lacunosum-moleculare, and shorter and thinner basal dendrites which arborize in the stratum oriens. The thick, myelinated main axons of the CA3 pyramidal cells arising from the soma and terminating in the stratum radiatum and oriens of the CA1 region are the Schaffer collaterals. Furthermore, CA3 pyramidal cells have recurrent collaterals terminating in the CA3 region itself. Axons of the CA1 pyramidal cells are thin, and provide part of the hippocampal output, projecting mostly to the subiculum, and sometimes straight to the entorhinal cortex. Both CA3 and CA1 pyramidal cells also have collaterals that descend to the septal area via the fimbria. For most pyramidal neurons, glutamate is the (excitatory) neurotransmitter, which binds to (at least) three different receptor subtypes, metabotropic, AMPA, and NMDA. Recently metabotropic excitatory amino acid receptors have also been taken into account. Data for the differential distribution of three types of glutamate receptor have been reviewed by Somogyi et al. (1998). Hippocampal "nonpyramidal" *interneurons* exhibiting local inhibitory effects have a decisive role in controlling electrical activity. Freund and Buzsáki (1996) reviewed the anatomical, neurochemical and pharmacological, cellular and system physiological data and showed the diversity of interneurons. Certain types of local interneurons control the activity of the principal cells, while others may form a network, and collectively exert the inhibitory effect. Interneuronal networks may exhibit network oscillations with different frequencies, and they control the synchronized operation of the principal cells and the formation of plasticity.

P7.2.3. Circuitry. According to the today classical scenario (Andersen et al. 1971) there is a unidirectional cortico-hippocampo-cortical loop formed by the excitatory pathways. The perforant path originates in the entorhinal cortex and terminates in the granule cells of the dentate gyrus. The axons of the granule cells, the mossy fibers, project to the proximal part of the CA3 pyramidal cell dendrites. There is an extensive axonal arborization within the CA3 region. The axon collaterals of the CA3 pyramidal cells, the Schaffer collaterals, innervate the dendrites of the CA1 pyramidal cells, which further project to the subiculum and then to the entorhinal cortex. The anatomical organization is far more complex, having several other projections. The entorhinal cortex also innervates a subfield of the CA3 and CA1 regions. The local inhibitory cells can receive innervations from the principal cells of the same

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

(feedback) or afferent (feedforward) subfield. They can be innervated also by extrahippocampal afferents as well (Freund & Antal 1988; Freund et al. 1990; Gulyás et al. 1990). As we cited earlier (Somogyi et al. 1998), the basic principle of the organization of hippocampal circuitry seems to be known, even though many specific details are under clarification.

P7.2.4. Afferents and efferents. It is well established that the major input (the perforant path) to the hippocampus arises from layer II of entorhinal cortex (ENT). The ENT itself is considered as a relay for information coming from multimodal association areas in the temporal, prefrontal, cingulate, and insular regions. It seems likely that olfactory information is relayed through the lateral entorhinal cortex, while the medial entorhinal cortex conveys visual information. The former terminates in the outer third of the molecular layer, the postsynaptic targets being the distal dendritic field of the granule cells. The latter terminates on the middle third of the molecular layer. For a newer review of the morphological features of the entorhinal-hippocampal connections see Turner et al. (1998). Besides cortical (and commissural) connections, different subcortical structures are identified as hippocampal afferents and efferents. Subcortical inputs, in general, may strongly modify the hippocampal activity patterns. After having a model of the cortico-hippocampo-cortical loop, the specific effects of different inputs can be studied.

Another type of fiber is found to be GABAergic and to exclusively innervate inhibitory interneurons of the hippocampus proper (Freund & Antal 1988; Gulyás et al. 1990). Since the interneurons contain mostly GABA, as transmitter substance, the GABA-GABAergic interaction implements the phenomenon called disinhibition. Though the number of fibers producing disinhibition is relatively low, their modifying effects are still strong. The quantitative details are not known, and extensive simulation experiments are necessary to discover them.

The raphé nuclei of the midbrain area innervate the hippocampus. The main neurotransmitter of the raphé-hippocampal projections is serotonin. Specifically, the median raphé projections selectively innervate a subclass of interneurons in the CA regions (Freund et al. 1990), namely those containing calbindin, but not exclusively (Acsády et al. 1993).

One important output field of the hippocampus is the subiculum; other projections exist to the presubiculum, parasubiculum, and the entorhinal cortex. The subicular efferents to the deep layers of the entorhinal cortex close the multisynaptic entorhinal cortex-hippocampus-entorhinal cortex loop. Subiculum also generates a massive projection that travels in the fornix to the anterior thalamic nuclei and the mammillary bodies lying at the posterior edge of the hypothalamus. Deep layers of ENT are innervated by the hippocampus and project to neocortex, especially to zones neighboring ENT and to the medial frontal areas.

P7.3. Global brain states and behavioral states

P7.3.1. Electrical activity patterns. Global brain states, in both normal and pathological situations, may be associated with spontaneous activities of large populations of neurons. Experimentally, these activities may be detected by recording both from large neural assemblies (as in the EEG) or from a single neuron of the cell population. Generally, behavioral correlates can be defined for electrophysiologically global brain states.

Two main and normally occurring global hippocampal states are known: the rhythmic slow activity called the theta rhythm, and the irregular sharp waves (SPW) (Buzsáki 1989). A pathological brain state associated with epileptic seizures, an epileptiform pattern, is also characteristic of the hippocampus. More precisely, a set of different types of collective neural behaviors are qualified as "seizures." Both normal brain states and epileptic states are related to some (not clearly defined) synchronous activity. While a certain degree of synchronization is characteristic of normal rhythmic activity, highly synchronized cellular activity is more characteristic of clinical disorders. Other oscillations, such as a fast (40–100 Hz) gamma oscillation found mostly in the hilus and transient high-frequency (200 Hz) oscillation in the CA1 region, have also been reported.

P7.3.2. Theta rhythms. The theta rhythm is a population oscillation with large (~1 mV) amplitude and with 4–12 Hz frequency. Originally, the theta rhythm was found to occur whenever the animal engages in such behaviors as walking, exploration, or sensory scanning, as well as in REM sleep. O'Keefe and Nadel (1978) suggested that displacement movements – but not stationary voluntary movements (e.g., bar pressing at low speeds) – in the rat, coincide with theta; moreover, the frequency of theta has been found to correlate with speed of movement (O'Keefe & Recce 1993). It can also be phase-locked to sensory stimuli. Buzsáki et al. (1994) speculated on the double functional role of hippocampal theta rhythm. First, a large-scale oscillation in the entorhinal-hippocampal network induced by the septum is maintained by phase locking. Second, because the majority of the pyramidal cells are silent during theta and their membrane voltage is kept close to but below the threshold, relatively few excitatory synapses are sufficient to discharge them. In addition, theta is involved in LTP generation.

P7.3.3. Sharp waves. Sharp waves (SPWs) have a very large amplitude (up to 3.5 mV), their duration is 40–120 msec, and their frequency can be between 0.2 and 5 Hz. Though maximal SPW frequencies do overlap theta frequencies, theta waves are much more regular than SPWs. SPWs also have behavioral correlates: they occur during awake immobility, drinking, eating, face washing, grooming, and slow wave sleep. During SPWs, pyramidal and inhibitory cells fire with increased frequency. Furthermore, there is a partial synchronous cellular activity of both pyramidal and inhibitory neurons. The degree of synchrony is, however, under the threshold for induction of epileptic seizure.

P7.3.4. Synchronization. While theta rhythms depend on septal input, SPWs are formed by internal processes. One important precondition for SPW generation is the occurrence of a population burst in a small set of CA3 pyramidal cells. Their synchronization is mediated by excitatory synaptic connections.

Epileptic activity occurs in a population of neurons when the membrane potentials of the neurons are "abnormally" synchronized. As we already know, a certain degree of synchrony is necessary for normal theta and SPW behavior, and the transition between normal and abnormal degrees of

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

synchrony is not clear. Rather arbitrarily, activity has been considered epileptic if more than 25% of the cells fire during 100 msec (Traub et al. 1992). In vitro models of epilepsy (Traub & Miles 1991; 1992; Traub et al. 1987; 1992) offer a means to study the cellular mechanisms of the different types of epileptic phenomena by combined physiological and simulation methods.

Both experiments and theoretical studies suggest the existence of a general synchronization mechanism in the hippocampal CA3 region. Synaptic inhibition regulates the spread of firing of pyramidal neurons. Inhibition may be reduced by applying drugs to block (mostly) $GABA_A$ receptors. If inhibition falls below a critical level, complete synchrony occurs. Collective properties of networks of pyramidal cells modulated by inhibition have been studied successfully by Traub and Miles (1991).

There are ongoing debates about the origin of cortical gamma oscillation. Gray and McCormick (1996) suggested that the source of the gamma frequency "chatter" may be an intrinsic property of the cell. Very recently Wang (1999) gave an ionic conductance model of chattering neurons (in the neocortex), where the propagation of the action potential from the soma back to the dendrite is a key element of the rhythm generation. Gamma oscillation, however, may be the network property of interneurons connected by $GABA_A$ mediated inhibition (Traub et al. 1997; Wang & Buzsáki 1996; Whittington et al. 1995).

P7.3.5. Modeling rhythmic activity in the CA3 region of the hippocampus. Structure-based bottom-up modeling has two extreme alternatives, namely multicompartmental simulations, and simulation of networks composed of simple elements. There is an obvious trade-off between these two modeling strategies. The first method is appropriate to describe the activity patterns of single cells, small and moderately large networks based on data on detailed morphology and kinetics of voltage- and calcium-dependent ion channels. The second offers a computationally efficient method for simulating large network of neurons where the details of single cell properties are neglected.

Traub and Miles (1991) simulate hippocampal (mostly CA3) population activity by building "bottom-up" models from data on anatomic connectivities, ionic conductances, and synaptic properties. In most of their simulations the aim is to reproduce the results of physiological measurements made on hippocampal slices. Physiological measurements (both intracellular recording from one cell or, mostly, from a pair of cells, as well as field potential recording from a localized cell population) and simulations under various circumstances contribute to discovering the mechanism of both normal and pathological phenomena (e.g., epileptogenesis). Neurons in the Traub-Miles networks are modeled with a Hodgkin-Huxley formulation which has been modified in numerous ways.

Two types of action potentials can be generated in the CA3 pyramidal cell: (1) fast, sodium-mediated, localized mostly to the soma and (2) slow, calcium-mediated, mostly in the apical dendrite. The role of the potassium channels is, roughly speaking, repolarization.

The response of CA3 pyramidal cells to injected currents, namely the intrinsic burst discharges, are reproduced by the model. The frequency, even the regularity, of the action potentials depends on the strength of the applied current. A burst consists of a series of fast spikes at intervals of 5–10 msec terminating in one or more slower action potentials. The burst is called intrinsic, since isolated neurons can produce it. Some characteristic features of the physiological responses that were reproduced were (1) an intrinsic burst followed by a long after-hyperpolarization (AHP); (2) the dependence of bursting on the resting potential; (3) summation of spike after-depolarization to produce a depolarizing envelope; and (4) the ability to prevent full burst generation by properly timed hyperpolarizing input.

In *Organization* we offered only a few comments about population models. The description of a large population of neurons requires a different methodological approach, namely, the application of population theories. Just as collective phenomena emerging in physical systems made from a large number of elementary components (spins, molecules, etc.) are treated by statistical mechanics, so, analogously, have statistical dynamic theories of neural populations been established (Amari 1974; Ventriglia 1974; 1994; Wilson & Cowan 1973). These neuronal population theories used oversimplified single-cell models. One important example is the lack of ability to generate burst mode.

In the last couple of years, Érdi's group has developed a population theory of bursting (and nonbursting) neurons (Barna et al. 1998; Érdi et al. 1997; Grôbler et al. 1998) and applied it to simulating large-scale hippocampal activities. In this framework (1) the activity (different levels of subthreshold membrane potential/refractory state) distribution of groups of otherwise indistinguishable neurons is considered, and the subpopulations of neurons communicate via packets of impulses (action potentials) which they can emit and absorb; (2) neurons and impulses (action potentials) form two distinct populations; (3) the neurons, excitatory and inhibitory, occupy fixed positions in space, and their state is characterized by probability density functions over two continuous variables: their membrane potential and internal calcium concentration; (4) impulses can move from the point of emission (a neuron) to the point of absorption (another neuron) either by homogeneous spreading (random connectivity) or along prespecified paths (specific connectivity), carrying a quantum of excitation or inhibition (depending on the character of the emitting neuron). The absorption of impulses by a neuron implies: (a) change of the membrane potential; (b) firing of the neuron with a probability determined by the value of the membrane potential; and (c) emission of new impulses as a result of firing.

Distribution functions for the probable number of (excitatory and inhibitory) impulses and neurons, and also for neurons in refractory state provide a statistical description of the system. To take into account the actual connectivity structure of the system, a set of absorption coefficients and emission coefficients are given. These values define the strength and efficacy of the excitatory and inhibitory effects at each point of the neural system. Further parameters incorporated into the model give the possibility of taking into account other specific biological details such as impulse generation from external sources, spontaneous decay of subthreshold excitation, refractory period, synaptic delay, and so on. To evaluate and to visualize the simulation experiments, we use such macroscopic variables as the local density of impulses, the local mean net excitatory effect, and the local mean subthreshold excitation. Several normal epileptic activities, such as the synchronized population burst and synchronized synaptic potential (the analogue of

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

SPW in slices) and the propagation of the stimulus, have been simulated and the behavior of an "averaged" single neuron was also shown.

P7.4. Brain states and long-term potentiation

Long-term potentiation (LTP) was first discovered in the hippocampus and is very prominent there. LTP is an increase in synaptic strength that can be rapidly induced by brief periods of synaptic stimulation and which has been reported to last for hours in vitro, and for days and weeks in vivo. This time-course may be insufficient to sustain long-term memory, but there appear to be multiple LTP mechanisms, and one dependent on protein synthesis might serve long-term memory: inhibition of protein synthesis disrupts the maintenance of LTP, but leaves the induction of LTP relatively or totally intact. It is possible to relate properties and mechanisms of long-term synaptic plasticity in the mammalian brain to learning and memory.

There is now evidence for both homosynaptic and heterosynaptic LTP in area CA1 of the hippocampus, and an associative form of LTP has been reported in hippocampal CA1 and dentate gyrus. Hebbian synaptic modification depends on the co-occurrence of pre- and postsynaptic activity, and this effect was found in the form of LTP occurring in the Schaffer collateral/commissural synaptic input to the pyramidal neurons of hippocampal CA1 (Bliss & Collingridge 1993).

Buzsáki et al. (1994, p. 168) argue that sequential potentiation mechanisms

> ensure that discharge of a given set of entorhinal neurons during subsequent visits to the same part of [a] maze (recall) will reactivate the same subsets of neurons in CA3 and CA1. The hierarchy of neuronal firing during the SPW-associated bursts, therefore, is precisely determined by the recent past of the neural network. The rules of burst initiation and reconvergent excitation, subserved by the anatomical-physiological organization of the CA3 region, ensure that the synchronized events during consummatory behaviors and slow wave sleep carry biologically meaningful information.

P7.5. Hippocampal function, cognitive maps, human memory

The functional view of the hippocampus – its role in the cognitive maps underlying navigation and spatial behavior in rats – should be combined with a dynamic view of synaptic plasticity, because Hebbian-like plasticity may enable hippocampal cells to learn to encode different "places" in a cognitive map.

P7.5.1. Place cells, navigation. Rats are highly exploratory. In a new environment, they tend first to explore outward from some base, then to shift to other bases until they become highly adept at navigating from one place to another, visiting sites where food has been taken, and returning to inaccessible hiding places. Rats entering one arm of a T-maze will tend to choose the other arm on the next exposure ("spontaneous alternation"). A landmark is not merely a stimulus to be approached for a reward. Rats remember headings relative to the landmark, and can use the position of a number of objects to navigate toward, for example, a food source or hiding place. In the "water maze" (Morris 1984), a rat can use such cues to swim to a platform located beneath opaque water.

Certain pyramidal cells of the CA1 and CA3 regions fire when the rat moves to a particular place in the environment, and these cell are called "place cells." There are some suggestions for how learning leads to the appearance of stable, bounded place fields as a result of exploratory behavior. Jensen and Lisman (1996) and Wallenstein and Hasselmo (1997) assume that the dominating factor of the information exploited during learning is consistency in the sequence of perceived sensory input. It is assumed that the important thing is not what the rat actually sees, smells, or touches, but that a particular input pattern is always followed by a (different) particular input pattern. According to this idea, place cells are sensitive to subsequences in the whole input stream, which we observe as spatial sensitivity due to the spatiotemporal continuity of the rat's movement. Models based on this temporal correlation in the input stream can reproduce many important features of place fields. However, they cannot account for the symmetric graded firing profile within a place field. Recently, we gave a learning-rule based model (Érdi et al. 1998; Szatmáry et al., submitted) to overcome the failure of previous models.

Experiments suggest that rats (1) have associative memory for complex stimulus configurations, (2) can encode the spatial effect of their own movements, and (3) are able to form sequences of actions to go from a starting location to a goal. In other words they have a cognitive map (Tolman 1932). The hippocampus may function as part of a local navigation system (i.e., the cognitive map).

Several neurobiologically (more or less) realistic models and algorithms have been suggested to solve the problem of orientation and navigation based on information obtained from place cell firing (Burgess & O'Keefe 1996; Burgess et al. 1994; Gerstner & Abbott 1997; Zhang et al. 1998). Burgess et al. (1994) constructed a multilayered network of different functionally defined cells (entorhinal, place, subicular, head-direction, and goal cells) supplemented with layer-specific activity and learning dynamics. Relevant neurophysiological phenomena (theta rhythm, phase coding, place fields) are incorporated into the model.

P7.5.2. Affordances, motivation, and the world graph theory. O'Keefe and Nadel (1978) distinguish two paradigms for navigation, the "locale system" for map-based navigation and the "taxon (behavioral orientation) system" for route navigation. Guazzelli et al. (1998) model both the taxon system and the map-based system, as well as their interaction; and they argue that the map-based system involves the interaction of hippocampus with other systems, not just the hippocampus alone. They relate *taxis* (movement toward some goal, as in phototaxis) to the notion of an *affordance* (a visual indication of a course of action, Gibson 1966; already presented for grasping in sect. P5.4). Just as a rat may have basic taxes (plural of taxis) for approaching food or avoiding a bright light, so does it have a wider repertoire of affordances for possible actions associated with immediate sensing of its environment. We propose that affordances are extracted by the rat posterior parietal cortex, which guides action selection by the premotor cortex and is also influenced by hypothalamic drive information. The Taxon-Affordances Model (TAM) for taxon-based determination of movement direction is based on models of frog detour behavior, plus expectations of future reward implemented using reinforcement learning. The specification of the direction of movement is refined by current affordances

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

and motivational information to yield an appropriate course of action. The World Graph (WG) theory (Lieblich & Arbib 1982) expanded the idea of a map by developing the hypothesis that cognitive and motivational states interact. Guazzelli et al. (1998) developed an integrated TAM-WG model which explains data on the behavior of rats with and without fornix lesions which disconnect the hippocampus from other neural systems.

P7.5.3. Memory systems. The role of hippocampal function in human memory is discussed in *Organization* by introducing the crucial dichotomies of procedural versus declarative memory and of skill versus episodic learning. The data suggest that human hippocampus is involved in declarative rather than procedural memory and in episodic rather than skill learning. It is also likely that hippocampus may form the memory traces but not store them indefinitely.

There are ongoing debates on the role of the hippocampus in episodic and declarative memories. Varga-Khadem et al. (1997; Mishkin et al. 1998) came out with the novel idea that the hippocampus might play a selective role in episodic memory, while the related cortical structures might support semantic memory even in the absence of hippocampal function. While Tulving and Markowitsch (1998) support the new proposal, Squire and Zola (1998) do not see sufficient proofs that episodic and semantic memory are differently affected in amnesia.

Although the hippocampus stores information, it seems that it also "installs" this processed information elsewhere in cerebral cortex for long-term availability. Buzsáki (1989) suggested an informal model of memory formation in which cortical information is processed by two stages. First, during the theta brain state, cortical activity weakly potentiates, via the granule cells, the CA3 pyramidal cells associated to a labile form of memory trace. This weak potentiation initiates population bursts implying a transition from theta to SPW state. Under the SPW state, excitatory synapses between pyramidal cells both within the CA3 region and between CA3 and CA1 regions are enhanced. These enhanced synapses would be the substrate of a long-lasting memory trace. Since SPW and associated high-frequency oscillation in the CA1 region yields discharge of neurons of deep layers of entorhinal cortex, it seems likely that hippocampal output may affect other neocortical targets, transferring information stored temporarily in the CA3 region to the neocortex for long-term storage.

This transfer may occur, at least in part, during sleep. Pavlides and Winson (1989) showed that hippocampal cells active during a waking period exhibit increased firing rates in the following sleep period. To investigate this effect in more detail, Wilson and McNaughton (1994) monitored the simultaneous activity of 50 to 100 CA1 cells during a running period (RUN) and during both the prebehavioral (PRE) and postbehavioral (POST) sleep periods. During the RUN period, cells with overlapping place fields exhibited highly correlated activity; those with nonoverlapping fields did not. Indeed, cells that were coactive during the RUN period showed a far greater correlation than during the PRE period. Moreover, this correlation was reactivated during the POST period but declined with a time constant of approximately 12 minutes. Wilson and McNaughton see this as support for the hypothesis that hippocampal activity during sleep exhibits a reactivation of population activity from the prior waking period. Since CA1 has little direct connectivity between pyramidal cells, they suggest the correlations arise in CA3 (which has many intrinsic connections) or entorhinal cortex.

In support of the idea that information is transferred from hippocampus to neocortex especially during the synchronized bursts ("ripples") of sharp wave (SPW) activity (Buzsáki 1989), Wilson and McNaughton (1994) found during the POST period that correlations during ripples were significantly greater than the correlations in the periods between ripples. Chrobak and Buzsáki (1996) have shown that SPWs and ripples are initiated in CA3 and that the output layers – but not the input layers – of entorhinal cortex exhibit neuronal activity correlated with CA1 SPWs. Wilson and McNaughton thus suggest that the induced correlations during SPWs arise from modifications within the hippocampus and are propagated to the output layers of entorhinal cortex.

P7.5.4. Functional imaging of the human hippocampus. The last few years have seen a tremendous development in brain imaging, including that related to the hippocampus (see, e.g., the recent issue of *Hippocampus* 9(1), 1999). Stern and Hasselmo (1999) integrated cellular and fMRI studies, while Horwitz et al. (1999) showed that there are only a few large-scale neuronal models relating PET data to neuronal activity. Much remains to be done. Functional imaging is certainly underrepresented in *Organization.* However, we offer our own work on synthetic PET (Arbib et al. 1994) and the population models developed in Érdi's group (Barna et al. 1998; Grôbler et al. 1998) as two steps toward a technique to understand the results of brain imaging in terms of detailed neural activity.

P8. Toward a cognitive neuroscience

Our studies of structure, function, and dynamics are located in a broad sweep that runs all the way from the motility of the embryo to the learning of visually guided behavior. In this concluding section we turn from retrospect to prospect, suggesting ways in which the ideas developed in *Organization* may contribute to future work in cognitive neuroscience. This prospectus embodies a strong philosophical position, namely that mind (at least that aspect of it known as cognition) can be explained in terms of the workings of matter (especially that structured as neural systems). This raises a methodological challenge since the categories of "mind talk" (function) and "brain talk" (structure) do not map directly one on to the other. *Organization* has provided a framework in which the study of structure and function may be integrated with dynamics. In doing so, we have reflected on the immense progress neuroscience has made in delimiting structure, whether the functional neuroanatomy that has, for example, used double labeling techniques to subdivide and chart the *terra* hitherto *incognita* of the primate association cortices, or the studies in neurochemistry and molecular neurobiology that reveal more and finer structures within the individual neuron. We have also seen conceptual advances in the study of large networks of (somewhat simplified) neurons, ranging from studies of low-level vision to the statistical mechanics of self-organization which emphasize the matching of a single function to a single network. Chapter 3 of *Organization* emphasized work at a different level, in which a network of

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

functions (schemas and schema assemblages) must be mapped to a network of neural networks. We now discuss the implications of this for cognitive neuroscience.

P8.1. Memory, perception, and intelligence

Among the properties that contribute importantly to intelligence are the following:

1. Possession of a modifiable model of the world, with its attendant adaptability: A system to act intelligently must not only be able to take properties of its environment into account, but must be able to update its record of these properties to take account of new observations and changing relationships.
2. Flexibility and generality: An intelligent system must not only use past experience to act adaptively, but must also be able to apply its past experience to situations that are not superficially similar to those encountered before. Again, techniques that have been developed to solve one type of problem should be recognized as applicable even when a very different domain of problems is involved.
3. Dynamic planning: An intelligent system should use its model to plan and evaluate alternative courses of action before committing itself to one of them. For a symbol-manipulation system there may be little real distinction between planning and action, but for a robot or an animal the distinction is very real and very important: it pays to recognize a precipice in advance and plan to avoid it rather than recognizing one's mistake after going over. However, it is crucial that the plan be dynamic in that it is rapidly and effectively updatable when new data reveal unexpected obstacles or make sought-for information available.

One important form of working memory is obtained by holding a particular pattern of firing during a delay task. Such neurons have been found in the dorsolateral prefrontal cortex as well as the hippocampus. What distinguishes these two systems? The answer is still far from clear, but we assert that a full analysis of the procedural/declarative distinction in humans requires a theory of consciousness that distinguishes conscious/declarative from nonconscious/procedural access to schemas. In schema-theoretic terms: the perception of a situation or the carrying out of a particular action requires the construction of a particular schema-assemblage (assimilation, in Piaget's term); to the extent that the perception or action is problematic, the schemas may become modified to increase the chance of success in similar situations in the future (accommodation). We learn both by storing schema-assemblages (memory of a specific situation and course of behavior) and by tuning extant schemas. However, it is too facile to say that the former corresponds to "fact memory" and the latter to "skill memory," because "skill memory" also partakes of some aspects of assemblage-formation. Skills are tuned versions of "programs" constructed from prior schemas, rather than only the tuning of single previously extant schema. The fact that a schema may be activated without conscious awareness emphasizes the notion that different neural processes must be involved in monitoring the use of a schema as distinct from the use per se of the schema.

The "what"/"how" distinction reviewed in Chapter 8 – one patient may be able to "declare" the size of an object yet not be able to preshape the hand appropriately to grasp it; another patient may exhibit the opposite symptoms – shows that some schemas are instantiated on paths to conscious awareness, and others are not. Moreover, some at least of the "working memory" systems of prefrontal cortex are tightly coupled to specialized areas of parietal cortex, and are thus tightly integrated into the procedural "how" system rather than the conscious/declarative "what" system. So: the loop of explanation must be closed back from the theory of consciousness. This is consistent with Rozin's (1976) view that procedural learning may be phylogenetically old, having developed as a collection of *encapsulated special-purpose abilities of specific neural systems* to register cumulative changes in their functioning. By contrast, the capacity for declarative learning reaches its full development only with the elaboration of medial temporal areas in mammals, especially the hippocampus and related cortical areas. *Organization* offers some relevant material in the discussion of Hippocampal Function and Human Memory in section 6.5.

Perception provides access to motor schemas to control interaction with the object, but this does not necessarily entail execution of even one of these motor schemas. Although an animal may perceive many aspects of its environment, only a few of these can at any time become primary loci of interaction. Planning is the process whereby the system combines an array of relevant knowledge to determine a course of action suited to current goals. In its fullest subtlety, planning can involve the refinement of knowledge structures and goal structures, as well as action per se. Novel inputs (e.g., coming upon an unexpected obstacle) can alter the elaboration of high-level structures into lower-level tests and actions which in turn call upon the interaction of motor and sensory systems (cf. the notion of dynamic planning). We seek to study schemas that are part of the internal state of the system and can flexibly guide ongoing action in terms of internal goals or drives and external circumstances. Note that we do not imply that planning is a conscious process, but do argue that planning goes beyond mere choice. In "choice," we suggest that a decision (whether conscious or not) must be made between a few clearly delimited alternatives. In planning, by contrast, solutions to many, possibly conflicting, subproblems will have to be constructed to yield a possibly quite novel course of action. Section 8.6, From Action-oriented perception to cognition, provides an evolutionary view of how visual perception may evolve into a distributed capability for planning.

An important gap in most computational analyses of the mind comes about because few neuroscientists think about the social nature of being a human (see Brothers 1997 for an entry point to that small literature which does begin to link neural activity to social cognition). To be human is not just to have a sophisticated "computer" in the head called the brain. It is also to have grown up as a member of society, and to have learned the nuances of that society. Neuroscientists and cognitive scientists emphasize what can be measured objectively, such as language where we analyze a string of symbols, or vision where there are particular patterns to which we can see how people or animals or neurons respond. Arbib's (1985) *In search of the person* emphasized that much of human experience, or, if you will, person-reality, of being a member of society, being aware, and having experience of love, hate, and anguish is normally not addressed at all within the framework of brain research or cognitive science. The point was not (yet?) to reduce these to current brain theory or cognitive science, but

rather to show how the science and the personal experience might be thought about in a unified framework in which understanding of each reality could come to shape that of the other.

P8.2. Neural-cognitive interaction: Self-organization, constructivism, downward causation, hermeneutics

Our intention was to write a "neural book" and to show how brain components are organized to implement "higher brain functioning." Still, we cannot avoid a brief mention of some philosophical issues. John Szentágothai (1993) argued that the essence of the neural is to be found in its self-organizing character. If all neural functions, including those from the simplest elementary reflexes to complex global functions of the whole organism have at their very basis spontaneous activities arising – in part at random, in part constrained by the genome – in individual nerve cells and, if all neural functions are integrated by self-organization into various activity patterns, our whole understanding of neural organization has to undergo rather fundamental changes. If the reflex paradigm of neural systems is to be abandoned for the new concept of "self-organization" of spontaneous (random or other) activity, this would be an entirely new challenge for "brain-mind philosophy."

In a recent *BBS* target article, the idea of constructionism (Quartz & Sejnowski 1997) emerged, providing arguments for the necessity of the interaction between cognitive and neural levels of description. The statement that cognitive level learning influences brain development may be associated to the argument on the problem of "downward causation." Downward causation – that is, the notion that mental agents can influence neural functioning – was suggested by Sperry (1969). It was not clear, however, how conscious processes directly influence physiological mechanisms in Sperry's scheme. Szentágothai (1984) was more cautious. He suggested that the nervous system can be considered as being open to various kinds of information and that there would be no valid scientific reason to deny the existence of downward causation (more precisely, a two-way causal relationship between brain and mind). Indeed, Érdi (1996) argued that the philosophical tradition of hermeneutics, that is, the "art of interpretation," which is a priori neither monist nor dualist, can be applied to the brain. Playing with the idea that the "device approach" to the brain and the philosophical approach can be reconciled, he concluded that the brain is a physical structure that is controlled and also controls, learns and teaches, processes and creates information, recognizes and generates patterns, and organizes its environment and is organized by it. It is an "object" of interpretation, but also it is itself an interpreter. The brain not only perceives but also creates new reality: it is a hermeneutic device. In *The metaphorical brain 2,* Arbib (1989) argued that our theories of the brain are metaphors, while the brain itself represents the world through schemas, which may themselves be viewed as metaphors (and see Arbib & Hesse 1986 for the role of hermeneutics in a schema-based epistemology that links the individual and the social, especially the notion of two-way reduction).

P8.3. Language

To end on a more concrete note, we discuss briefly the controversial issue of whether or not the study of language can be illuminated by approaches to neural organization of the kind presented in *Organization.*

Connectionist modelers distinguish two main architectures for their networks. In a layered feedforward network, the "neurons" are arranged in a series of layers, with the only connections being from neurons in one layer to neurons in the next layer. Because there are no loops, there is no possibility of a "reverberating memory," and thus, after a suitable propagation delay, each input pattern yields a unique output pattern. By contrast, a network with loops acts as a dynamic system: for each fixed input pattern, patterns of activity may move around the network, creating dynamic sequences of internal states. In many studies, the behavior of such a network is not characterized in terms of input-output pairs, but rather in terms of settling into a "steady state" such as a *point attractor,* a *limit cycle* (yielding a sustained oscillation), or even a *strange attractor* (deterministic chaos). Networks studied from this point of view are thus often referred to as *attractor networks* (cf. *Organization,* sect. 4.5.4, Computation with attractors: Scope, limits and extensions). Studies of lesioned attractor networks (e.g., Hinton & Shallice 1991; Plaut & Shallice 1993) provide further insights into the clinical observations of neuropsychology. Simulated damage to attractor networks can qualitatively mimic some of the deficits seen following human brain damage. In particular, such studies suggest how there may be an appearance of functional modularity (i.e., two functions may be differentially impaired by network damage) even when the functions are implemented by a single network.

Despite its real contributions, such studies make little progress in explaining the contributions of specific brain regions to language capabilities. Recent work on "mirror" neurons (di Pellegrino et al. 1992) provides promise in this direction by suggesting a new path for integration between the study of human language and the study of detailed neural mechanisms of visuomotor coordination. These mirror neurons are a subset of the grasp-related neurons of the F5 region of premotor cortex (sect. 8.4.3). They discharge when the monkey observes meaningful hand movements made by the experimenter, such as placing or taking away objects from a table, grasping food from another experimenter, or manipulating objects. There is always a link between the effective observed movement for a mirror neuron and the effective executed movement.

These data suggest that area F5 is endowed with an observation/execution matching system and led Rizzolatti et al. (1996b) to seek evidence for an observation/execution matching system in humans. In a PET study of brain activation of humans observing hand gestures, they found a highly significant activation of the posterior part of the left inferior frontal gyrus – the rostral part of Broca's area! While homologies between cortical areas of different species are always difficult, a good case can be made that Broca's area is in part homologous with F5. These data led Rizzolatti et al. (1996a) to a bold hypothesis: namely that the functional specialization of human Broca's area derives from an ancient mechanism related to production and understanding of motor acts. To this we would add that this specialization may correspond to verbs or verb phrases, but seems separate from the functions of naming and of noun phrases more generally – concordant with our emphasis that language, like other functions studied in this volume, is to be seen more as a distributed function ("cooperative

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

computation") rather than being a "unitary faculty." Rizzolatti et al. argued more specifically that the sophisticated capacity of action analysis shown by mirror cells is at the basis of the evolutionary prevalence of the lateral motor system over the medial (emotion-related) one in becoming the main communication channel in higher primates and man. Much work is currently under way (e.g., Rizzolatti & Arbib 1998) to turn this hypothesis into a rigorous neurolinguistic model subject to coherent testing that integrates monkey neurophysiology and human brain mapping within a framework offered by the current debate over language evolution (Pinker & Bloom 1990; Wilkins & Wakefield 1995).

The point for our current claim – that *Organization* provides powerful tools for Cognitive Neuroscience – is that we see here an approach to language which does not treat it in grand isolation in the style of Chomsky (1995 onward), but instead (without denying the special character of these higher mental functions) sees language and other cognitive processes within the framework of neural organization in general that we have charted with John Szentágothai.

ACKNOWLEDGMENT

This work was supported by the OTKA grants T025472 and T25500 (P.É.), and by the Human Brain Project under P20 Program Project Grant HBP: 5-P20-52194 (M.A.A.).

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

The organization of *Organization:* Neuronal scaffold or cognitive straitjacket?

A. J. Amos & C. D. L. Wynne

Department of Psychology, University of Western Australia, Perth, WA 6009, Australia. **andrews; clive@freud.psy.uwa.edu.au**

Abstract: We praise Arbib et al.'s *Neural organization* for its support of the integration of different levels of analysis, while noting that it does not always achieve what it advocates. We extend this approach into an area of neuropsychological activity in need of the structure offered by *Organization* at the intersection of the conflated fields of executive function and frontal lobe function.

Traditional views of the relationship of brain to behavior have fallen into either top-down or bottom-up camps. Top-down theorists (e.g., Staddon 1999) argue that, since a major new discovery in basic neuroscience may always be just around the corner, the behavioral theorist is better off explaining behavior in its own terms and leaving it to the neuroscientists to figure out how these behavioral laws are implemented in the brain at some future date. Bottom-up theorists (e.g., Koch & Crick 1994) believe that when the last piece of the jigsaw of neuronal activity fits into place, the behavioral and cognitive questions will simply solve themselves. Top-down theorists take comfort every time accepted wisdom in neuroscience is overthrown (such as the recent demonstration that the primate cortex continues to make new nerve cells postpartum Gould et al. 1999), but are vulnerable every time a useful model of behavior is built from quasineural components. Bottom-up theorists are vulnerable to the observation that not all the features of neural units are relevant to the behavior of neural networks, and that such networks have emergent properties that could not be predicted from the action of the individual elements.

Neural Organization (henceforth *Organization*) is a curse on both these houses. Arbib et al. argue persuasively for a vertically integrated approach that pursues bottom-up, top-down (as well as "middle-outwards") directions of argument simultaneously. In the long run they must surely be right – the only question is whether the state of the component sciences is ready yet. At times the gaps between the levels are still all too evident even in their own work (e.g., in their discussion of the hippocampus, Ch. 6). This caveat notwithstanding, we believe that the organization proposed in *Organization* is very promising and our remaining comments illustrate how useful we have found that structure in conceptualizing our own research, past and future.

An area of neuropsychological activity greatly in need of the structure offered by *Organization* is the intersection of the conflated fields of executive function and frontal lobe function. It has often been assumed that the study of executive function is isomorphic with the study of frontal lobe function because many of the tests used to measure executive function have been poorly performed by patients with frontal lobe lesions (Stuss et al. 1994). *Organization* suggests that the levels of structure and function be specified separately until evidence motivates their integration. Recent computational work illustrates the utility of this approach.

Levine and Prueitt (1989) and Amos (in press) described putative mechanisms in normal and impaired frontal lobe function with computational models of the Wisconsin Card Sorting Task (WCST). Levine and Prueitt's (1989) top-down approach specified the model in terms of functions performed while carrying out the task, while Amos (in press) based his model on the neuroanatomy of the frontal lobe, basal ganglia, and thalamus – a bottom-up approach.

Levine and Prueitt's (1989) model simulated normal and frontal patient performance on the WCST. However, because it was specified at a functional level, it could not incorporate evidence that subjects with subcortical damage also perform poorly on the WCST. The model based on neuroanatomical information relevant to diseases of subcortical structures could accommodate this data by modifying parameters analogous to dysfunction at the neuronal level (Amos, in press).

Organization provides a structure clearly lacking in both papers. An appreciation of the importance of integrating different levels of analysis would have allowed either the Levine and Prueitt (1989) or the Amos (in press) model to achieve a more complete and compelling account of function and dysfunction in WCST performance. It might also allow for the disambiguation of the various executive functions, and a more systematic mapping of these functions to diverse brain areas. There exists abundant evidence to completely specify the WCST at a functional level and to completely specify possible anatomical substrates of the test. Despite this, Levine and Prueitt (1989) made claims about the anatomical substrate of their functionally defined model, and Amos (in press) made claims about the functional level of his anatomically defined model, without fully specifying the domains of knowledge within which they were working. In this they were certainly following the approach taken by most of the papers published on executive functions and the cognitive functions of the frontal lobes; because of this they may simply have repeated the common error of conflating two levels of analysis, obscuring the functional and structural insights that disambiguation might afford.

Thus we contend that the confused state of research into executive function, and frontal lobe function, could be greatly relieved by sensitivity to different levels of analysis. *Organization* provides a structure that could be used to motivate a principled integration of multiple levels of neuropsychological data, models, and theories. The structure both clarifies previous endeavors and identifies

valuable avenues for future investigation. This would be most useful in the debate on frontal/executive function, but it could be fruitfully applied to all areas of neuropsychological research.

In summary, it appears to us that *Organization* both blueprints a scaffold upon which to erect a stronger understanding of the brain, and cautions against the premature integration of levels which can unprofitably constrain scientific endeavor.

From reductionism to reductionism

Fred L. Bookstein

Center for Human Growth and Development, University of Michigan, Ann Arbor, MI 48109. **fred@brainmap.med.umich.edu**

Abstract: *Neural organization* attempts to thwart, at least in part, modern neuroscientists' tendency to focus reductionistically on ever smaller microsystems. But although emphasizing higher levels of systems organization, the authors end up enforcing reductionisms of their own, principally the reduction of their domain to the study of invariable normal functioning, without explicit modeling of the deviations that constitute disease states or aging. This reductionism seriously weakens the authors' claims about the truth of their quantitative models.

Like any other conscientious systems scientists, Arbib et al. are concerned about reductionism. They write on page *ix:*

> A "good" model is responsive to available data; an "interesting" set of data will test hypotheses that are theory laden. . . . Much of modern neuroscience seems to us excessively reductionist, focusing on the study of ever smaller microsystems to the exclusion of an appreciation of their contribution to the behaving organism. We do not reject the data gained in this way but are concerned with restoring some equilibrium between systems neuroscience, cellular neuroscience, and molecular neuroscience.

Now it is entirely appropriate to wish "equilibrium" restored, but that is not the only pertinent reductionism bedeviling this domain. After all, the neurosciences arose in neurology, and applications there still supply their primary fiscal apologia. Then in eschewing any concern for disease states, the authors are inadvertently no less reductionist than the level-specific approaches they would supersede. The reader finds no explanations of disease states that follow from predictable behaviors of the models developed here, even though such consequences are hinted at. For instance (p. 309), in Parkinson's disease, "output from GPi may increase from 80 Hz to 100 Hz, and this may account for the hypokinesia." But the text does not supply any argument pertinent to this "account." The few other diseases mentioned at all are treated just as curtly: the claimed "accounts" are neither argued nor figured.

Thus the authors' rejection of reductionism appears somewhat inconsistent. Inasmuch as their work remains overwhelmingly normative, concerned with "healthy" or otherwise typical behaviors of the networks they study, the principal variation that is studied can only be variation over (normal) developmental time. The discussion of the visual cortex in section 8.2, for instance, which links "two quite different themes of the book: modular architectonics and the dynamics of development" (p. 222), does not include any sort of natural variability in its subject matter. While the explicit modeling includes both environmental noise and an experimentally induced disease state (monocular vision), nevertheless I cannot find any match between quantifications of the simulation and quantifications of the claimed system being simulated. There results a novel reductionism all its own, from the actual neurohistologic or neuroelectrical data to the meager qualitative summaries (textures, catastrophes) that seem to match equally meager qualitative reports of the behavior of their models.

Thus a book that begins with a clarion call to avoid reductionism displays its own reductions nearly as troubling as those it takes pains to circumvent. There is the reduction of the subject matter to normative (healthy) behaviors, and the more admirable of those behaviors at that – nothing here about rage, lust, or the hallucinations of schizophrenia – yet there is also the reduction of both model behavior and data analysis to the most superficial qualitative descriptions. For a book that claims to be concerned with "the behaving organism," this disregard of quantitative aspects of that behavior, as well as quantitative aspects of the fit between the authors' models and that behavior, is unfortunate. In effect the book asks only, "How are nervous systems possible?" without inquiring further as to differences among the different nervous systems that are all evidently equally possible: the nervous systems that fill our psychiatric wards right alongside our neuroscience laboratory chairs.

This may be a collective decision on the part of a community of the like-minded, what Ludwik Fleck called *Denkkollektiv* (Fleck 1979). But if a good model is "responsive to available data" and interesting data test "hypotheses that are theory-laden" (p. *ix*), it is no less important that the data match the model than that the hypothesis be "theory-laden"; and such theories ought to deal with a continuum of behaviors, not just one. I hope that the authors' reductionism in this sense can be overturned via a determined balancing of the study of normal processes and normal development with the equally important study of the abnormal. It is the latter study, after all, that pays the bills.

Encyclopedia of computational neuroscience: The end of the second millennium

Roman Borisyuk

Centre for Neural and Adaptive Systems, School of Computing, University of Plymouth, Plymouth, PL4 8AA, United Kingdom. **borisyuk@soc.plym.ac.uk**
www.tech.plym.ac.uk/soc/research/neural

Abstract: Arbib et al. describe mathematical and computational models in neuroscience as well as neuroanatomy and neurophysiology of several important brain structures. This is a useful guide to mathematical and computational modelling of the structure and function of nervous system. The book highlights the need to develop a theory of brain functioning, and it offers some useful approaches and concepts.

Arbib, Érdi & Szentágothai unite under the same cover a brilliant and nicely illustrated introduction into neuroanatomy/neurophysiology and an encyclopedia of mathematical/computational models in neuroscience. The book contains two parts; Part I gives three overviews: structural, functional, and dynamical and Part II describes the interacting systems in the brain: olfactory, hippocampus, thalamus, cerebral cortex, cerebellum, and basal ganglia.

Consideration of neural organisation in the book is based on the triad of structure, function, and dynamics and the authors consider each of these subjects separately in detail. There is an important philosophical question about the relationships among structure, function, and dynamics in this triad. In the 1970s, my supervisor, Professor Albert Molchanov, said that a biological "function" is the result of the kinetic/dynamical development/realisation of "structure" and that today's structure is a consequence/fixation of yesterday's kinetics (Molchanov 1967).[1] These relations might be realised on different hierarchical levels. Figure 1 gives a graphical representation of the relations in the triad.

A functional overview presents mostly the schema theory, which is rooted in Immanuel Kant's philosophy and was further developed in Jean Piaget's work. From a modern point of view, schema theory is a "language" to describe different mental functions. Computational and mathematical models of neural networks are particular examples of schemas for realising some particular function. For example, the authors consider in detail the schemas for reaching and grasping, which are based on a cybernetic model with a feedback loop. It is interesting to note that in the 1970s the

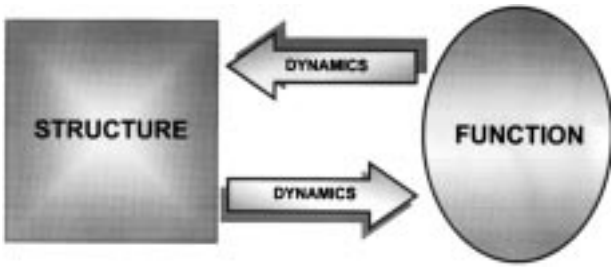

Figure 1 (Borisyuk). Graphical representation of interrelations between elements of philosophic triad: structure, function, and dynamics.

outstanding Russian physiologist academician Petr Anokhin developed the theory of "funcitonal systems." That theory is still the basis for many neurophysiological studies (Sudakov 1997). However, it is not reflected in the book at all.

In "A Dynamical Overview," the methodology of neural network models is considered. The authors define deterministic models and consider attractors of dynamical systems: equilibrium point, limit cycle, and strange attractors. In principle, one more attractor type exists in phase space: a torus, which corresponds to quasiperiodic (envelope) oscillations. Computational models with multifrequency dynamics look very promising for modelling of information processing in brain structures (e.g., in Borisyuk et al. [1999] the feature binding problem is described using multifrequency oscillations).

Different models of a single neuron, neural population, and neural networks are considered as appropriate building blocks/modules for designing particular computational models of brain structure or function. Several types of connection architectures between modules corresponding to different anatomical structures are studied. Note that the authors do not mention an important type of connection architecture, the "neural network with a central element" (Kazanovich & Borisyuk 1999). However, for the description of the anatomy of different brain structures in the book, it follows that the neural network with a central element is a very common architecture for connections between brain structures.

A significant part of the dynamical review chapter is devoted to chaotic behaviour in neural systems. Neural models with chaotic dynamics, models with control of chaotic behaviour, and methods for analysing chaotic neuronal activity (e.g., EEG activity) are presented with many details. Nevertheless, the key question about the functional role of chaotic dynamics in information processing in the brain is left open. The authors discuss the broad spectrum of opinions existing in the literature from the optimism of Freeman to the doubts of Glass. Probably the truth is, as usual, in the middle: chaotic dynamics arises naturally as a possible mode of brain activity. Moreover, chaotic activity in the brain might be synchronised or partially synchronised, and this is a possible basis for information processing.

The authors study brain dynamics at many hierarchical levels and discuss some principles of neural organisation as a necessary component in future brain theory: (1) The principle of modular architectonics (nervous system is composed of building blocks) and (2) the principle of topographical ordinal representation of sensory information in primary sensory cortices and a number of other structures (it is not clear how this principle could be used for auditory or olfactory sensory information). The authors also discuss some other principles of brain information processing indirectly and without a clear formulation: the principle of synchronisation of neural activity; the principal of spatiotemporal information encoding; the principle of flexibility of neural circuits (the same neural elements participate in the implementation of different functions); and so forth. Further progress in this direction seems to be extremely important because the list of principles is likely to be a foundation for the future development of theoretical ideas about brain functioning.

The book has several limitations, such as some misprints and insufficiently careful organisation and systematisation of material. Nevertheless, these limitations do not diminish the contribution the book has made.

It is likely that the book's publication process was too long and took more than three years. The majority of referenced papers were published before 1996. Computational neuroscience is a very rapidly developing field, and during the last three to four years a huge number of interesting and important papers have been published. We can say that the main results and achievements in brain modelling in the last millennium are summarised in the book. Now is the right time to write a new book, which will provide the concepts and ideas for the further development of computational neuroscience in the third millennium. The present book highlights the need to develop a theory of brain functioning and offers some useful approaches and concepts.

NOTE

1. At that time, scientists in the USSR understood the role of oscillations in biological and chemical systems and studied the relations between "structure" and "function."

Neuroanatomy and function in two visual systems

Bruce Bridgeman

Department of Psychology, University of California at Santa Cruz, Santa Cruz, CA 95060. **bruceb@cats.ucsc.edu**
zzyx.ucsc.edu/Psych/Faculty/bridgeman.html

Abstract: Neuroanatomy and neurophysiology are insufficient to specify function. Modeling is essential to elucidate function, but psychophysics is also required. An example is the cognitive and sensorimotor branches of the visual system: anatomy shows direct cross talk between the branches. Psychophysics in normal humans shows links from cognitive to sensorimotor, but the reverse link is excluded by visual illusions affecting the cognitive system but not the sensorimotor system.

For those interested in the physiological mechanisms of human behavior and experience, the findings of neuroanatomy and to some extent neurophysiology have always provided more promissory notes than explications. We can juxtapose detailed descriptions of cytoarchitecture and neural connectivity with analyses of function, but the links remain tenuous. The neuroanatomical analyses can tell us where to look for a particular mechanism, and neurophysiology can tell us when the critical information is being processed, but current techniques leave the "how" of the process tantalizingly elusive. Arbib, Érdi, and Szentágothai have come as close as anyone can to making the two fields relevant to function.

The key intermediate step is modeling, testing algorithms for functionality on one side and consistency with anatomy and physiology on the other. It is the models that tell us whether we understand a process and where the uncertainties remain. Arbib et al.'s style of modeling is to begin with the anatomy and physiology, make a few reasonable assumptions, and attempt to account for a restricted range of functions.

Modeling, however, can be an empty exercise in the wrong contexts. Since Turing (1936) we have known that an astonishingly simple bit of hardware, the Turing machine, can compute any function that is computable (given enough time and a long enough tape), and since McCulloch and Pitts (1943) we have known that a Turing machine can be fashioned from simplified neurons. After that, it is only an exercise to construct a model of made-up neurons to do any given job. Indeed, the mathematics shows us that an indefinite number of different models can be designed to simulate any given data set. The key to productive modeling, then, lies

not in simulating the data but in the additional assumptions and restrictions encountered along the way. This is where Arbib et al.'s work is especially productive and stimulating.

I will discuss one example, however, of where the anatomy and physiology lead us astray, showing that there is no substitute for teasing apart systems psychophysically as well physically. The issue arises in Chapter 8, on the cerebral cortex. Here it is particularly difficult to relate structure and function: all cortical areas look similar, regardless of their functions. This is a strength of brain design, for cytoarchitecture does not have to start anew for each new function, but it makes homeomorphic modeling difficult.

The issue arises in the recent discovery that vision can be divided into two broad functions, a "what" or cognitive system that handles pattern recognition and visual experience, and a "how" or sensorimotor system that controls visually guided behavior. The two systems share a common input and early-vision stage, at least up to the primary visual cortex. In broad anatomical outline the "what" system consists of a series of cortical areas extending into the temporal cortex, culminating in the inferotemporal (IT) cortex. The "how" system passes visual information dorsally and anteriorly, culminating in several parietal areas concerned with motor control. The details of the anatomy are not laid out as conveniently as this summary implies, for the anatomy gets a bit tangled in extrastriate areas; these are only broad generalizations.

The importance of this distinciton is that it takes what seems to be a unitary capacity, the sense of vision, and divides it up into at least two quasi-independent modules. Neurological patients with damage to only one system will be able to manipulate objects but not recognize them, or vice versa, depending upon which system is damaged (Milner & Goodale 1995). The double dissociation establishes the existence of two separate uses of visual information but does not reveal how the systems work in intact humans. Using visual illusions, my laboratory has demonstrated independent representations of visual space in the two systems (Bridgeman et al. 1997; Bridgeman 1999). The observer is aware only of the spatial values in the cognitive system (Castiello et al. 1991). Our dissociations are based on an induced Roelofs effect, a tendency to perceive a target's position as deviated in the direction opposite the offset of an asymmetrically positioned background frame. Open-loop pointing to the same target, however, is unaffected by background position. In these experiments an observer sees a target in one of several possible positions, and a frame either centered before the observer or deviated left or right. Verbal judgments of the target position are altered by the background frame's position. An immediate jab at the target never misses, however, regardless of the frame's position. The cognitive system is misinformed about the target's position, but the sensorimotor system is not.

In some conditions the two systems must communicate with one another. For example, motor activities are normally initiated when the cognitive system tells the sensorimotor system what to do. It might grasp one of several available objects, poke a particular target, and so on. Using the lack of a Roelofs effect as an indicator of the use of information in the sensorimotor system, we have found that a symbolic message about which of two targets to jab can be communicated from the cognitive to the sensorimotor system without communicating the cognitive system's spatial bias as well. Thus cognitive-to-sensorimotor communication can be observed between the two representations.

Arbib et al., however, assert on the basis of neuroanatomy that the two systems engage in direct cross talk (p. 244). They cite extensive anatomical and neurophysiological evidence. If the sensorimotor system were able to pass spatial values to the cognitive system, however, humans would experience neither the Roelofs effect nor induced motion, for the sensorimotor system's spatial representation is not affected by these illusions. We still have a long way to go before neurophysiology and neuroanatomy can reliably specify cognitive processes.

Cognitive incrementalism: The big issue

Andy Clark

Philosophy/Neuroscience/Psychology (PNP) Program, Washington University in St. Louis, St. Louis, MO 63130. **andy@twinearth.wustl.edu**

Abstract: *Neural organization* raises, in an especially clear way, a major problem confronting contemporary cognitive science. The problem (the "big issue" of my title) is: What is the relation between the strategies used to solve basic problems of perception and action and those used to solve more abstract or "cognitive" problems? Is there a smooth, incremental route from what Arbib et al. call "instinctual schemas" to higher-level kinds of cognitive prowess? I argue that, despite some suggestive comments, Arbib et al. do not resolve this issue.

A major feature of *Neural organization* is, as its subtitle suggests, the delicate and principled attempt to interweave structural, functional, and dynamical themes. Getting this interweaving right is, I am inclined to believe, the single greatest challenge facing contemporary cognitive science. One reason it is so important is that (or so I shall argue) getting this balance right may be the key to solving an increasingly pressing problem, that is (and this is the "big issue" of my title), What is the relation between the strategies used to solve basic problems of perception and action and those used to solve more abstract or "cognitive" problems (problems of long-range planning, deductive reasoning, and thought about the distal, the absent, the abstract, the nonexistent, etc.)? I shall first comment on the general problem, then on a suggestion made by Arbib et al.

The general problem is that much of the recent literature on "embodied cognition" often seems committed to a rather dubious notion that I'll call *cognitive incrementalism.* This is the idea that you get full-blown, human cognition by gradually adding bells and whistles to basic (embodied, embedded) strategies of relating to (adaptively coupling with) the present-at-hand. It is just such a principle of continuity that prompts Thelen and Smith, for example, to comment that there is in principle no difference between the processes engendering walking, reaching, and looking for hidden objects and those resulting in mathematics and poetry; cognition is seamless and dynamic (Thelen & Smith 1994, p. xxiii). Much depends, of course, on what is meant by the phrase "no difference between." In many interesting instances we can discern both a kind of (often structural) continuity alongside some quite radical functional discontinuity. As a result, some cognitive functions may depend not on the tweaking of basic sensorimotor processing, but on the development of relatively (functionally) independent and (functionally) novel kinds of neural processes.

A case in point is the two visual systems hypothesis of Milner and Goodale (1995) treated in Chapter 8 of Arbib et al. Milner and Goodale's claim, very (very!) briefly is that on-line visuomotor action is guided by neural resources that are quite fundamentally distinct from those used to support conscious visual experience, off-line visual reasoning, and visually based categorization and verbal report. The latter complex of activities depends, it is argued, upon a ventral processing stream and the former, upon a largely independent dorsal stream. Milner and Goodale's (admittedly quite contentious) explanation thus invokes a radical dissociation of codings-for-on-line action and for off-line reason and imagination. Here, then, is one concrete case in which we seem to confront not a simple incremental process in which off-line reason exploits the very same basic mechanisms as on-line action guidance, but something more dramatic and different: a case, perhaps, where nature adds functionality by developing whole new ways of processing and exploiting sensory input.

Interesting to note, however, Arbib et al. opt for a somewhat different story: one designed to do justice to a kind of continuity while attempting to come to grips with what is special and distinctive about advanced human reason. The story has two main components:

The first is the "schema theory" developed in Arbib (1992) and

elsewhere. Schema theory seeks to characterize the functional (indeed, often representational) role of mental (ultimately neural) resources. But unlike (some) classical approaches, schema theory does not require that there exist inner, static, textlike symbol structures ready to be operated upon by a single kind of "central processing." Instead, a schema is tied to the execution of a type of task in an environment and incorporates both knowledge and procedures for applying the knowledge. It is not invoked by top-down control but by a process of cooperative computation, and is thus designed to function as part of a larger distributed network of such functional elements. Much of *Neural organization* concerns itself with basic schemas used for perceptual and motor tasks. Note that a schema, thus understood, is a functional unit and may be anatomically distributed throughout different structural brain regions.

But the big issue rapidly looms. How can schema theory, which seems so well-suited to understanding basic, often preprogrammed behaviors (e.g., the famous work on frog visuomotor coordination and more generally on what Arbib et al. call "instinctive schemas"), deal with the higher reaches of human thought and reason? The second part of the Arbib et al. story offers a sketch of a mechanism. The idea is that instinctive (basic, special-purpose, perceptuo-motor) schemas are joined by somewhat more abstract (learned) schemas and that these two types of resources then interact in flexible ways to support intelligent human behavior. It is in the interaction that much of the real work, it seems, gets done. Thus, commenting on Milner and Goodale's proposal with its quite severe dissociation of vision-for-thought and vision-for-action, Arbib et al. comment that "it is not enough to dissect more carefully a variety of specialized subsystems. In general, the various aspects of our visual recognition of, and motor interaction with, an object are joined seamlessly" (p. 244). Drawing on Jeannerod (1994), it is suggested (and see Clark 1999 for a similar account) that a rather more complex web of interactions unites the two streams, and (in Ch. 8) that the effective connectivity patterns between component schemas may often be dynamic and malleable.

The trick to high-level cognition, if I understand the authors correctly, lies largely in this kind of seamless joining and dynamic reconfigurability (accomplished perhaps by neural gating mechanisms and/or various binding and unbinding techniques). Hence, in section 8.6, with the alluring title of "From action-oriented perception to cognition," we are told that the instinctive schemas (basic perceptuo-action routines) provide "a basis for, and are intertwined with" (p. 260) rational behavior. The final picture is one in which a variety of basic ("embodied, action-oriented") resources are combined with some more abstract learned schemas and the whole caboodle orchestrated (not top-down but by cooperative computation) into temporary ensembles according to the demands of a current task: one that may call, for example, for the fine interweaving of planning and action, recruiting ventral and dorsal resources in some complex, task-dependent mixture.

As a broad story about the nature of the continuities and more important, the discontinuities that make human cognition the flexible tool that it is, this is an appealing model. But it leaves one crucial element unexplained: how does the process of cooperative computation actually solve the recruitment problem? How are the RIGHT schemas assembled into the *right* temporary wholes at the right times? With the right "coordinated control program" (p. 250) in place, some very complex problems (see e.g., the VISIONS example in sect. 8.6.1) can be solved. But how do we learn these control programs and how is that learning neurally implemented? Must the right control program be in place in advance or can we generate it "on the fly"? Regarding "how the various assemblages that exercise . . . restructuring are themselves acquired and updated through experience," we are told only that all this "remains to be discovered" (p. 250). The trouble is that as long as that is so the big issue remains unresolved. For what really counts is the unknown process by which the right stuff (schemas) gets recruited at the right time so as to meet higher-level demands. But that is, in essence, a version of the notorious frame problem itself! The question about cognitive incrementalism then becomes simply the question whether the fundamental principles that lie behind nature's solution to the "flexible schema recruitment" version of the frame problem are functionally continuous with those developed to subserve embodied, embedded cognition. And since we don't know the answer to the recruitment problem yet, we cannot really say one way or the other. But it looks (and here I must agree with Fodor 1983) increasingly unlikely.

I don't wish this to sound too negative. The basic picture that Arbib et al. paint is one I find deeply attractive (see, e.g., Clark 1997). But I do think that deep questions about cognitive incrementalism remain unresolved, and that the final value of the embodied, action-oriented approach, as an approach to *cognition*, depends almost entirely on the outcome.

Let me end, however, with a mere speculation. It strikes me as possible that we shall need *two* fundamentally distinct frameworks for understanding cognition. One framework would explain how, given a set of temporarily recruited resources (given an effective connectivity pattern), the organism solves a specific problem. Here, we may well be on familiar ground, able to exploit computational and representational stories of a moderately conventional kind (though with distributed encodings, processes of cooperative computation, partial, action-oriented representations, etc.). The other framework, still barely imagined, would deal with the way multiple specialized neural resources exploit complex feedback and feedforward pathways, and dynamically modifiable links, so as to create the right instantaneous cognitive architecture to solve the problem at hand. Comprehending this latter process requires us to do simultaneous justice to the (genuine) specialization of neural structures and to the complex interdependencies and interactions that create new, temporary functional wholes from this underlying cloth. It is my suspicion that this latter process, though arguably at the heart and soul of higher-level cognition, is itself either *not a cognitive process* at all, or else is a cognitive process of some fundamentally different stripe – one perhaps best investigated using the rather different resources of some (information-processing-friendly?) version of dynamical systems theory.

The basic framework sketched by Arbib et al., with its valiant attempt to balance considerations of anatomical structure, potentially distributed functionality, and complex self-modifying dynamics, offers a detailed, valuable, and timely window onto this kind of new and puzzling territory. But the big issue (here surfacing as the superficially innocent question, how do we acquire, modify, and appropriately activate coordinated control programs?) remains unresolved.

ACKNOWLEDGMENTS

Thanks to Ned Block, Jesse Prinz, Susan Hurley, and Dominic Murphy for useful input concerning these issues.

A moveable feast

Peter F. Dominey

Institut des Sciences Cognitives, CNRS UPR 9075, 69675 Bron Cedex, France. **dominey@isc.cnrs.fr** **www.isc.cnrs.fr**

Abstract: *Neural organization* achieves its stated goal to "show how theory and experiment can supplement each other in an integrated, evolving account of structure, function, and dynamics" (p. ix), showing in a variety of contexts – from olfactory processing to spatial navigation, motor learning and more – how function may be realized in the neural tissue, with explanatory and predictive neural network models providing a cornerstone in this approach.

In *Neural organization*, Arbib, Érdi, and Szentágothai undertake the quite noble but difficult task of weaving into a coherent tissue a multidimensional characterization of their collective views on

modern neuroscience. At first sight of the table of contents, the reader will be quite pleased with the variety of subjects treated, from nervous system development, dynamics, and chaos to the olfactory system, cortex, thalamus, basal ganglia, and more – a true intellectual feast. But the reader may also wonder how such a broad range of topics could be treated in a single volume, if not just as a collection of unrelated chapters.

The final result of the authors' efforts here, however, has yielded a volume that will be of significant value in the training of new and confirmed neuroscientists for two related reasons. First, many of the chapters stand alone as excellent treatments of topics of central importance. At the same time, through the "meta chapters" and extensive cross-referencing between chapters, the book does achieve the difficult task of providing an excellent and unique view of the connections among these topics, showing in a variety of contexts, from olfactory processing to spatial navigation and motor learning, how these functions may be realized in the neural tissue with explanatory and predictive neural network models providing a cornerstone in this approach.

In the remainder of this review, I will not repeat the contents of the book, but rather comment on specific aspects of special noteworthiness. Thus the first two major parts treats the thematic issues of structure, function, and dynamics. The structure chapter provides among other things a quite useful perspective on the adult nervous system by way of embryological development. In the function chapter, schema theory is introduced as a methodology for describing function and allowing the mapping of function onto neural systems, which can then be tested in corresponding neural network models. Of particular value here were case studies demonstrating the utility of the interaction between experimentation and modeling in frog approach versus avoidance behavior and in primate reaching and grasping. The chapter on dynamics provides an excellent treatment of this topic (including clear introductions to ideas of state spaces, attractors, stability, chaos, etc.) and its relevance and application to neuroscience. Relating back to ideas developed in the treatment of structure in Chapter 2, self-organization is addressed within this framework.

While Part I thus introduced the thematic foci, Part II then treats six of the major neural systems in modern neuroscience, the olfactory system, hippocampus, thalamus, cerebral cortex, cerebellum, and basal ganglia, and then terminates with a synthesis and prospective for the study of the neuroscience of cognition. At first glance it might seem rather presumptuous to treat, in single chapters, topics as vast as the olfactory system or the cerebral cortex! But within the framework established for this book I think that the authors have done an excellent job in treating these topics from the perspective of systems neuroscience. That is, they have looked at the defining neuroarchitectural properties of the system in question, characterized the functional role of the system and presented – when possible – neural network models that achieve the specified function within the identified structural constraints.

Thus a central strength of the book is this systems neuroscience approach that is applied to a rather overwhelming collection of brain systems. The importance of this approach should not be underestimated. While it is true that great progress is being made in more fine-grained approaches to neuroscience, including molecular biology, progress in understanding system-level function can only proceed by putting these elements together into system-level models. A second major strength of this book comes from the effort, already alluded to, of the authors to assure the connectivity among the interwoven topics. A clear example of this is in the chapters devoted to thalamus, cortex, and basal ganglia, with extensive referencing among these chapters.

The obvious criticism that can be levied at this book is that by treating, in a series of chapters, topics that merit complete volumes themselves, the authors must have been forced to cut some corners. Indeed, it is clear that a treatment of "the basal ganglia" in thirty pages cannot be complete. Thus, for example, it would have been interesting to see how data from the treatment of Parkinson's disease by selective lesions and stimulation in the basal ganglia nuclei can provide important constraints for improved models of basal ganglia function. To such criticism the authors can offer two responses. First, while Chapter 10 itself is only thirty pages long, this is in fact misleading, since the basal ganglia and their fundamental role in thalamocortical interaction is already illustrated in the chapters on thalamus and particularly cortex. In other words, the whole is greater than the sum of its parts.

The second response to the criticism that "they tried to do too much" is related to the stated objective of the book. It does not pretend to be a complete reference volume on neuroscience – such animals already exist. Instead, in its stated goal to "show how theory and experiment can supplement each other in an integrated, evolving account of structure, function, and dynamics" (p. ix) the book is a clear success. It provides a series of examples for affirmed and developing neuroscientists of exactly this – the fruitful interaction between theory and experiment, providing a guide for the systems neuroscience methodology that will clearly play an important role in the ongoing challenge of understanding the brain and its structure, function, and dynamics. What may now be required is an effort to integrate the pieces (i.e., the models of cerebellum, basal ganglia and cortex, hippocampus, olfactory system, etc.) into a system model, with "macaca computatrix" as the descendent of "rana computatrix."

Brahe, looking for Kepler

Shimon Edelman

Department of Psychology, Cornell University, Ithaca, NY 14853-7601.
se37@cornell.edu **kybele.psych.cornell.edu/~edelman**

Abstract: Arbib, Érdi, and Szentágothai's book should be a required reading for any serious student of the brain. The scope and the accessibility of its presentation of the neurobiological data (especially the functional anatomy of select parts of the central nervous system) more than make up for the peculiarities of the theoretical stance it adopts.

> . . . inkhorns, arrows, loaves, cruses, fetters, axes, trees, bridges, babes in a bathtub, shells, wallets, shears, keys, dragons, lilies, buckshot, beards, hogs, lamps, bellows, beehives, soup-ladles, stars, snakes, anvils, boxes of vaseline, bells, crutches, forceps, stags' horns, watertight boots, hawks, millstones, eyes on a dish, wax candles, aspergills, unicorns.
>
> – Joyce, *Ulysses,* p. 325

1. A scrambled mosaic. One bright June day in graduate school, I asked my advisor to recommend some reading material for the approaching summer. One of the articles thus recommended left me with an impression sufficiently vivid to prompt total recall on the slightest provocation (such as groping for an opening line for the present review). The article in question (Gilbert 1983), albeit informative and well-written, evoked a lingering feeling of disappointment, no doubt because its title – "MicroCircuitry of the Visual Cortex" – had initially sounded misleadingly suggestive to a literal-minded ex-electrical engineer such as myself. That article listed all manner of neurons and their distribution throughout the visual cortex, but, alas, did not quite specify wiring diagrams.

Satisfyingly, *Neural organization* does provide some actual wiring diagrams, and not just for the visual cortex. The anatomy of other neocortical, archicortical, thalamic, and cerebellar structures as well as their development and models of their function are discussed. All this adds up to a grand tableau resembling nothing as much as a Joycean procession of saints "bearing symbols of their efficacies," complete with buckshot, beards, bellows, and beehives (not to mention eyes on a dish).[1]

The attempt undertaken by Arbib et al. to introduce some order into the masses of currently available findings on brain function is highly commendable. That the result appears less orderly than

hoped for can be attributed in part to the sheer scale of the enterprise. It is tempting to compare the present state of knowledge in neurobiology (and, on a different level, in cognitive psychology) to that of pre-Ptolemaic astronomy: there are observations to be explained, but theories (even wrong ones) are hard to come by. This analogy, however, is strained: the amount of neurobiological data to be explained is vaster than anything that Ptolemy (or even Kepler) had to confront, and the explanation is highly unlikely to consist of a few equations accompanied by a simple diagram. The problem of commenting on books such as *Neural organization* is, therefore, this: how to evaluate an attempt to reconstruct a huge scrambled mosaic, without having a clear idea of what the big picture is supposed to show, or, for that matter, whether or not there *is* a single big picture at all.[2]

In lieu of speculations concerning the existence and the possible nature of the big picture, the book offers three threads common to all the chapters: structure, function, and dynamics (as reflected in the book's subtitle). I shall comment on these in the reverse order, starting with dynamics.

2. Dynamics. The argument underlying the book's construal of dynamics as an all-encompassing explanatory paradigm is that of unity of approach across time scales. The basic premise – that phenomena involving change over time are properly described by differential equations – is unimpeachable. I wonder, however, whether it is productive to lump together, say, compartmental models of membrane potential in single neurons on the one hand and models of development of neural wiring on the other hand. Differential equations are ubiquitous in the sciences; most scientists, however, do not take this to mean that all natural phenomena, even in their own field, are merely different aspects of a single whole.[3]

Fortunately (I think), the book does not actually adhere to this extreme approach: apart from Chapter 4 (which offers a very good survey of "neurodynamical system theory") and a few paragraphs in the last chapter, "dynamics" does not play too prominent a role in the book. Typically, dynamics and differential equations tend to crop up where explanatory (rather than descriptive) approaches are as yet unavailable. For example, toward the end of Chapter 4, "invariant" pattern recognition is finessed into the "binding problem" and is offered a solution in the form of von der Malsburg's (1985) dynamic link architecture (p. 102) – a model that only becomes relevant if one accepts the prior assumption that binding is indeed a problem.[4] Likewise, in Chapter 8 (which is devoted to the cerebral cortex), one finds dynamical models of ocular dominance formation (nothing is said about *why* there are ocular dominance columns in the first place), and of thalamocortical oscillations (another phenomenon whose function remains obscure).

3. Function and schemata. If we relegate dynamics to the status of a mathematical means rather than an explanatory end, the issue of function is brought to the fore. Regrettably, *Neural organization* takes "function" consistently to mean "how does this bit of the brain function?" rather than "what does this bit of the brain do for a living?" In other words, no clear distinction is made between explanation of operation and explanation of goals and means.

The dearth of this latter kind of explanation is apparent throughout the book. For example, the most succinct and explicit description of the function of the visual cortex in all of Chapter 8 is found in a quote from J. Maunsell (1995), suggesting that "while the early stages of processing in the visual pathway provide a faithful representation of the retinal image, later stages of processing in the visual cortex hold representations that emphasize the viewer's current interest" (p. 223). To be sure; but what are these representations, and why do they have the properties they do, and how do they support visual behavior?

The preference for operational rather than functional explanation that one finds in the book is explicitly justified in section 11.2.2 ("Brain theory"): "The issue for the brain theorist, then, is to map complex functions, behaviors, and patterns of thought either on the interactions of these rather large entities (anatomically defined brain regions) or on these very small and numerous components (the neurons)" (p. 338). It seems to me that such a mapping would constitute an incomplete theory of the brain, unless it includes an explanation of its function (over and above its operation). The theoretical concepts which Arbib et al. bring to bear on this issue – dynamics, self-organization, and schemata – are inherently incapable of filling this lacuna. The reason for this is simple: dynamics, self-organization, and schemata all belong to a level of explanation that deals with the operation of a system, rather than its function. In particular, attempting to explain the function of a system by saying that it employs schemata is like explaining how a computer program fulfills its function by saying that it employs subroutines (to make this analogy work better imagine applying it to some really complicated piece of software, such as SABRE, the flight reservation system).

Chapter 3 is entirely devoted to an explanation of the concept of schemata. The overview is excellent and wide-ranging, but it does not convince me that recruiting schemata as an explanatory aid (let alone making the concept of schemata an explanation *an sich*) can advance a theory that does not otherwise address itself to the functional level. Intuitively, the idea of multiple schemata, agents, subroutines, etc., operating in parallel seems to be a very plausible framework for trying to understand brain and behavior. In itself, however, it is not an explanation, and attempts to present it as such tend to give rise to oxymoronic titles such as "A robot that walks: *Emergent* behaviors from a *carefully evolved* network" (Brooks 1989); the emphasis is mine. In other words, if you want your flock of schemata to cooperate and do something useful, these days you still have to carefully engineer your system, and that takes understanding. Moreover, if you ever succeed to have the schemata evolve without supervision, you'd still want to analyze the emergent behavior to gain understanding of what is going on – just the problem we have with explaining brain function.

4. Anatomy. The discussion of structure (that is, anatomy) is truly excellent throughout the book (which is only expected, given Szentágothai's contributions to this field over the past half century). The anatomical data are presented in a lucid form, and are accompanied by outstanding illustrations and a good discussion. I found the chapter on the hippocampus, with its "systems view" (p. 170), especially illuminating. The problem of integrating current ideas about the function (not merely the operation) of hippocampus in animals and in humans is well presented. The prospects of linking the idea of cognitive maps with more general notions of memory and other cognitive processes are very intriguing. The possibility of the involvement of the hippocampus in the representation of relational information (discussed on pp. 182–84)[5] is fascinating, in view of the present attempts of brain theorists to address the so-called problem of compositionality (Bienenstock & Geman 1995). I hope this section will prompt further studies and will generate more discussion in the future.

5. Summary. The last chapter presents some wonderful opportunities for a lively debate. To quote just one example, the reader is told that "Digital computers . . . have low adaptability" (p. 340); this *must* mean that a compartmental simulation of a neural circuit running on a digital computer is equally handicapped. For the sake of brevity, I shall forego all these opportunities,[6] and proceed to summarize my impression of the book.

I believe that this book is a required reading for any serious student of the brain. Even if one disagrees with the theories *Neural organization* propounds, the scope and the accessibility of its presentation of the neurobiological data make it more useful than most textbooks and many reviews.[7] My guess is that it will stimulate many readers to try their own hand at the big mosaic, perhaps playing Theon and Hipparchus to a future Ptolemy – or maybe even Brahe to someone's Kepler.

NOTES

1. I wonder to what extent this is intentional; on p. 337, this approach to the study of the brain is characterized as "nonmonolithic."

2. Marr, in one of his last papers, expressed concern regarding the prospects of developing comprehensive, analytically explicit (Type I) the-

ories of some visual submodalities, the alternative being the idea that these may turn out to be describable only as bags of unrelated (Type II) tricks (Marr 1981).

3. As one should expect, there are exceptions, such as S. Grossberg's recurrent invocation of "neural dynamics" as an explanatory device.

4. Arbib et al. do not define the binding problem until p. 244, where it appears in italics, and is described as "linking representations of a single object, task, or action widely distributed over many brain regions." For an alternative view of binding, see (Edelman 1999), p. 246.

5. A central reference discussed in the text, (Eichenbaum et al. 1994), does not appear in the book's bibliography.

6. I must mention, however, a disturbing typographical error in three places on p. 339, where Descartes' *res cogitans* is spelled "*res cognitans.*"

7. *Neural organization* would have been even more useful if the Web links mentioned in Appendix B worked. None of the links I tried allowed me to get much farther than the front page of the project; some links were dead, others pointed back to the front page, and some were password-protected.

A multidimensional approach to the mind-brain: Behaviour versus schemata versus cognition?

Jonathan K. Foster

Programme in Neuropsychology, Department of Psychology, University of Western Australia, Nedlands, Perth WA 6907, Australia.
jonathan@psy.uwa.edu.au

Abstract: Arbib, Érdi, and Szentágothai's book seeks to present a multidisciplinary, multistrategied approach to the study of the mind-brain, encompassing structural, functional, and dynamic perspectives. However, the articulated framework is somewhat underspecified at the cognitive level. The representational level of analysis will need to be fleshed out if the explanatory potential of Arbib et al.'s framework is to be fulfilled.

In their volume, Arbib, Érdi, and Szentágothai valiantly attempt to develop a multidimensional, integrative, and multistrategied approach to the study of brain-behaviour relationships. In so doing, they seek to integrate data from a wide range of empirical sources, ranging from anatomical tracing studies via single cell electrophysiology to computational network modelling. Arbib et al. attempt to synthesize structural, functional, and dynamic approaches to the mind-brain, and they sensibly consider the extent to which these complementary perspectives can be truly dissociated. Their attempt to explore such a broad canvas is ambitious, and it is especially laudable and refreshing in the current era, given the strictures placed on many research workers by sometimes obsessively focused funding bodies and superspecialised academic meetings and journals. The book would be most suitable for those at the early graduate student or later stage of their career, working within neuroscience (broadly defined).

As a whole, the book is strongly rooted in neuroscience, and in that tradition the emphasis – from a psychological perspective – is on behaviour rather than on cognitive functioning. This is understandable when one is basing one's conclusions on data derived from studies involving a number of animal species, but this approach also presents certain inevitable constraints. The emphasis on behaviour rather than cognitive representations is also mirrored at the anatomical level, with more coverage of subcortical structures than of the cerebral cortex, albeit within a framework of interactive brain systems. At the level of the neural substrate (the "hardware"), the authors consider a range of perspectives, encompassing anatomical, physiological, and computational frameworks. However, at the representational ("software") level, the approach taken by Arbib et al. is somewhat underspecified. This is perhaps unsurprising given the respective scientific backgrounds of the three coauthors, but it is significant nevertheless.

From the top-down perspective, there is extensive reference by the authors to the role of "schemata." However, the use of this catchall terminology corresponds with little exploration in any real detail of specific intervening variables, hypothetical constructs, or psychological component processes and mechanisms, and there is consequently limited exploration in the book of the full range and depth of the human cognitive landscape. This has important implications, in terms of the kinds of top-down constraints that are considered by Arbib et al.

In the past 20–30 years, cognitive neuropsychology has made tremendous strides in delineating the "architecture of cognition" and in "carving cognition at its joints," so that the cognitive mechanisms underlying processes such as attending, reading, speaking, remembering, and identifying are now quite well specified. Although theirs is a bold and wide-ranging endeavour, Arbib et al. do not consider adequately findings in cognitive neuropsychology in particular (and neuropsychology in general) when articulating their theoretical framework of mind-brain function. The cognitive neuropsychology school of scientific research has placed particular emphasis on functional dissociation as the methodology of choice for identifying specific functional units (or "modules") subserving particular cognitive and computational processes. The way in which distinct functional modules map onto physiology and anatomy (how the "software" can be interrelated with the "hardware") provides an ongoing challenge, and one that needs to be fleshed out if the explanatory potential of Arbib et al.'s framework is to be fulfilled. In so doing, it will be important to think of how particular, detailed cognitive representations are embodied by particular hardware elements of the brain, and whether these are represented in a localist, a distributionist, or some hybrid manner. In the future, a more complete interdigitation and dovetailing of neurobiological, computational, and cognitive data and concepts will be critical.

More specifically, there is, unfortunately, little consideration in the chapter on the hippocampus of the ongoing debate between those researchers who regard the hippocampus as being specifically involved from a cognitive perspective in the mediation of recall memory (for example, John Aggleton) and those who argue that it has a more wide-ranging role in subserving memory, encompassing both recall and recognition (for example, Larry Squire), nor of the contentious role of the hippocampus in the phenomenon of retrograde amnesia. These are significant omissions when one is considering how best to interrelate structural, functional, and dynamical aspects of this important brain region and highlight the problems one faces when top-down constraints are not adequately considered in formulating an ambitious framework of mind-brain functioning.

Analyzing the brain

Peter Gouras

Department of Ophthalmology, Columbia University, New York, NY 10032.
pg10@columbia.edu **cpmcnet.columbia.edu/dept/eye/retina**

Abstract: *Neural organization* describes an approach to analyzing neural function in anatomically defined subsystems in the brain, the hippocampus, cerebellum, sensory systems, thalamus, basal ganglia, and cerebral cortex, combining information on neurocircuitry with mathematical models that link structure with function. It is an up-to-date source on the major schemes and background for neural modeling of the central nervous system and is combined with a Web site that includes tutorials and on-line modeling possibilities.

Neural organization describes one of the best methods to understand how our brain works but at the same time illustrates how difficult the problem still is with current technology. The book concentrates on the three main strategies needed to define brain function: the anatomical connections between neurons, the time-varying excitatory and inhibitory signals within these multiple neural channels, and the logic of these signals and connections when considered in concert.

The authors emphasize that one must distinguish the forest from the trees in considering global brain function. But the book shows that this is still impossible. The authors had to chop the forest down into subunits of operation, defined by anatomy: hippocampus, cerebellum, sensory systems, thalamus, basal ganglia, and cerebral cortex. Many important links that connect all of these structures are broken in the analysis. Therefore, seeing the forest in its entirety remains difficult. Nevertheless the strategy described and well developed in this book is a promising and tractable one to achieve the ultimate end result, understanding how the brain works as a total organ.

Each of these neural centers is considered in separate chapters. The known neural connectivity within, from, and to the center is described. In general, the connectivity is far from complete for many of the subunits, for example, the cerebral cortex. Then a theoretical model relating input to output is suggested in a mathematical framework. This approach has been most successful in analyzing neural centers involved in stereotypical quasireflex operations like the vestibular ocular reflex and the saccadic oculomotor system, which have obvious motor corollaries in their programs. This strategy becomes procrustean when applied to more arcane structures like the hippocampus or the cerebellum. Nevertheless the exercises of analysis in each of these areas is valuable for all participants in this field.

There are certain topics that are included, presumably to be complete but which complicate the picture. For example, the consideration of epilepsy or oscillations or the pharmacological ingredients that are liberated at synapses seem irrelevant. It is really important to the purpose of this book to know only whether a particular synapse is excitatory or inhibitory and not what chemical determines this effect. That an epileptic focus can develop in certain areas, presumably due to some defect in negative feedback, also seems irrelevant. Modeling epilepsy or oscillations seems of secondary importance. The need to describe the Hodgkin-Huxley equations for propagated axonal conduction, though a highpoint in the history of mathematical descriptions of neural function, is of little relevance to understanding the logic of how neural connectivity defines brain function. Instead, more space could have been devoted to describing new insights into wide-field network properties of the living brain using functional neuroimaging.

What is invaluable in this book is its attempt to define these functional links with a semistandardized mathematical logic and describe everything that is known about the connectivity within each of the neural systems considered. This is therefore a useful source for students as well as advanced investigators in the field of brain science. There are some minor errors in the bibliography, but the illustrations and format are, on the whole, good.

Few analogies are made between the brain and the evolution of the World Wide Web, another example of how standardized sequences can be linked in a universal way to create much more complex abstractions. But the authors describe a useful format for heuristically integrating and expanding their approach on the Web at http://www-hbp.usc.edu.

The book is dedicated to the wife of one of the authors, Janos Szentágothai, who died during the writing of the book. Dr. Szentágothai was a giant in the field of neuroanatomy and was a driving force in unraveling the brain's complex connections. His lifetime's work and his insight into the importance of anatomy for understanding the brain are aptly commemorated in this publication.

Chaotic dynamics and psychophysical parallelism

Robert A. M. Gregson

Division of Psychology, School of Life Sciences, Australian National University, Canberra, ACT 0200, Australia. **ramg@macquarie.matra.com.au**
www.psy.anu.edu.au/psychophysics

Abstract: An impressive review of brain neurophysiology provides the basis for modelling the dynamics of transmission in neural circuits, using appropriate nonlinear mathematics. The coverage is unbalanced, however: the parallel dynamics at the level of behaviour and sensory-cognitive processes are sparsely addressed, so the final chapter fails to indicate the complexity and subtlety of relevant modern work.

Neural organization is so impressive and comprehensive in its coverage that it may confidently be expected to serve as a preferred reference source for some time to come. Arbib et al. have initially set out the plan of their work so that the intricate interrelation of the sections, both neuroanatomical and mathematical, can be followed even by the novice in the area. A work like this could not be written, nor reviewed, by a single person as the necessary erudition cannot today be found even in one polymath. The rapidity with which new physiological discoveries and mathematical subtleties are being presented in the relevant literature, scattered over a diversity of disciplines, means that though this work covers results in some areas up to 1995, inevitably the hypercritical can discover gaps that would need remedial treatment in a second edition.

My review focuses on Chapter 4 and its linkage to Chapter 11, treating other chapters as supportive evidence to the arguments pursued. Consider the two levels of treatment, neurophysiological and mathematical, as the two sides of the dynamics. The objective is to match, not necessarily as strict isomorphisms, what the events in the neural pathways and the trajectories in the preferred mathematical models can together tell us about such matters as sensory acuity, timing, plasticity, learning capacity, memory, adaptation, and even consciousness. In short, the body-mind problem is still with us and is honestly though cursorily raised in the last chapter.

Precisely because dynamics is applied mathematics, here applied to cortical neurophysiology, what a dynamic treatment requires to carry conviction is four things A,B,C,D, which may be thought of as being in a 2×2 arrangement. One contrast (A-C, or B-D) is between physiological and psychophysical, the other (A-B or C-D) is between substantive data and mathematics. The coverage of substantive-physiological (cell A) is the strength of the book, the other three cells are subordinate to it. For example, in cell A the critical role of timing within closed but leaking loops of excitation can be traced, in Chapter 6, in the pathways involving the hippocampus; and how such pathways will function when they are themselves linked and arranged in a dominance hierarchy can be modeled (in cell B) by simultaneous nonlinear difference-differential equations. Chapter 4 is their cell B coverage; Arbib et al. give various examples and prefer the more traditional differential equations written in continuous functions.

More recent mathematical developments in nonlinear dynamics should feature in cell D but are absent. Consequently the cells C,D do parallel what the authors are attempting, but instead of using nonlinear dynamics to model behavioral data from cell C, which is needed to complete the picture, this course is ignored. This is perhaps presently excusable because such developments are recent and specialized (Gregson 1995; 1999), but if the body-mind problem (or a reductionist variant) is the justification for considering the dynamics of whole subsystems in the brain for their part in consciousness, then the book has still a built-in yawning gap.

I am, like others such as Haken or Kelso whom Arbib et al. cite in their bibliography, in complete sympathy with getting away from single idealized neurons and using the properties of recursive pathways as the elements of modelling. Such pathways need

augmentation by gating functions, whose dynamics have been studied from both physiological and mathematical perspectives. The practical difficulties of identifying what the simplest underlying dynamics actually are need emphasis; the roles of delay and noise are ubiquitous (Aguirre & Billings 1995; Arhem et al. 1999; Chapeau-Blondeau & Chauvet 1992).

Not only in terms of convenient computer simulation, but in psychological dynamics, discrete time processes may be more plausible and tractable (Dehaene 1993; Geissler et al. 1999). The treatment of chaotic dynamics (sect. 4.3) introduces, appropriately, Lyapunov exponents and their spectra, but is weak on one-dimensional complex systems (Gregson 1988; Milnor 1992) and may give the misleading impression that sensitivity to initial conditions is a legitimate defining property and test of chaos, which it is not (Crannell 1995). It is the capacity of nonlinear dynamics to jump around, reversibly or irreversibly, in the phase space between the Fatou and the Julia sets and exhibit what is observed as nonstationarity in outputs with hidden stationarity in higher-order system parameters, which is potentially most useful in modelling in biological systems (Christiansen et al. 1997). This leads to new statistical methods that are needed to cope with sparse and short data strings (Buhlmann 1999); such sparseness is a perennial problem in confirming dynamical modelling when real data are from cell A, as opposed to being from physics, and certainly from cell C.

Now that we have a marriage of cortical microphysiology and dynamics, which computer simulation has made possible, our understanding of systems that are chaotic at one level of resolution and stationary and almost linear at another will have to be explored more, both substantively and mathematically (Contopoulos 1998). That such multilevel systems exist, as the authors note, in part explains why simpler models in psychology have survived so long but only as locally plausible, but why it has been so difficult to construct plausible global models after the paradigm of the physical sciences.

Schema theory: Very promising

Alexander Grunewald

Division of Biology, California Institute of Technology, Pasadena, CA 91125.
alex@vis.caltech.edu **www.vis.caltech.edu/~alex**

Abstract: A direct equivalence between neural function and neural structure does not provide a fruitful approach to understanding brain functioning. Arbib et al. describe a new and powerful approach to circumvent this problem, which they call schema theory. However, in examples they fall prey to the tradition of finding such equivalences, not doing schema theory justice.

The goal of artificial intelligence (AI) was to build machines that could think. To do this, human thinking was used as a guide to develop better thinking machines. AI never really succeeded in this goal, but it had a profound impact in the converse direction: many notions about how the brain works come from AI. For example, in AI it was assumed that there is a representation of the sensory world that is multimodal, not tied to a specific sensory coordinate system, and from which information could be used to perform any movement (Marr 1982). While this notion is attractive for its simplicity, it has been difficult to find psychophysical or physiological evidence in favor of such a central representation.

The first major defeat for the notion of a central representation was provided by compelling evidence that suggests that there are two visual pathways, one for perceptual information and one for action information (Goodale & Milner 1992; Ungerleider & Mishkin 1982). Recent evidence suggests that there are also two auditory pathways (Romanski et al. 1999), separated by their perceptual or motor functions. It is possible that analogous auditory and visual action pathways converge at one central action representation in frontal cortex. However, that representation is not likely to play a central role for action since several parietal areas that are part of the dorsal pathway are preferentially active before arm movements but not eye movements, and vice versa (Snyder et al. 1997). If the earlier stages of visual processing already have a movement preference, then it seems unlikely that there is a later stage that is independent of movement planning.

With all these developments one wonders whether a central representation exists at all. It is in this discussion that Arbib et al. provide a great contribution with their book. They abandon the notion of a central representation, and instead propose schema theory, where schemas are functions that the brain carries out. The authors suggest that there is no isomorphism between schemas and brain structures, but rather that this is a many-to-many mapping. This mapping can be restricted to a one-to-one mapping depending on the context. In other words, a given brain structure may participate in various schemas over time, but in only one schema at a given time.

Similarly, any schema may arise due to the cooperation of several neural structures. Recently I have suggested that to understand neural processing it is useful to think of various areas providing different competencies (Grunewald 1999). The words "competencies" and "schema" denote similar ideas: some specific functionality that can be combined with other functions as the context requires. I prefer the term "competency" only because it is more intuitive. In essence I am suggesting that instead of there being just two neural pathways (one for action and one for perception), there may in fact be many pathways, and that each pathway is called upon depending on the contextual demands, for example, due to task requirements. These pathways do not by necessity exist in parallel, in the sense that they do not overlap, but rather they may share common neural structures, and the assignment to which neural pathway a structure belongs in a given situation occurs dynamically, again as a function of contextual demands. This dynamic reassignment of the specific role of a given brain area is in line with the central role that Arbib et al. have in mind for neural dynamics. Arbib et al. have to be commended for bringing this point out very clearly.

Schema theory is first introduced in section 3.1. The book introduces perceptual and motor schemas, which are subdivided into subschemas. Schema theory provides a great framework to study brain functioning, but as always the difficulties come with the details. The authors provide a detailed discussion of how such schemas work, for example, the *saccade* schema. In that model (described in sect. 8.4) they explain how saccades are generated in a variety of contexts. As an example for their approach, they show how memory arises out of the interaction between the frontal eye fields (FEF) and the thalamus. Thus neither the FEF nor the thalamus alone is responsible for memory, supporting the notion (and need for) schema theory. However, in other aspects they assign specific functions to specific brain areas. For example, they suggest that the role of the posterior parietal cortex (PPC) is to provide a remapping mechanism. Yet it is known that many neurons in the PPC code a variety of signals at different times including memory signals during visual memory saccades (Gnadt & Andersen 1988) or auditory signals in fixation or saccade tasks (Grunewald et al. 1999; Linden et al. 1999).

Arbib et al. reassure the reader that with more experimental data a more careful assignment of functions may be possible. However, one is left wanting, since many aspects of this model do not distinguish it from other models without the benefit of schema theory. Moreover, many of the puzzling details are left unexplained. For example, the role of the FEF cannot be restricted to saccade or memory functions alone, since recent physiological experiments suggest that the FEF is also involved in other functions (Bichot & Schall 1999; Thompson & Schall 1999). In a sense, to make schema theory concrete, Arbib et al. find it difficult to escape the traditional approach of assigning specific functions to specific brain areas. I interpret this failure not as an argument against schema theory, but rather as evidence how prevalent the

view of functional specificity is, how difficult it is to escape it, and how important schema theory is.

At present there are many questions regarding schemas that address the identity of schemas: What constitutes a schema? What differentiates two schemas that are related? Can two schemas be hierarchically organized in one context, but be parallel in another? While the book by Arbib et al. does not answer these questions, it provides a good starting point to ask these questions, which will be very important for neuroscience in the years to come.

A dynamical system theory approach to cognitive neuroscience

D. Heinke

Cognitive Science Research Centre, School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom.
d.g.heinke@bham.ac.uk **web.bham.ac.uk/heinkedg/**

Abstract: *Neural organization* contains a wealth of facts from all areas of brain research and provides a useful overview of physiological data for those working outside the immediate field. Furthermore, it gives a good example that the approach of dynamical system theory together with the concepts of cooperative and competitive interaction can be fruitful for an interdisciplinary approach to cognition.

In *Neural organization*, Arbib, Érdi, & Szentágothai aim at an interdisciplinary approach to understanding the functioning of the human brain. In the first part of the book they present a "schema-based theory" as a framework for achieving this objective. In this approach, there is an initial decomposition of behaviour into coarse processing modules (e.g., short-term memory) – the schemas. Schemas operate in parallel and interact with each other mainly in a competitive and cooperative fashion. In a second step, this functional level of analysis is linked with a structural decomposition of the brain, for instance, derived from studies of brain lesions. Neural networks are then applied that attempt to implement the bottom-up constraints from neural circuitry. Here, according to the authors, dynamical system theory (Ch. 4) plays an important role as a mediator between the functional-structural level of analysis and the neural level. The authors stress, rightly, that satisfying the constraint from the two levels is not a straightforward process and that it requires numerous iterations. The authors develop schema-based theories for various behaviors, for example, approach and avoidance in frogs and toads, eye movements, visual scene interpretation, and reaching and grasping. In Part II these schema-based theories are applied to particular brain regions, such as the olfactory system, the hippocampus, the thalamus, the cerebral cortex, the cerebellum, and the basal ganglia. Here there is an emphasis on the role of subcortical structures in cognition.

The book contains a wealth of facts from all areas of brain research. Although probably not up-to-date in all areas covered – inevitable considering the number of facts it covers (see the following paragraph for a concrete example) – it certainly provides a useful overview of basic physiological data for those working outside the immediate field. However, there are also shortcomings. For example, there is a paucity of data from behavioural (whole system) levels of analysis – from experimental psychology and neuropsychology, for instance – which is unfortunate, since the behavioural level, as pointed out by Arbib et al. themselves, forms the starting point for the whole modelling enterprise.

Overall the approach to brain research by Arbib et al. has a Marrian flavor (Marr 1982). Their suggestion that different levels of analysis are useful to understanding brain function shows some resemblance to Marr's argument concerning the role of the computational, algorithmic, and hardware level theories. The main difference is that Arbib et al. show a commitment to the hardware or neural level (using Marr's terms) with a dynamical system theory approach to it. Here I agree with the authors that the approach of dynamical system theory together with the concepts of cooperative and competitive interaction can be fruitful for an interdisciplinary approach to cognition.

For example, recently Duncan et al. (1997) developed a new "integrated competition hypothesis" of visual attention, successfully bringing together numerous research disciplines. In their theory, delayed reaction time for attending to one object, if many objects are present in a scene, results, on the neural level, from competition between different brain systems, each responsible for a different relevances of objects for behaviour. In a computational underpinning of Duncan et al.'s theory with a dynamical systems theory approach, Heinke and Humphreys (1998) linked the competition process to physiological properties of neurons, showing that, indeed, competition between different brain regions can lead to delayed responses. However, here some deficits of the book have to be noted. The "integrated competition hypothesis" and other contemporary theories have stressed the crucial contribution of covert selective attention to human visual perception. Arbib, Érdi, and Szentágothai cover two aspects of visual perception: eye moments and visual scene interpretation. In the case of eye movements, experimental evidence supports the view that selective attention guides eye movements to new fixations by extracting peripheral information prior to the movement (see Hoffman 1998 for a recent summary). As for visual scene interpretation, the "window of attention" as put forward by the authors is nowadays considered an oversimplification. Looking at selective attention as an object-based process has gained substantial experimental support over the last years (see Baylis & Driver 1993 as one example of the experimental evidence). In this context, newer discussions of the involvement of the pulvinar in the control of selective attention (e.g., LaBerge 1998) are ignored by the authors.

A final point concerns the structure and organization of the book. The book is written in a lucid style, making easy reading of often difficult and technical topics. The references between each section and the repeated summaries will help readers coming from different interdisciplinary backgrounds. In addition, the division of the book into discrete sections makes it approachable for readers who are only interested in a subset of topics covered by the whole book. Two things lacking are an author index and a glossary of technical terms, which would be the icing on the cake.

Studies of synaptic elimination identify an intersection of neurocomputational and neurodevelopmental perspectives

Ralph E. Hoffman

Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06520-8038. **ralph.hoffman@yale.edu**

Abstract: In order to reach a better understanding of brain function, conceptual synergies linking empirical neurobiological studies and neurocomputational studies should be pursued. I describe an example of a potential synergy based on studies of neural network pruning. Simulations demonstrate that selective elimination of connections enhances the computational capacity of networks capable of temporal processing. These findings may shed light on the functional significance of postnatal neurodevelopmental pruning of cortical connections that occurs in mammals.

Neural organization reviews a large number of neuroanatomic, biological, and neural simulation studies in an attempt to mold a comprehensive view of mammalian brain function. Some of the concepts presented are undoubtedly right, while some are certainly wrong. How does one know if one is on the right track in this vast and still largely uncharted domain of observations, theories, and speculation? How does one know that one is paying attention to the right things?

My short answer to this question is that we will know when we are on the right track when a synergy emerges – when laboratory and human neurobiological studies seriously inform and constrain computational and simulation studies of neural systems and vice versa. Only then will theory and observation lift each other to another level, enabling us to say, finally, yes, now we understand. I don't think we are there yet. The scope of Arbib, Érdi, and Szentágothai's book demonstrates how elusive this synergy remains even as neurobiological and neurocomputational studies explode in number and complexity.

One example of a potential intersection of simulation studies of neural networks and empirical studies of mammalian brain systems pertains to the developmental process of network pruning. It is now well known that postnatal mammalian development of the cerebral cortex is characterized initially by extensive overelaboration of neuritic processes, that is, axons and dendrites, in the cerebral cortex. Subsequently there is a gradual reduction in connectivity, with synaptic density plateauing to about 60% of maximum levels (Huttenlocher 1979; Rakic et al. 1986). In humans, this process is largely complete by the second year in sensory areas, but is much more extended in prefrontal and probably other areas of association cortex and does not reach adult levels until midadolescence (Huttenlocher & Dabholkar 1997). Early in development it is likely that synaptic connections are created more or less randomly, with subsequent selective elimination based on environmental experience as well as endogenous factors. In adulthood, production of new synapses continues but is matched by a similar rate of synaptic elimination so that synaptic density of different cortical areas remains relatively constant.

Why the ubiquity of this developmental process? Physiological studies indicate that synaptic elimination reduces local metabolic requirements (Mata et al. 1980; Roland 1993). Along these lines, imaging studies of postnatal human brain development have shown downward shifts in local cerebral metabolic rates that parallel developmental shifts in synaptic density (Chugani et al. 1987). One wonders, however, if the only advantage of developmentally programmed cortical pruning is a metabolic one. To address this issue, we have undertaken two studies that examined effects of progressive elimination of network connections based on a Darwinian algorithm (i.e., weaker connections are selectively eliminated). We found, using two very different network architectures, that network pruning enhanced computational performance. In the first case, we used a recurrent, backpropagation network based in part on a prior simulation described by Elman (1990). The network was designed to learn grammatical rules implicit in sequential word presentation to disambiguate noisy inputs, a process that emulates some aspects of normal speech understanding. We found that selective pruning of recurrent connections dramatically enhanced the network's ability to utilize sequential expectations in processing inputs (Hoffman & McGlashan 1997). Estimations of the magnitude of synaptic elimination optimizing network performance closely matched adult reductions of synaptic density relative to peak childhood levels based on the postmortem study reported by Huttenlocher (1979).

These results paralleled an earlier study of a recurrent network designed to learn transition rules of finite-state automata reported by Giles and Omlin (1994). Their study also found that network pruning enhanced generalization performance of the network. A second study by our group utilizing an attractor network system was designed to emulate some aspects of semantic processing. We found that local pruning of connections in this architecturally distinct system enhanced its ability to access semantic categories (Siekmeier & Hoffman, submitted). Of interest is that all three network simulations were designed to process inputs in sequence rather than singly. This observation suggests that networks capable of temporal processing might preferentially benefit from selective pruning.

An additional payoff of our simulation findings is that they may provide insights into the pathophysiology of schizophrenia. There is now a growing body of evidence supporting the view that this disorder, which generally emerges during late adolescence or young adulthood, arises from excessive, developmentally-driven pruning of connections in association cortex (for reviews see Hoffman & McGlashan 1993; McGlashan & Hoffman, submitted). Our simulations demonstrate how symptoms characteristic of schizophrenia, such as hallucinated and disorganized speech, can arise from overzealous pruning of cortical networks during adolescence.

These simulation findings provide a compelling computational rationale for postnatal cortical pruning and raise important questions regarding the timing of pruning, the specific types of connections that are lost, and the relationship to learning capacity. Further efforts to delineate intersections of neurodevelopmental, neuroanatomical, and neurocomputational studies will, I believe, produce beneficial conceptual synergies. Although it is impossible for a book of the scope of the Arbib text to cover all relevant areas, I would suggest that the dramatic changes of brain organization that unfold during postnatal development is an area deserving additional scrutiny.

ACKNOWLEDGMENTS

Studies by the author are supported in part by NIMH grant MH-50557 and a NARSAD Independent Investigator Award.

Lifting the screen on *Neural organization:* Is computational functional modeling necessary?

Damian Keil and Keith Davids

Psychology Research Group, Manchester Metropolitan University, Alsager, Cheshire ST7 2HL, United Kingdom. **{k.davids; d.keil}@mmu.ac.uk**
www.mmu.ac.uk/c-a/exspsci/welcome

Abstract: Arbib et al.'s comprehensive review of neural organization, over-relies on modernist concepts and restricts our understanding of brain and behavior. Reliance on terms like coding, transformation, and representation perpetuates a "black-box approach" to the study of the brain. Recognition is due to the authors for attempting to introduce postmodern concepts such as chaos and self-organization to the study of neural organization. However, confusion occurs in the implementation of "biologically rooted" schema theory in which schemas are viewed as computer programs. The inclusion of an additional functional level between structure and dynamics is unnecessary in a postmodernist perspective of brain and behavior.

The explosion of research on the structure and function of the human nervous system during the past decade has produced a vast array of information, the synthesis of which constitutes a worthy feat. Arbib, Érdi, and Szentágothai have to be commended for such a comprehensive amalgamation of ideas and data from current research on neural organization. This text differs from others in the liberal mixture of theoretical modeling from modernist and postmodern scientific paradigms and a commendable stand against excessive reductionism. Description at the structural level is complemented with conceptualization at the systemic level, incorporating postmodern concepts such as self-organization and chaos in the nervous system. At the same time, modernist assumptions are made about a functional level of description invoking hypothetical constructs such as schemas, schema instances, and schema assemblages. The result is that the underlying tone of the book is modernist, with postmodern concepts woven into the theoretical fabric. It becomes clear that integration of the two philosophies has not been directly approached. Rather, an attempt is made to posit schema theory within a postmodern framework without reference to some of the thorny philosophical issues raised by mixing metaphors.

Psychology has been called the "last resting place" of the modernist scientific paradigm (Pickering 1997). The traditional modernist view of the brain as a computer can be traced from the

metaphysics of Socratic Greece and the Aristotelian symbolic categorization of the "essence of things," through to the Cartesian dualism of the eighteenth century, and the metaphysical materialism of Hobbes and Liebniz. The modernist approach has been influenced by the engineering and computer sciences in explaining processes of brain and behavior, exemplified by the software-hardware distinction in psychology, in which computation is assumed to be independent of the structural organization and evolutionary development of the nervous system. A weak form of computation is invoked by Arbib et al. in the dubious claim that schema-based modeling needs little commitment to the localization of individual schemas, and needs only to "be linked to a structural analysis as and when it becomes appropriate." (p. 33). On the one hand it is claimed that schemas are "biologically rooted," and yet it is also argued that "schemas are composable programs in the mind. A schema is like a computer program" (p. 41), ignoring the philosophical morass in implementing modernist and postmodern concepts in a unitary framework. What is lacking in the current account of neural organization is integrative modeling, in which functional and structural levels of analysis are mutually constraining.

In modernist science, problems occur due to the lack of an explicit, empirically verified interpretation of terms like "coding" and "transformation," and specification of the role of "computations" and "representations." Explicit definitions would avoid assumptions in the modeling process, captured metaphorically as "loans" and "mortgages" in the psychology literature. In some respects, the book does try to advance schema theory by placing it in a biological context, that is, neural anatomy and mechanism related to a particular schema. However, at critical points in the explanation, there is a reversion to modernist terminology and assumptions.

Superficially, a computer might seem a useful analogy for how brains perceive and act in information-rich environments. However, just because an algorithm can be written for a machine to act as a sensing device, this does not constitute an explanation for perception and action in biological systems. Despite the pervasiveness of the computational view, neurobiology has failed to provide reliable evidence for an architecture suited to symbolic manipulation and syntactical communication within biological nervous systems. For example, there is no existing evidence specifying the translation process between representational codes in biological systems. Another problem with computational accounts is that often several types of representations and several types of codes in the CNS are required, invoking the need for "translators" and "controllers," exemplified by the need for schema instances and schema assemblages in the current book. With the introduction of one hypothetical construct, that is, schemas, more are needed to clarify and qualify the function of the original representation, thereby perpetuating the need for additional system sublevels. The notion of schema is indeed recursive.

The problem is not just that the phenomenology of cognitive science has been inordinately influenced by modernism, but also that it has proceeded unconstrained by neuropsychology. The constraints imposed on cognition by other systems in the body, and the typical environments for action, have been ignored. Cognitions and intentions towards environmental objects need to be understood in a biological context, and are not static, discrete, and private affairs.

Postmodernism promotes the scientific goal of examining dynamic stability in natural phenomena. It is argued that all processes, including consciousness and intentionality, should be studied within natural order, not apart from it. Mental life in humans is viewed as having evolved from physical, biological and psychological processes, and not produced by special substances composing representations and codes. Postmodern frameworks for understanding neural organization encompass data and theory from psychology, biology, physics, and the neurosciences, rejecting the notion of the mind as a machine. Modernist concepts of "representations," "computation," and "coding" are rejected for an understanding of the stability and instability of pattern formation in complex, natural systems and a neo-Darwinian emphasis on exploration, assembly, variation, and selection in brain and behavior. thoughts, emotions, ideas, beliefs, images, and actions are merely the neural traffic constantly being produced between the billions of neurons in the CNS.

A postmodernist approach attempts to explain processes of brain and behavior within the boundaries of natural laws. Their delineation should not be screened by the unnecessary introduction of an additional layer of hypothetical constructs, which a modernist framework imposes. Lifting the screen could help us to fully understand brain, behavior, and, ultimately, consciousness.

Multiple personalities and views of neural organization

Rolf Kötter

C. & O. Vogt Brain Research Institute and Institute of Anatomy II, Heinrich Heine University, D-40225 Düsseldorf, Germany.
rk@hirn.uni-duesseldorf.de www.hirn.uni-duesseldorf.de/rk

Abstract: Neural organization has many facets; multiple descriptive levels and multiple analytical strategies coexist. Although most neuroscientists agree that a multidisciplinary, multistrategic approach is necessary to understand neural organization, diverse individual approaches make it difficult to find the optimal mixture and priority list.

Arbib, Érdi & Szentágothai's *Neural organization* is a conventional title for an unconventional book. Its scope is very broad and its views are quite specialized. This is no surprise to anyone familiar with the work of the three authors who are to be credited for their best intentions and their ceaseless efforts to promote systems approaches to the central nervous system.

The book touches many aspects that such a broad title may encompass – "structure, function, and dynamics" of biological as well as artificial neuronal networks, reaching from the single-cell level to a range of complex behaviors. It consists of two parts: the first introduces the authors' notions of structure, function, dynamics, and their integration; the second exemplifies their approaches in six specific brain regions, namely, the olfactory system, hippocampus, thalamus, cerebral cortex, cerebellum, and basal ganglia. The development of the chapters is quite systematic – to the limit of didactic overkill. Part I, in particular, requires a long breath to find, between preliminary remarks and multiple cross-references, a continuous stretch of text that has some contents to convey. Thus, the preface provides a general overview, followed by the overview of Part I, which gives overviews of the overviews to follow.

Within the structural overview, for example, very important aspects of neural organization are highlighted: an embryological perspective, the modular architectonics principle, and multiple models of modularity. These interesting and important aspects are sketched with unsatisfactory incoherence and brevity. Consider section 2.3.2 on the cerebral cortex as an example: we hear about staining methods revealing different aspects of cortical organization and are provided with some historical views of cortical modularity before learning that "the architecture . . . is more sophisticated in reality than could have been imagined earlier" (p. 30). The concluding diagram from Szentágothai 1993 (Fig. 2.17) is annotated to emphasize multiple levels of modularity, but it unfortunately lacks the clarity needed to convey a precise concept of cortical organization. Both theme and sequence recur in the corresponding chapter on cerebral cortex in section 8.1 of Part II, leading to a discussion of cortical modularity from Szentágothai's detailed view of neuronal connectivity.

In section 8.2, as another example, we take a big step to large-scale mathematical models generating modularity in the visual cortex: some detail is provided from the work of Érdi and col-

leagues on the development of ocular dominance columns in afferent-driven cellular sheets; models by others describing the formation of orientation columns are nicely contrasted for the different underlying principles. Other aspects of visual cortical organization are only hinted at or completely ignored (e.g., topological analyses of long-range connectivity by Young 1992, 1993, and the more recent work of his group). Among the unavoidable errors, the definition of forward and feedback connections on p. 223 as both terminating in layer IV of their target areas but originating from cell bodies in supra- and infragranular source layers, respectively, is particularly misleading in view of an ongoing discussion of whether delineation of retrogradely stained cell bodies is sufficient to establish hierarchical relationships (see Batardière et al. 1998).

By contrast, I am quite impressed by the breadth and level of discussion presented in Chapter 10 on the basal ganglia. For neuronal mechanisms, the authoritative chapter by Wilson (1998) in *The synaptic organization of the brain* is hard to beat, but the present chapter recalls not only fairly recent insights into the generation of up and down states in striatal medium spiny neurons, but also discusses experiments on motor imagery in Parkinsonian patients and basal ganglia models of reinforcement and sequence learning.

The appendix to the book introduces Neural Simulation Language as a convenient programming language for large-scale networks of single-compartment neurons, and Brain Models on the Web (BMW) as an online repository primarily for models written in this language. The URL given for BMW fails to produce the desired page, which at the time of writing seems to reside at http://www-hbp.usc.edu/Thrusts/bmw.htm instead. The available collection of models clearly requires further development by the Arbib group, but there is hope because many of them feature in the announcement of another book titled *Neural simulation language,* by Weitzenfeld et al. (2000). *Neural organization* was published in 1998 and ends with an index and a reference section that has only a few entries after 1995. Thus, we are looking back five years, which means that many details will have to be updated or revised. As a general remark, the print quality of many (in particular the anatomical) figures is not sufficient to show their intricate details – quite in contrast to the attractive dustcover design.

In my opinion, this book by Arbib et al. contains interesting but quite diverse material. The prominent personal components are not, and could not be, sufficiently glued together to serve the intention to "point out the advantages of a multidisciplinary, multistrategied approach to the brain" or to "offer a plan for the use of their methods" as stated on the dustcover. The lack of coherence of the material and the unequal style of exposing the topics causes an indigestion that is not ameliorated by the stitches of multiple overviews and cross-links. I suspect that it would have been better to publish the interesting historical material by Szentágothai (and the many references to other eminent Hungarian neuroscientists) separately, and to provide the framework and cross-linking in an electronically accessible format that lends itself better to accessing multiple threads. In addition, the book contains readable introductions to systems modeling and worthwhile reviews of previous modeling work, which help to clarify the state of the field and to foster systematic progress in building working models of the brain.

Meanings of "function" in neuroscience, cognition, and behaviour analysis

Julian C. Leslie

School of Psychology, University of Ulster at Jordanstown, Newtownabbey, Northern Ireland BT37 0QB. **jc.leslie@ulst.ac.uk**

Abstract: Different sciences approach the brain-behaviour system at various levels, but often apparently share terminology. "Function" is used both ontogenetically and phylogenetically. Within the ontogeny it has various meanings; the one adopted by Arbib et al. is that of mainstream cognitive psychology. This usage is relatively imprecise, and the psychologists who are sceptical about the ability of cognitive psychology to predict behavioural outcomes may have the same reservations about Arbib et al.'s cognitive neuroscience.

It is possible to maintain a "unified-science" view of the brain and behavioural sciences, and to assert, for example, that cognitive psychology, behaviour analysis, neuropsychology, neuroanatomy, ethology, and evolutionary biology are all studying the same system, but from different perspectives and using different levels of analysis. If this view is to be sustained, it is important that the concepts currently used in each discipline be regularly examined from the perspective of the other disciplines, in the hope of indicating meaningful linkages.

What do we, in the brain and behavioural sciences, mean by "function"? In the tradition of behaviour analysis, the term "function" has important and resurgent meaning. The function of a class of behaviours of an organism is described in terms of the outcomes, or changes in the environment of the individual organism, that it produces. Applied behaviour analysts have in recent years been particularly concerned with functional analysis as a behavioural assessment technique: it is desirable, and perhaps essential, to determine the current functions of classes of behaviour if successful plans are to be laid for changing the frequency of target behaviours (Leslie & O'Reilly 1999). For those concerned with developing behaviour analytic accounts of language and cognition (e.g., Barnes-Holmes et al. 2000), it is of central importance that operant stimulus classes and response classes are entirely functionally defined, and there is thus no reason to suppose that topographic similarity (that is, similarity of form) will be a guide to class membership.

There is a different but equally clear tradition in ethology of ascribing a phylogenetic meaning to the term "function." The ethological function of a class of behaviours lies in its utility for the species. Both behaviour analysis and ethology have "selectionist" views of behaviour and of the use of the term "function," but they relate to selection at different levels. Evolutionary biologists assert, and it widely agreed, that characteristics of species, including those of behaviour, arise and persist through the process of natural selection. Behaviour analysts assert that the behavioural characteristics of individuals arise and persist through the reinforcement contingencies that those individuals encounter in their environments. (This position is widely ignored, however, or misunderstood by those who approach psychology from other perspectives.)

What do Arbib et al. mean by function? They distinguish (e.g., p. 35) between "mental function" and "neuronal function," in the course of presenting a case for the use of schemata as functional entities that will enable them "to bridge from mind to neuron" (p. 41). Further details (of a representative account of schemata) are:

> A schema constitutes the "long-term memory" of a perceptual or motor skill or the structure coordinating such skills, whereas the process of perception or action is controlled by active copies of schemas called *schema instances.* . . . For certain behaviors, there may be no distinction between schema and instance; a single neural network may embody the skill memory and provide the processor that implements it. However, in more complex behaviors, the different mobilizations of a given "skill-unit" must be distinguished carefully. A *schema assemblage* is a network of schema instances, and its characteristics are similar to that of a single schema. (p. 41)

What type of explanatory system is proposed here? It is, in this form at least, an ontogenetic rather than phylogenetic one (although Arbib et al. make clear their concurrent interest in phylogeny in other important passages), and it thus belongs in the domain of psychological rather than evolutionary biological explanation. The authors themselves categorise it as cognitive neuroscience, which is generally seen as a close relative of cognitive psy-

chology. The notion of a schema as a purely psychological entity is of course one that has been current in cognitive psychology for a considerable time. As with other versions of cognitive neuroscience, the approach of Arbib et al. is explicitly neuroreductionist, and steps are taken towards specifying the links between the psychological processes, which have a function in the sense defined above within behaviour analysis, and the neural systems which support these processes.

The notion that it should be possible to identify neural systems that are functionally equivalent to behavioural processes has been enthusiastically pursued at least since Hebb (1949), and Arbib et al. provide an extended account of the neural architecture that may underpin a substantial number of important behavioural processes. [See also Amit: "The Hebbian Paradigm Reintegrated" *BBS* 18(4) 1995.] Their aspiration is to define function at the level of behaviour, using the terminology of cognitive psychology. In each area discussed they indicate what they believe to be the linked cognitive processes, and they give a central role to schemata. They conclude (e.g., p. 344) that minds consist of many schemata that interact in complex and hierarchical ways. From an ontogenetic perspective, the authors can therefore be classed as mainstream cognitive psychologists, and those of us who doubt the predictive value of much contemporary cognitive psychology will have the same types of concern about the psychological applicability of the work of Arbib et al. Enthusiasts for the use of the type of explanatory system favoured by cognitive psychology may retort that the types of neural model proposed by Arbib et al. can potentially provide an account of how and why schemata interact to determine particular behavioural outcomes. This may be the case, but, so far, the neural modelling approach provides hints about specific psychological processes rather than specific predictions.

Is schema theory an appropriate framework for modeling the organization of the brain?

Pietro G. Morasso

University of Genova, Department of Informatics, Systems, Telecommunications, 16145 Genova, Italy. **morasso@dist.unige.it**
www.laboratorium.dist.unige.it/~piero

Abstract: This review evaluates pros and cons of the schema theory as a general framework for expressing what Arbib et al. call "systems neuroscience." We discuss the software/hardware duality of the schema concept and the relative neglect of the mechanical properties of muscles. We propose a computational alternative to the functional decomposition in terms of schemas.

Neural organization is a bold and ambitious attempt to outline some of the common threads in the constant flux of neuroscience research. Readers with backgrounds in computer science who are not so familiar with the multifaceted complexity of computational neuroscience can find an inspirational and wide-ranging tour d'horizon of neuroanatomy, neurophysiology, neuroembryology, neurophylogeny, and so on. Readers with a more pronounced inclination toward the neurosciences are offered a framework for modeling brain organization and activities. A recurrent concept is that, despite undeniable advances, cellular and molecular neurosciences are ill-suited to capture the overall picture, but their reductionistic short-sightedness might find a guiding hand in what Arbib et al. call "systems neuroscience." Schema theory is proposed as a general framework for designing models and interpreting experimental results. However, this is the less convincing part of the book, even if one agrees with the premise, that is, the need of a nonreductionist brain theory on the side of cellular and molecular neuroscience.

A schema is a functional unit, based on the action-perception cycle. Hence "defined functionally, a given schema may be distributed across more than one brain region; conversely a given brain region may be involved in many schemas" (p. 45). In other words, the brain is viewed as a network of interacting schemas.

At this level of generality, nobody can disagree. The notion of a schema is simply a concise recapitulation of the diverse body of research in which, after the pioneering work of Anokhin (1974), Bernstein (1957), Gibson (1950), Piaget (1963), von Holst (1950), and others, it has become clear that perception and action are not independent functions but are always intermingled in complex and peculiar ways. However, this is not the stuff of a scientific model, which is supposed not only to recapitulate what is already known but also to set the stage for predictions and novel experiments that can in principle falsify the model. As a matter of fact, the authors admit that the schema concept is just a container whose content must come from other sources of information: "When a schema is provided with a precise functional characterization and a neurological localization, it sets the appropriate framework for neural network modeling" (p. 66). The point is that if we have a "precise functional characterization" of some task coming from a suitable set of behavioral experiments and a reliable "neurological localization" provided by electrophysiological/brain imaging/neurological/neuroanatomical evidence, what do we gain by rephrasing the evidence in terms of formal schemas?

At the philosophical level, the notion of schema is traced back by the authors to Kant and, indirectly, to platonic idealism that by its very nature favors pure over empirical knowledge and the immutable container over its volatile content. In terms of more modern scientific metaphors, a schema is a "software object" that can precede and in a sense be independent of the "implementation" in a specific experimental environment. Matching brain regions to schemas is then defined as a kind of "neuralization" of schemas (p. 234): neural models follow schemas and not the opposite. This explains the criticism of neural network modeling and connectionism (p. 33 and other parts of the book).

We might also find a parallel between the "software" nature of schemas and David Marr's approach toward perceptual-motor modeling, although the authors explicitly criticize Marr's well-known 2½-D primal sketch of early visual processing. Marr (1982) posits that in attempting to understand perception and action one has to follow a logical order by asking (1) *what* is the problem from the computational point of view, (2) *why* it needs to be solved, and only last (3) *how* would such a solution be "implemented" in the brain. One of the fundamental contributions of the neural network modeling revolution of the 1980s is just to negate the conceptual necessity or even the usefulness of the software/hardware distinction in computational neuroscience. Understanding the ways in which brain structures learn and process information is the necessary prerequisite for constraining functional models like Marr's primal sketch or Arbib's schemas in an empirically plausible way and not the opposite: neuralizing a functional model on the basis of empirical knowledge. In the connectionist way of thinking it still makes sense to conceive of schemas as emergent patterns, arising from the complex nonlinear dynamics of interacting neural networks coupled with the equally complex dynamics of the outside world. But this is a quite different kind of schema.

Probably the same bias toward an idealistic interpretation of neural processes, implicit in the schema theory, can justify the authors' neglect of the role of the mechanical properties of muscles (and of biomechanics in general) in the neural organization of action and the action-perception cycle. Curiously, this contrasts with their view, which I deeply share, that "the key for analyzing the brain . . . is . . . to understand how local interactions can integrate themselves to yield some overall results without explicit executive control" (p. 42). After the work of Bernstein (1957), Feldman (1995), Bizzi (1992), and many others it has become clear that the organization of motor patterns cannot be understood without taking into account the mechanical properties of muscles. One can agree or disagree with specific aspects of the many versions of mass-spring models, lambda or alpha models, and the like, but the basic idea is now well established. Internally generated patterns alone are insufficient to determine purposive ac-

tions in a reliable way, and mechanical compliance carries out a double computational function: it is at the same time a source of information on external dynamics and a tunable coupling device for allowing such processes to become "partners" rather than obstacles in the formation of complex sensorimotor patterns. In general, one can observe that the action-perception cycle implies a bidirectional flow of energy between biological organisms and their environments as well as a circular flow of information. Both flows are essential for the self-organization of purposive actions, but their interplay cannot be understood in terms of functional schemas alone.

The authors are well aware of the importance of self-organization. For example, they speak of cooperative computation as a pattern of strengthening the alliance between mutually consistent schema instances. But how this might be carried out in the actual neural machinery is not even hinted at. The problem is in the level of analysis. Choosing the software modular level (implicit, for example, in the notion of "instantiating" multiple schema copies for "schema assembly") automatically sets the insurmountable obstacle of the "credit assignment problem" that has motivated much of connectionist research in the first place. As a matter of fact, understanding the ways in which self-organizing processes shape the formation of brain structures and functions is clearly a work in progress; however, the key is to emphasize the constraining role of the nonlinear dynamics of the organism-environment interaction, on the one hand, and the interactions among neural assemblies as in the formation of cortical maps, on the other, rather than an abstract self-organization of schemas.

Braitenberg et al. (1997) have clearly identified the distinction, but also the deep complementarity between the brain representations of morphological-geometric objects, on the one hand, and dynamic-physical objects, on the other, suggesting that the cerebral cortex and the cerebellar circuitry, respectively, might be the candidate brain structures for the two types of computational functions. As a further clarification of the same theme, we wish to cite recent work by Doya (1999), who has pointed out that the learning paradigm, rather than specific perceptual or motor functions, characterizes in a unique way the computational nature of these two brain macro-areas, as well as their natural partners (the basal ganglia): (i) an unsupervised self-organizing paradigm in the cerebral cortex, for building cortical maps, (ii) a (self)-supervised paradigm in the cerebellar circuitry, for learning internal dynamical models, and (iii) a reinforcement-learning paradigm, in the basal ganglia, for learning sequential aspects of actions in complex tasks.

Regarding cortical maps, Arbib et al. appear to share the opinion of some neurophysiologists (summarized in a recent review paper by Rizzolatti et al. 1998) that beyond primary cortical areas somatotopic structure is hardly significant, because we observe a fragmentation of cortical representations of body parts and highly specialized neural clusters, such as "mirror neurons." This is deemed to rule out the possibility of unitary computational functions of associative cortical areas, such as the classic concept of "body schema." In fact, this view is consistent with the articulated and somehow fragmented nature of the schema assembly concept, but its Achilles' heel is the formation of coherent assemblies of cortical clusters (i.e., again the credit assignment problem). The authors do not take into account a computational alternative for explaining the apparent fragmentation while maintaining a common representation, which is outlined in Morasso and Sanguineti (1997): if cortical maps are trained to represent higher-dimensional manifolds, as is possibly the case in associative parietal areas, then the apparent fragmentation is a side effect of a hardware constraint (hosting an N-dimensional grid on a 2-dimensional substrate) but is compensated by using long-distance corticocortical connections (which are known to be present in a massive and organized manner).

Summing up, a feasible alternative to a "functional decomposition" of brain activity in terms of schemas is a "computational decomposition" in terms of maps, dynamic models, and sequences that is grounded in the organization of the neural hardware and exhibits functional regularities as "emergent properties." The difference between the two alternatives can also be appreciated in relation to the notion of intelligence, whose evolution is discussed in pages 260–61 of the book in relation to the "Great Move" idea of Newell (1990). [See also multiple book review of Newell's *Unified Theories of Cognition. BBS* 15(3) 1992.] His observation is that as biological organisms became more and more complex as a result of natural selection, learning an exponentially increasing set of specialized mechanisms became inefficient and might have prompted the "Great Move" of evolution: "establishing a neutral, stable medium that is capable of registering variety and then composing whatever transformations are needed" (p. 260). The core of the argument is the *neutral computational medium;* we suggest that from what we know about neuroscience it should be able to build internal representations of the outside world in terms of maps, dynamic processes, and sequences, independent of any functional schemas.

Synthetic approaches to cognitive neuroscience

Olaf Sporns

The Neurosciences Institute, San Diego, CA 92121. **sporns@nsi.edu**
www.nsi.edu/users/sporns

Abstract: Cognition and behavior are the result of neural processes occurring at multiple levels of organization. Synthetic computational approaches are capable of bridging the gaps between multiple organizational levels and contribute to our understanding of how neural structures give rise to specific dynamical states. Such approaches are indispensable for formulating the theoretical foundations of cognitive neuroscience.

Finding the neural correlates of mental or cognitive states is one of the foremost goals of cognitive neuroscience. But many theoretical and practical challenges remain, facing any meaningful attempt to forge connections between brain and mind. Two of these challenges are posed by the *levels problem.* First, nervous systems function across multiple levels of organization, ranging from molecular and synaptic events to the activity of cells and circuits, the integrated function of large-scale neural systems, and the behavioral activity of an organism in an environment. Clearly, no single level suffices to fully explain how brain and cognition interrelate. Processes at all levels interact. For example, synaptic changes give rise to modifications in network activity, which leads to changes in overt behavior. In turn, behavior influences what is sensed and experienced. and experience shapes synaptic and cellular processes within the organism's brain. Second, at any of these organizational levels, the conceptual linkage between structure and dynamics remains insufficiently understood. Obviously, neuroanatomy (structure) constrains very strongly the activity and correlational patterns (dynamics) of a given circuit. Yet, which anatomical patterns underlie specific patterns of functional connectivity (thought to be good candidates for correlates of cognitive or perceptual states) is still unknown. A variety of theoretical and computational approaches have been developed in recent years to address the challenges posed by the levels problem.

Biologically based computational approaches are ideally suited for bridging gaps between multiple levels of organization, as well as between structure and dynamics. Computational models encapsulate characteristic properties of neurons in simple equations and connect them in anatomically defined patterns. Such neuronal networks show temporal dynamics and may engage in synaptic plasticity or organize into coherent states. If connected to inputs and outputs, such networks may perform sensory processing or control motor behavior. Models that incorporate diverse neurobiological knowledge, from realistic synaptic rules to patterns of interconnectivity that mimic those of real brain areas, are *syn-*

thetic, in that they "put together" individual components across levels (Reeke et al. 1990). Their study provides the opportunity to correlate synaptic, cellular, network, and behavioral aspects of brain function, something that is hard (if not impossible) to do with real organisms. Synthetic neural models offer insights into what kind of higher-order phenomena can be generated given an explicit set of elementary rules and components.

A wealth of empirical information on the structure, function, and dynamics of neural systems has become available in recent years, much of it reviewed in *Neural organization.* A major accomplishment of this book is to have exposed the interrelationship between structure, function, and dynamics for all major functional subdivisions of the vertebrate brain. Particular emphasis is given to computational models of olfactory, hippocampal, cerebellar, and cortical networks (among others). In the past, computational approaches have scored some significant successes, but there have been disappointments as well. "Classical" connectionist networks with their highly stereotypic architectures and algorithmic learning procedures have proven to be inadequate as realistic models of brain circuits and processes. Network models based on attractors emphasize content-addressable memory and fixed-point dynamics and may be useful parallel or physical implementations of particular algorithms. However, their status as adequate brain models is questionable. There is very little experimental evidence suggesting that storage or retrieval of memories involve fixed-point attractors. Rather, brain dynamics (at least of cortical networks) appears to be characterized by dynamic transients, abundant spontaneous activity, and time-varying inputs rather than by attractors selected by constant input patterns. Brain responses occur on a time scale of tens to hundreds of milliseconds, while it may take considerably longer for a complex interactive system to "find" its attractor state from an initial condition. A more in-depth discussion of what the dynamical concept of "attractor" means in the context of realistic brain models or analysis of neural data sets would have been a useful addition to the book.

Neural organization does not concern itself much with statistical approaches to neural activity based on information theory (Rieke et al. 1997) or with approaches based on functional connectivity, correlation, or coherency analysis. However, investigating the relationship between the statistics of neural population activity and functional (e.g., perceptual) states has yielded significant insights. From a theoretical perspective, information-theoretical measures have been used to characterize global dynamical states of neural systems, capturing how information is integrated across multiple functional subdivisions of the cerebral cortex (Tononi et al. 1998). Analytical techniques based on information theory may prove useful in the analysis of multidimensional data sets from neurophysiology or neuroimaging. They may also be useful in relating anatomical and functional connectivity (Sporns et al. 2000), adding to our theoretical understanding of major aspects of cortical organization such as clustering of areas, conserved wiring, and reciprocity of pathways.

While the analysis of structure and dynamics is largely confined to the nervous system itself, the realm of function (visual scene analysis, sensorimotor coordination, reaching and grasping, and navigation) necessarily deals with an organism embedded in an environment. The organism's body structure, its movements, and the sensory content of its environment must be taken into account. *Neural organization* largely centers on Arbib's "schema theory" as a functional framework for expressing behavioral processes, allowing the mapping of functional relationships onto structural networks. In its emphasis on the active organism, continually engaged in action-perception cycles, schema theory is in agreement with more recent theoretical developments toward active perception and embodied cognition. These approaches, exemplified by recent neural (Almassy et al. 1998) and developmental (Thelen et al. 2000) modeling studies, link the structure and dynamics of the nervous system to the structure and dynamics of body and world. Behavioral, perceptual, and cognitive function are more fully understood when viewed from a perspective that takes into account body *and* brain, both subject to development and experience in the course of an individual history.

Arbib, Érdi, and Szentágothai write that a multilevel analysis of brain function "is a necessary complement to the reductionist program of empirical neuroscience" (p. 4). This notion is all the more important since the research program of reducing mental phenomena to molecular or cellular processes has shown no signs of succeeding. As a theoretical framework, reductionism is in perhaps terminal decline, as the complexity, interconnectedness, and integrated function of biological systems are increasingly recognized as central and irreducible problems of modern biology. Instead, computational approaches are on the rise, capturing and disentangling the many parallel interactions present in biological systems, including the brain. *Neural organization* points the way to a cognitive neuroscience based on a computational and synthetic view of how neural structure and dynamics interrelate across multiple levels.

Self-organisation or reflex theory?

George Székely

Department of Anatomy, University Medical School, Debrecen, Nagyerdei krt 98, Hungary. **szekely@chondron.anat.dote.hu**

Abstract: Neuromodelling is one of the techniques of modern neurosciences. The "at a distance" type of triadic synapse is probably the prevailing form of impulse transmission in many parts of the brain. If the genetically controlled cell-to-cell neuronal interconnections are abandoned, self-organisation may be the mechanism of structure formation in the brain. This assumption weakens the position of the reflex arc as the basic functional unit of nervous activities.

Contemporary research has accumulated such a large body of data and revealed such a high degree of complexity in the organisation of the nervous system that model building, realistic or abstract, is an indispensable complement to the armament of modern neuroscience. Fruitful modelling that helps us orient in the abundance of data and in the maze of complexity of nervous structures requires comprehensive knowledge, deep insight into the nature of the nervous system, and masterful handling of sophisticated modelling techniques.

I think the book under review is a fortunate combination of these conditions. The backbone of *Neural organization* is the nervous system. This is very good because it guarantees the continuous contact between models and the brain and introduces the brain from the fundamental principles to the outmost complexity of cortical structures. From an introductory chapter we learn that the nervous system consists of modules that can be separately studied. The formidable structures of the functional systems are stepwise unwrapped, and the text is complemented with self-explanatory drawings, which are themselves inviting for model studies. The presentation of the structure of the nervous system is not just a description, it is rather a humble confession of a person who spent his lifetime in intimate contact with the brain. This enabled Szentágothai to provide a "ready-made" brain for model studies. I would like to select two points for further comment from the rich selection of the book.

Szentágothai did pioneering work in describing synaptic glomeruli, also known as encapsulated synapses. An extra section is devoted to the question of impulse transmission in such complex synaptic arrangements. The activity of individual components cannot be investigated with the present techniques; their probable function can only be investigated in model studies. The simplest of complex synapses, the synaptic triad, is chosen for this purpose. Attributing chaotic behaviour to the synapsing elements, the synaptic triad acts as an input-dependent ON-gate in the visual system. I think the successful treatment of triads in model studies has special further significance from the point of view that serial

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

section electron microscopic studies reveal the "at a distance" type synaptic triads in great abundance in the spinal cord where the conventional "closely packed" triads are relatively rare (Székely et al. 1989). Sensory boutons are often surrounded by a number of presynaptic profiles, and boutons establish synapses with about twice as many dendrites. Many of the component elements are members of "at a distance" triadic arrangements, which are probably the prevailing type of synaptic engagements in many parts of the brain. The model presented can also treat interacting multiple triads, and subsequent studies may provide suggestions about the function of other complex couplings as well.

The most interesting paragraphs for me are those in which the eternal questions of how nervous structures evolve in phylogenesis and develop in ontogenesis are discussed. As long as one adheres to the precise cell-to-cell interconnections among neurons forming reflex arcs, allegedly the functional units of the nervous system, it is very difficult to find a rational answer to that question. On the basis of morphological complexity of neuronal interconnections, Szentágothai (1978) questioned the rigid deterministic structure of the brain more than 20 years ago.

As one comes to think of the triadic arrangement, the simplest of complex synapses, it is very difficult to conceive a kind of genetic determination mechanism that could guarantee that the "corresponding" synapsing profiles find one another in the throng of fibre terminals. Several years ago, Szentágothai (1978) suggested that the structural organisation may come about by some kinds of self-organisation mechanisms among neurons in tissue culture (see also pp. 14, 73). Probably because of its vague definition, self-organisation is not a generally accepted mechanism in the organisation of the nervous system, although it is emphasised in the first part of the book that complex dynamic systems are capable of self-organisation. In the book there are a number of structures in which this possibility is pointed out. The most beautiful example is the formation of modules in the basal ganglia in which the peculiar form of the component neurons almost present themselves to get organised "spontaneously" into a structure. In a quantitative neuromorphological study characterising the dendritic geometry by thirty-two numerical variables, we succeeded in showing a strong form-function relationship in the fifth and seventh motor nuclei (Matesz et al. 1995), and simple qualitative observations indicated similar relationship in the spinal motor column (Székely 1976).

It seems that neuronal morphology plays a significant role in the organisation of neural structures. The assumption of self-organisation in the brain has an important implication. The formation of interacting reflex arcs cannot be conceived without the genomic determination of "correct" neuronal interconnections, and genetic determination leaves a very narrow margin for self-organisation. I think the book, especially the last chapter, can be a good starting point for revisiting the classical reflex theory.

Difficulties with synaptic theory of learning and memory and possible remedies

Mikhail N. Zhadin

Laboratory for Neurocybernetics, Institute of Cell Biophysics, 142292 Pushchino, Moscow Region, Russia. **zhadin@online.stack.net**

Abstract: The absence of a clear influence of an animal's behavioral responses to Hebbian associative learning in the cerebral cortex requires some changes in the Hebbian learning rules. The participation of the brain monoaminergic systems in Hebbian associative learning is considered.

Neural organization by Arbib, Érdi, and Szentágothai is a unique encyclopedia of basic achievements in theoretical neurobiology. One of the main subject matters of the book is analysis of a wide complex of experimental data and basic lines in development of the synaptic theory of learning and memory. A series of sections in the book is devoted to these questions. However, despite the success of this theory, one cannot but infer that it has serious difficulties. The capacities for learning and the memory are evolutionally formed mechanisms for adapting the behavioral responses of the whole organism to the perpetually changing milieu in which some elements of previous situations vital to the animal occasionally occur. In learning, it is important that information about the usefulness or harmfulness of an animal's behavioral responses should continually flow into the brain. The information manifests itself as positive (PR) or negative reinforcement (NR). This process is an obligatory element in any model of learning. Unfortunately, in most models of associative learning, this function is not considered in its entirety.

Reinforcement is most often introduced in the form of an unconditioned input in Hebbian associative learning at the level of a single neuron. But in the real brain, the learning of complicated behavioral responses to complicated polymodal stimuli embraces a great many neurons and interneuronal connections. In these conditions, unconditioned inputs originating from unified centers of PR and NR must terminate on the overwhelming bulk of neurons to provide a wide variety of conditioning forms. Similar polyneuronal inputs are revealed in the cerebellum in the form of climbing fibers (Arbib et al., sect. 9.3) and in the basal ganglia in the form of dopaminergic fibers (sects. 10.3 and 10.4), but the role of climbing fibers in the reinforcement system is not quite so evident.

Nothing of this sort is seen in other brain structures, however, including the neocortex and hippocampus. It is also worth noting the strange finding that many Hebbian synapses are detected in the hippocampus, exhibiting the most primitive form of memory (habituation), and in the developmental neocortex, but they are very seldom seen in the mature neocortex, which is responsible for the highest mental activity: recognition and complicated forms of memory and learning. This suggests that there may be something wrong in either the way we are searching for them or in the concept of Hebbian synapses. Monosynaptic LTP (sects. 4.5 and 6.3) is unlikely to have anything to do with the learning at all, having no association with the reinforcement process.

The following features of a multiple synaptic input of the reinforcement could be stated a priori: (1) These inputs must terminate on an essential population of neurons and must originate from a single center of PR or NR. (2) Their chemistry and postsynaptic receptors must be elsewhere than in internal connections inside the structures where the memory traces are formed, and the chemistry of the PR input must differ from that of the NR input. (3) For each Hebbian associative neuron, this input should have a possibility to influence all the multitude of synapses on the neuronal membrane, no matter how far they are located from this input synapse.

Among all brain structures, such features are inherent only in the monoaminergic systems (MS) (Ungerstedt 1971): the serotoninergic one (SS) coming from the raphe nuclei, the norepinephrinergic one (NS) from the locus coeruleus, and the dopaminergic one (DS) from the substantia nigra. Fibers of these systems are widespread over all diverse brain structures; moreover, synapses with such chemistry never occur among internal interneuronal connections in the structures where the MS endings exist.

What is most interesting is the unique structure of the MS synapses in the neocortex, hippocampus, and other laminar structures (Beaudet & Descarries 1978). These synapses are simple varicosities including synaptic vesicles without any postsynaptic parts, which just saturate the intercellular space with the appropriate monoamine for a time until its inactivation by monoamine oxidase. So monoamines have free access to all other synapses upon activation of the appropriate monoaminergic centers.

In several laboratories in the mid-1970s (Libet et al. 1975; Freedman et al. 1977; Zhadin 1977) there independently arose an idea that it is the MSs that convey the information on usefulness or harmfulness of animal's behavioral responses; they constitute

an essential part of the reinforcement system and change the efficacy of active synapses according to appropriate rules. Gromova (1980) showed that destruction of the raphe nuclei impedes learning with PR (food), but does not affect learning with NR (pain). Destruction of the locus coeruleus, conversely, impedes learning with NR, not affecting learning with PR. Hence it follows that SS mediates PR, and NS mediates NR.

This contradicts the view of Libet et al. (1975) and Freedman et al. (1977) who consider the NS to mediate PR according to Olds and Olds's (1963) experiments with autostimulation. Olds used rather high intensity stimulation, hence they had no reliable way to localize nuclei as small as the raphe nuclei and locus coeruleus.

In our mathematical model (Zhadin 1987; 1991; 1993; Zhadin & Bakharev 1987), the electrical activity of cortical neurons during prolonged action of PR or NR was studied. It was shown that under these conditions the neurons must manifest typical and well-distinguishable patterns of their activity. With PR, the neural excitation tends toward its highest level upon initial relatively high excitation or to complete inhibition upon initial low excitation with possible accidental transitions from one extreme level to another. With NR, it tends to some intermediate excitation level, regardless of its initial level. These conclusions have been verified in experiments (Mamedov 1987; Zhadin 1987; 1993) involving a long-term increase in concentration of serotonin or norepinephrine in the intercellular medium with direct application of these monoamines to the neocortex. These patterns in neuronal activity occurred: serotonin induced the patterns typical of PR and norepinephrine of NR. In experiments (Zhadin & Karpuk 1996) on neocortical tissue slices, increase in the concentration of serotonin in the intercellular medium caused a rise in the efficacy of relatively highly activated synapses and a fall in the efficacy of less activated ones.

In light of all the foregoing, some modification of Hebbian learning rules seems to be necessary. Two changes might be possible:

(1) An increase in intercellular concentration of serotonin or dopamine facilitates Hebbian heterosynaptic associative learning. An increase in intercellular norepinephrine prevents Hebbian associative learning or facilitates anti-Hebbian learning. According to this alternative, Hebbian or anti-Hebbian synapses in the neocortex or other brain structures should be sought, along with natural or artificial introduction of the appropriate monoamine into the intercellular medium.

(2) Monoaminergic influence on neuronal membrane receptors could play the role of a direct reinforcement input. Arbib et al.'s section 10.4 describes a model and experimental evidence showing that usual dopaminergic synapses provide unconditioned inputs in Hebbian associative learning. One might expect that in the neocortex and other laminar brain structures the similar synapses would have more complicated design with a presynaptic component in the form of a host of varicosities on all proper monoaminergic fibers within this structure, with a postsynaptic component in the form of a magnitude of corresponding monoamine receptors on the whole membrane of the neuron, and with a synaptic cleft in the form of the entire intercellular space of the structure saturated by the monoamine for a short time. Here serotonin or dopamine would mediate PR, as norepinephrine does NR.

Perhaps direct experimental tests on this hypothesis will reveal new lines in the development of the synaptic theory of learning and memory.

Authors' Response

Organizing the brain's diversities

Michael A. Arbib[a] and Peter Érdi[b]

[a]*USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520;* [b]*Department of Biophysics, KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, H-1525 Budapest, Hungary.* **arbib@pollux.usc.edu**
www-hbp.usc.edu/erdi@rmki.kfki.hu
www.rmki.kfki.hu/biofiz/biophysics.html

We clarify the arguments in *Neural organization: Structure, function, and dynamics,* acknowledge important contributions cited by our critics, and respond to their criticisms by charting directions for further development of our integrated approach to theoretical and empirical studies of neural organization. We first discuss functional organization in general (behavior versus cognitive functioning, the need to study body and brain together, function in ontogeny and phylogeny) and then focus on schema theory (noting that schema theory is not just a top-down theory and discussing the transition from action-oriented perception to cognition). We then turn to dynamical organization, with a focus first on neural modeling and dynamics (clarifying the multiple functions of neurons and brain regions, and looking further at various forms of dynamics) and second on learning, development, and self-organization (looking at monoaminergic systems, reinforcement, self-organization, postnatal development, and disease). We close with a brief philosophical discussion of postmodernism and reductionism.

R1. Functional organization

R1.1. Principles for brain theory

Two of our principles of neural organization are listed by **Borisyuk.** (1) Principle of modular architectonics – the nervous system is composed of building blocks; and (2) Principle of topographical order representation of sensory information in primary sensory cortices and a number of other structures. It is not clear whether the principle of topographical order can be used for olfactory sensory information, but for the auditory system we have tonotopy in auditory cortex and a spatial map of auditory location in superior colliculus. Some other principles of information processing in the brain noted by **Borisyuk** are the principle of synchronization of neural activity; the principle of spatiotemporal information encoding; and the principle of flexibility of neural circuits (the same neural elements participate in the implementation of different functions). In *Organization,* other principles include the action-perception cycle, the existence of collective population phenomena to be described by statistical dynamic theories, the occurrence of hierarchical and recurrent connections, and self-organization. We agree with **Borisyuk** that the list is far from complete.

Székely applauds the way in which (in Ch. 2) the formidable structures of the functional systems are unwrapped stepwise by Szentágothai with the text complemented by self-explanatory drawings that invite model studies. But this applause shows up one of the book's shortcomings – its

models are not explicitly based on Szentágothai's drawings. Nonetheless, these drawings offer much to inspire the brain theorist.

Edelman notes Marr's (1981) concern regarding the prospects of developing comprehensive, analytically explicit ("Type I") theories of some visual submodalities, the alternative being the idea that these may turn out to be describable only as bags of unrelated ("Type II") tricks. Our search for general organizational principles may suggest a commitment to grand Type I theories, but in fact our principles summarize general perspectives on neural organization, which can be used to constrain theorizing, rather than offering logical axioms from which neural organization can be rigorously inferred. Nonetheless, as we chart the immense diversity of regions and cell types within the human brain, let alone the incredible diversity of specialized brain structures revealed by neuroethology, it is our hope that the use of such principles will show that the "bags of tricks" are far from unrelated, and that brain theory will employ an increasingly powerful set of principles and theories to provide a taxonomy in which many of these tricks can be related to yield a general understanding.

Marr (1982) posited that in attempting to understand perception and action one has to follow a logical order by asking (1) *what* is the problem from the computational point of view, (2) *why* it needs to be solved, and only finally (3) *how* such a solution would be "implemented" in the brain. **Morasso** seems to believe that our approach to schema theory follows this line, for he thinks it a criticism of *Organization* when he states that

> One of the fundamental contributions of the neural network modeling revolution of the 1980s is just to negate the conceptual necessity or even usefulness of the software/hardware dualism in computational neuroscience. Understanding the ways in which brain structures learn and process information is the necessary prerequisite for constraining functional models like Marr's primal sketch or Arbib's schemas in an empirically plausible way and not the opposite: neuralizing a functional model on the basis of empirical knowledge.

Heinke understands that we espouse a more subtle approach than Marr's "logical order," when he notes our commitment to the hardware or neural level (using Marr's terms) with a dynamical system theory approach to it. He agrees with us (and **Morasso!**) that the approach of dynamical system theory together with the concepts of cooperative and competitive interaction can be fruitful for an interdisciplinary approach to cognition. We shall have more to say on this topic below.

R1.2. Behavior versus cognitive functioning

Foster states that "The emphasis on behavior rather than cognitive representations is also mirrored at the anatomical level, with more coverage of subcortical structures than of the cerebral cortex." It *is* true that *Organization* provides relatively limited coverage of such higher cognitive capacities as attention, executive functioning, social cognition, and the many aspects of memory and language, but Chapter 11 is very much concerned with showing how the approach to neural organization, developed throughout the book, paves the way for new developments in cognitive neuroscience. However, the book makes clear that it is a mistake to associate cognitive capacities with cortical to the exclusion of subcortical structures. All too many researchers within human neuropsychology neglect the roles of neurochemical systems or subcortical mechanisms, of which the latter receive especial emphasis in *Organization*.

Foster suggests that cognitive neuropsychology has made tremendous strides in delineating the "architecture of cognition" and in "carving cognition at its joints," so that the cognitive mechanisms underlying processes such as attending, reading, speaking, remembering, and identifying are now quite well specified (see, e.g., Ellis 1996; McCarthy & Warrington 1990; Parkin 1999). He notes that the cognitive neuropsychology school has placed particular emphasis on functional dissociation as the methodology of choice for identifying specific functional units (or "modules") subserving particular cognitive and computational processes. The ways in which distinct functional modules map onto physiology and anatomy provides an ongoing challenge. In *Organization* we were at pains to distinguish functional units – *schemas* – from structural units – *modules* – and to stress that a given schema may rest on the cooperative computation of many modules, while a given "cognitive behavior" may in turn rest on the cooperative computation of many schemas. All too often, "carving cognition at its joints" yields entities that have no neural reality, especially in those topics related to language and planning, with over-interpretation of human brain imaging coupled to computationally incomplete models that have little relation to the findings of animal neurophysiology. The discipline of iteratively reformulating a schema analysis in the light of varied neuroscience data is crucial to the future success of cognitive neuroscience.

Section 11.3.1 (Memory, Perception, and Intelligence) of *Organization* articulates our theoretical framework for mind-brain function. The section is very brief considering its strong title. In his approach to cognitive architecture, Anderson's (1983) main technique is spreading activation which, going back to Collins and Quillian (1969), was considered as a general mechanism for the dynamics of the retrieval processes. The dynamic model for spreading activation has been formulated based on some "quasineurological" allusion. In this spirit Érdi et al. (1992a) used the concept of a "double architecture," for example, memory systems: namely, neural and mental architectures. Connectionist models have evolved in different levels (i.e., both at the neural and mental or cognitive) of natural information processing systems, as well as in artificial intelligence researches. Psycholinguistics has been an important field of connectionist applications. Word perception (Rumelhart & McClelland 1982; 1986) and word/sentence production (Dell 1986) are illustrative early examples. High-level cognition systems have been simulated with varying degrees of neural constraint by "production systems," such as SOAR (Newell 1990), a new version of the spreading activation hypothesis (Anderson 1993), and the different versions of CAPS (Just et al. 1999). Although the state of the art now is to connect brain imaging data to cognitive level modeling, there is a clear missing link between conventional neural modeling and data produced by brain imaging; a link that our Synthetic PET technique seeks to supply (see below).

Amos & Wynne, in discussing how to avoid the conflation of executive function and frontal lobe function, stress that *Organization* provides a structure that could be used to motivate a principled integration of multiple levels of

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

neuropsychological data, models, and theories. The structure both clarifies previous endeavors, and identifies valuable avenues for future investigation. However, they lament that the gaps between the levels are still all too evident even in our own work (e.g., in Ch. 6 on the hippocampus). They suggest [PCC] that in the hippocampus chapter integration occurred in limited domains, whereas most of the discussion seemed to indicate fairly strict separation of levels in the reported research. **Foster** sees little consideration in this chapter of the ongoing debate between researchers who regard the hippocampus as being specifically involved, from a cognitive perspective, in the mediation of recall memory (e.g., John Aggleton) and those who argue that it has a more wide ranging role in subserving memory, encompassing both recall and recognition (for example, Larry Squire), or of the contentious role of the hippocampus in the phenomenon of retrograde amnesia. By contrast, **Edelman** found the chapter on the hippocampus especially illuminating. For example, he finds the possibility of the involvement of the hippocampus in the representation of relational information (discussed on pp. 182–84) fascinating, in view of the present attempts of brain theorists to address the so-called problem of compositionality (Bienenstock & Geman 1995).

Here, the Précis goes beyond the book, since it both updates the sampling of data, and summarizes our own recent work on "schema-level" analysis of the hippocampus as part of a far larger brain system involved in place cells, cognitive maps, and navigation (Guazzelli et al. 1998; see Redish 1999 for related material). Other models worthy of mention include Rolls and O'Mara (1993) and Treves and Rolls (1994), as well as the study by Traub et al. (1999) of the hippocampal and cortical fast gamma oscillations. As for the role of hippocampal regions in encoding/retrieval of memory, there are ongoing debates on the eventual division of labor between the caudal and rostral portions of the hippocampal formation. Lepages et al. (1998) reported that, at least based on PET studies, the rostral regions are more active in the encoding process, while the caudal portions take part predominantly in retrieval processes. Results deriving from fMRI recordings (Schacter & Wagner 1999) seem to be partially contradictory to these findings and interpretations.

Borisyuk states that, consistent with the anatomy of different brain structures in the book, the "neural network with a central element" (Kazanovich & Borisyuk 1999) is a very common architecture of connections between the brain structures. For **Borisyuk** (PCC), the thalamus is a central element according to the information distribution in the cortex – information from the thalamus goes to the cortex by divergent connections and comes back to the thalamus by convergent connections. To this we would just add that the thalamus has many nuclei, each with very different patterns of connections with cerebral cortex and other regions (*Organization*, Ch. 7). Thus we might argue that different nuclei of thalamus serve as a central element for different neural networks.

R1.3. Body and brain together

As **Sporns** notes, the realm of function (visual scene analysis, sensorimotor coordination, reaching and grasping, and navigation) necessarily deals with an organism embedded in an environment. The organism's body structure, its movements and the sensory content of its environment must be taken into account. In its emphasis on the active organism, continually engaged in action-perception cycles, **Sporns** sees *Organization* as in agreement with recent work on active perception and embodied cognition exemplified by recent neural (Almássy et al. 1998) and developmental (Thelen et al. 2000) modeling studies, which link the structure and dynamics of the nervous system to the structure and dynamics of body and world. For example, Almássy et al. (1998) have modeled the generation of translation invariant and object selective properties of inferior temporal neurons. The neural model was embedded in a mobile robot equipped with a camera. The continuity of the visual input due to self-generated movements of the robot was crucial in setting up an architecture within IT that resembled the one found in adult primates. If the continuity (the "movie-like" quality) of the visual input stream was disrupted (without changing the input patterns themselves) no translation invariance developed. This is a case where being "embodied" and active in an environment actually made a difference to the nervous system.

However, **Morasso** takes us to task for our neglect of the role of the mechanical properties of muscles (and of biomechanics in general) in the neural organization of action and the action-perception cycle. **Morasso** (PCC) notes that all the evidence is that motor patterns for locomotion (walking as well as running) are tuned to the body-environment interaction (the pendular movements due to gravity), so that the mechanical properties of the musculoskeletal system are fundamental for the "neural organization" of the skill (Inman et al. 1981; Morasso et al. 1999). His slogan is that external physics is a co-processor of the computational processes from which the skill unfolds. We agree completely. *Organization* has provided an advance on most overviews of brain function in that it places great emphasis on the motor system from the start (note the frequent discussion of MPGs, motor pattern generators), but **Morasso** is correct that it would be salutary to look more at muscle models and spinal circuitry, for the brain does not act on kinematics, but on a dynamic system of a body interacting with the world, and these dynamics are not experienced directly but rather through the medium of brainstem and spinal circuitry. As **Morasso** (PCC) further observes, mechanical compliance is at the same time a source of information on external dynamics, and a tunable coupling device for allowing such processes to become "partners" in the formation of complex sensorimotor patterns. In general, one can observe that the action-perception cycle implies a bidirectional flow of energy between biological organisms and their environments as well as a circular flow of information. All this provides an excellent challenge for another chapter, perhaps with the title of the paper by Chiel and Beer (1997), "The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment." In this regard, we need more computational models that show how the structure of NNs may change radically (as distinct from some weight tuning) when "dynamic coprocessing" is taken into account. However, there may be cases where the dynamics makes the brain's work harder than it might be. For example, our recent work on cerebellar involvement in control of arm movements pays particular attention to the fact that interaction forces mean that one cannot treat the control of each joint kinematically, but must take these dynamic interactions into account (Schweighofer et al. 1998a; 1998b).

R1.4. Function: Ontogeny and phylogeny

Leslie argues that our approach to "function," at least as embodied in schema theory, is an ontogenetic rather than phylogenetic one, and thus belongs in the domain of psychological rather than evolutionary biological explanation. Similarly, **Edelman** notes that *Organization* most often takes "function" to mean "how does this bit of the brain function?" rather than "what does this bit of the brain do for a living?", making no clear distinction between explanation of operation and explanation of goals and means. *Organization* (sect. 11.2.2) states that "The issue for the brain theorist . . . is to map complex functions, behaviors, and patterns of thought either on the interactions of these rather large entities (anatomically defined brain regions) or on these very small and numerous components (the neurons)." **Edelman** feels that this would constitute an incomplete theory of the brain, unless it includes an explanation of its function (over and above its operation), and that the theoretical concepts that *Organization* brings to bear on this issue – dynamics, self-organization, and schemas – are inherently incapable of filling this gap.

We think that **Edelman** is partly right. Each brain is partly the way it is because we are the products of evolution, and thus a complete brain theory should include an account of how the brain got to be the way it is. This would include an account of how parts of the brain have evolved within a system for achieving a certain goal, and would help us understand to what extent the brain structure was optimal, and to what extent its structure is suboptimal because evolutionary "design" is more a matter of tinkering (bricolage) than of global analysis. *Organization* does discuss evolution – but only briefly – and it would certainly be valuable to further explore this topic within the framework that *Organization* provides. Materials for such an exploration could include the work of Allman (1999), Kaas (1993), and Krubitzer (1998), as well as our current attempts to understand the evolution of language (Arbib 2000; Rizzolatti & Arbib 1998).

We think that **Edelman** is partly wrong. Much of the structure of the adult brain is shaped by patterns of developmental "self-organization," and by learning from experience. The former may shape structures in ways that reflect general constraints on "growing a brain" more than they are related to the function of the region. The latter helps us understand how circuitry may become tuned to better meet a range of goals. *Organization* has much to say about both of these (see, e.g., sects. 4.4 and 8.2 for the self-organization of modular architectonics of visual cortex; and sects. 4.5, 6.4, 6.5, 8.5, 9.3, 9.4, and 10.5 for a variety of approaches to learning).

Amos & Wynne (PCC) suggest that comparative neuroscience and developmental neuroscience are equally relevant in reaching understanding of the links between structure and function because the emergence of new functions in new structures will depend on the functions performed in older structures. Disease research (more on this below) seems useful as a means of testing hypotheses generated by such integration. They also note that it may be impossible to control for either different "social experiences" or for reorganization of damaged neural tissue. Our limited knowledge of chaos theory has left us with the dim apprehension that in complex dynamic systems such as the social networks of evolving animals and the developing brains of multicellular creatures, small disturbances can cause large and unpredictable changes in equilibrium states, making attention to the dynamic level necessary. This accords very well with the point made in *Organization* that integration of multiple levels is necessary.

As **Leslie** notes, there is a tradition in ethology of ascribing a phylogenetic meaning to the term "function." The ethological function of a class of behaviors is in terms of its utility for the species. Evolutionary biologists assert that characteristics of species, including those of behavior, arise and persist through the process of natural selection. (This is true in part. Much of what makes us human today is the result of cultural, rather than biological, evolution.) Behavioral analysis assert that the behavioral characteristics of individuals arise and persist through the reinforcement contingencies those individuals encounter in their environments. This characterization of behavioral analysis is too rigid. Am I different from a cat because I have a different genetically predisposed brain-body matrix, or am I only different because I experience different reinforcement contingencies than a cat does? We would seek (and **Leslie** [PCC] agrees) a "joint solution" in which the "genetically predisposed brain-body matrix" interacts, from conception on, with the environment. Those interactions change, or become part of, the genetically predisposed brain-body matrix, and thus this interaction is a very strong one.

Leslie sees *Organization* as exemplifying "a type of cognitive approach which I feel is bound to fail because of its imprecision." But consider, for example, the work of Goodale and Milner (1992) (*Organization,* sect. 8.4.3) on the role of parietal cortex in setting parameters for motor control. Is that a failure? If so, will conditioning studies be the right path to precision or (as indeed **Bridgeman** suggests) will more precise psychophysics solve the problem? **Leslie** notes that *Organization* provides an extended account of the neural architecture that may underpin a substantial number of important behavioral processes. Yet, he then states that "From an ontogenetic perspective, [Arbib et al.] can therefore be classed as mainstream cognitive psychologists, and those of us who doubt the predictive value of much contemporary cognitive psychology will have the same types of concern about the psychological applicability of [*Organization*]." Leslie (1996) offers examples of what, presumably, are more satisfying neural approaches for behaviorists, such as the

> remarkable effects . . . demonstrated by Stein, Xue and Belluzzi (1993). Using a slice of hippocampal tissue removed from the brain of a rat, they found that a type of neural activity (pyramidal cell bursting responses) could be reinforced with applications of dopamine, a neurotransmitter, to the slice of brain tissue. This was the result of operant conditioning.

But this very example seems to show the strength of *Organization* rather than weakening it. For example, in Chapter 9 on the cerebellum, we progress from the role of cerebellar circuitry in classical conditioning to its role in the adaptation and coordination of movement to close with an all too brief look at "The Cerebellum and Mental Activity." More generally, the book gives a variety of well-grounded case studies of the brain's varied learning mechanisms. In short, we provide a bridge from the controlled but cognitively bland studies of the strict behaviorist through controlled semicognitive behaviors to the general prospectus for a cognitive neuroscience rooted in schema theory, dynamics, and anatomy provided in the last chapter of *Organization.*

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

R1.5. The worldwide web

Gouras suggests the interest of exploring analogies between the brain and the evolution of the worldwide web as another example of how standardized sequences can be linked in a universal way to create complex abstractions.

Kötter found that URL given for the Brain Models on the Web of the USC Brain Project (which Arbib directs) fails to produce the desired page while **Edelman** found that some links were dead, while others pointed back to the front page. Arbib is happy to report that much progress has been made by the USC Brain Project since *Organization* went to press, and the website for the Brain Project is now greatly improved even since the time when **Edelman** wrote his commentary. However, since the website is still in a state of flux, we suggest the reader turn to the home page at *http://www-hbp.usc.edu,* rather than depending on the stability of links to specific pages. An overview of the work of the USC Brain Project, Arbib and Grethe (2000).

R2. Schema theory: The next step

R2.1. Schema theory is not just a top-down theory

For **Morasso,** the notion of schema is "simply a concise recapitulation of the diverse body of research in which, after the pioneering work of Anokhin (1974), Bernstein (1957), Gibson (1950), Piaget (1963), von Holst (1950), and others, it has become clear that perception and action are not independent functions but are always intermingled in complex and peculiar ways. However, this is not the stuff of a scientific model." Our first response is that schema theory has two senses: (1) a general framework like differential equations, and (2) a set of specific models expressed in that framework, such as, to take a specific example from Chapter 3 of *Organization,* the Hoff-Arbib model of reach-grasp coordination. The former is useful to the extent that it helps one express the latter; it is the latter that meet the test of experiment. However, at either level, the schema theory of our book goes well beyond the work of **Morasso's** progenitors. We shall say more about Anokhin in the next section. Arbib (1984) pays explicit tribute to Bernstein's notion of synergy, but also notes that it lacks many of the crucial element provided by the explicit notions of "motor schema" and "coordinated control program" of our schema theory. Gibson (1966) gave us superb intuition into the diversity of information that we can "pick up" from the environment, but he resolutely spoke of "direct perception" and it has taken much work, including our own, to relate his insights to plausible computational (Prager & Arbib 1982) and neural (Arbib 1997) mechanisms. Piaget emphasized the role of "assimilation" and "accommodation" in developing a form of schema theory to describe his observations on the child's progression from infantile behavior to logical thought, but emphasized "one schema at a time" rather than the crucial role of multiple interacting schemas in giving a processing account of the child's behavior. Thus the schema theory offered in *Organization* is far more than a recapitulation of the work of the pioneers – it pays homage to their work, yes, but also adds to the thrust of each of them within a new and powerful framework which makes far closer contact with modern neuroscience than they would have aimed for.

As noted earlier, **Morasso** is mistaken in his claim that a schema is a "software object" that can precede and in a sense be independent of the "implementation" in a specific experimental environment. Yes, matching brain regions to schemas is defined as a kind of "neuralization" of schemas (p. 235) but section 3.2.1 is explicitly designed to counteract the view that "neural models follow schemas and not the opposite." The section shows that an initial functional decomposition when linked to a "neural localization" of specific schemas can yield hypotheses that can be empirically tested by neuroscientific experiments, and that this may lead to explicit restructuring at the schema level. **Morasso** suggests that a feasible alternative to a "functional decomposition" of brain activity in terms of schemas is a "computational decomposition" in terms of maps, dynamic models, and sequences that is grounded in the organization of the neural hardware and exhibits functional regularities as "emergent properties." **Edelman** seems on somewhat the same wavelength as **Morasso** when he asserts that "if you want your flock of schema[s] to cooperate and do something useful, these days you still have to carefully engineer your system, and that takes understanding. Moreover, if you ever succeed to have the schema[s] evolve without supervision, you'd still want to analyze the emergent behavior to gain understanding of what is going on – just the problem we have with explaining brain function." But we do not see the second sentence as a criticism. If we simply take a very large network and let it self-organize we gain no insight unless we can provide some high-level analysis of the outcome. Thus it may well be a fine strategy (one might even call it neo-Piagetian) to confront a functional decomposition of an adult behavior with developmental data, detailed neurophysiological recording, and modeling studies of self-organization (as indeed *Organization* itself does on several occasions). But, when the dust settles, we still need a language that bridges from an overall description of behavior or cognition to the differential contribution of a multitude of neural circuits. Schema theory (sense i) is our bridging language for this purpose, and we have offered specific models (schema theory, sense ii) to show how schema theory may maintain a dialog with a wealth of data as we try to build ever more comprehensive models of the roles of multiple brain regions in such overall behaviors as "saccade control" or "reaching and grasping."

Organization critically reviews and integrates the concepts of structural, functional, and (much more implicit) dynamic modules. We adopted a pluralist strategy combining a functionalist top-down approach and a structuralist bottom-up approach, and we offered many ideas for what might be called structure-based modeling of schemas. Morasso and Sanguineti (1997) entitle their book *Self-organization, cortical maps and motor control.* Unlike **Morasso,** we do not see these concepts as antithetical to schema theory; rather, we see them as full partners with schema theory in the development of Brain Theory, a partnership very much exemplified in *Organization.* As **Heinke** notes, our dynamical system theory (Ch. 4) plays an important role as a mediator between the functional-structural level of analysis and the neural level. Satisfying the constraints from these two levels is not a straightforward process and it requires numerous iterations.

Grunewald suggests that to understand neural processing it is useful to think of various areas providing different *competencies* (Grunewald 1999). The words *competence* and *schema* denote similar ideas; some specific functional-

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

ity that can be combined with other functions as the context requires. However, we do not find it useful to introduce new terminology (running a risk of confusion with Chomsky's "competence"), and urge **Grunewald** to use his insights to clarify the concept of schema. (It was in the latter spirit that, back in 1975 or so that, at the urging of the late Richard Reiss, Arbib adopted Piaget's word, rather than using a neologism of his own.) **Grunewald** (PCC) says that he would tend to shy away from a relationship to Piaget's schema, since that refers to the whole S-R process, but our use of "schema" in *Organization* is precisely aimed at providing a hierarchical refinement of Piaget and linking it to the brain. Our schemas are recursive. Chomsky's work seems inappropriate to us because (a) Chomsky's "competence" is innate, whereas the brain and mind develop, and (b) it is explicitly distinguished from performance – the rules, rather than the implementation of them. In any case, we agree with **Grunewald** that "there may in fact be many pathways [in the brain, and] each pathway is called upon depending on the contextual demands, for example due to task requirements," but we stress – as *Organization* shows – different coalitions forming, rather than a switching from one pathway to another, as when we recognize what an object is to determine how to grasp it.

R2.2. Anokhin

Borisyuk notes that in the 1970s the Russian physiologist Petr Anokhin developed the theory of "functional systems," and that the theory is still the basis for many neurophysiological studies (Sudakov 1997), but is not reflected in the book. However, this work has certainly influenced the authors of *Organization.* In the pages of BBS, Arbib and Caplan (1979) wrote:

> Luria transfers the emphasis from the brain-damaged patient to how brain regions interact in some normal performance – but using data on abnormal behavior to provide clues about neurological processes. Luria's approach is founded upon the idea of a *functional system* (Anokhin 1935) in which an invariant task can be performed by variable mechanisms to bring the process to an invariant result, with the set of mechanisms being complex and, to an important degree, interchangeable. From such a background, as well as from the developmental studies of Vygotsky (1934, translated 1962) and his own wide experience in neurology and psychology, Luria (1973, pp. 33–34) formulated the following program for neuropsychology: 'It is accordingly our fundamental task not to localize higher human psychological processes in limited areas of the cortex, but to ascertain by careful analysis which groups of concertedly working zones of the brain are responsible for the performance of complex mental activity; what contribution is made by each of the zones to the complex functional system and how the relationship between these concertedly working parts of the brain in the performance of complex mental activity changes in the various stages of its development.

This 1979 paper was part of an attempt to develop schema theory within the context of neurolinguistics, and has clear resonances (though not explicit acknowledgment) in Chapter 3 of *Organization.* Szentágothai (1979) wrote (in Hungarian) that

> P. K. Anokhin had already worked on the concept of the nervous system as a self-organizing system during the late forties. He assumed, based on results from repeating earlier experiment of Ten Kate under modified conditions, that at least some unspecific afferent input is necessary to generate movement patterns in the spinal cord. While Anokhin was hindered by the dogmatic atmosphere of his time and by the Pavlovian school being strongly reflex-oriented (he still got in trouble in 1950), I was less affected by these scruples. . . . Although those, aiming to explain neural function as a whole, and especially P. K. Anokhin, had been faced with this contradiction since the last decades of the 19th century, still no one has so far been able to either solve the problem or propose a clear experimental setup for solving it. However, the theory of dynamic patterns has substantially altered this situation, and can be easily approached by studying the highest center of the nervous system, the cerebral cortex.

This perspective is reflected in section 2.1.

R2.3. Schema theory: From action-oriented perception to cognition

Clark asks whether the strategies used to solve basic problems of perception and action also apply to more abstract and "cognitive" problems (problems of long range planning, deductive reason, thought about the distal, the absent, the abstract, the nonexistent, etc.). He labels as "cognitive incrementalism" the idea that you get human cognition by gradually adding bells and whistles to basic embodied strategies of relating to the present-at-hand. By contrast, he notes that the two visual systems hypothesis of Goodale and Milner (1992; Ch. 8 of *Organization*) holds that on-line visuomotor action is guided by neural resources that are quite fundamentally distinct from those used to support conscious visual experience, off-line visual reasoning, and visually based categorization and verbal report. Goodale and Milner argue for a radical dissociation of codings for on-line action and off-line reason and imagination. *Organization* tries to do justice to a kind of continuity while attempting to get to grips with what is special and distinctive about advanced human reason.

But how, **Clark** asks, can schema theory, which seems so well-suited to understanding basic, often pre-programmed behaviors (e.g., the work on frog visuomotor coordination, *Rana computatrix,* sampled in Ch. 3 of *Organization*), deal with the higher reaches of human thought and reason? We argue that instinctive (basic, special-purpose, perceptuomotor) schemas are joined by somewhat more abstract (learnt) schemas and that these two types of resources then interact in flexible ways to support intelligent human behavior. In section 8.6, we suggest that a variety of basic ("embodied, action-oriented") resources are combined with some more abstract learnt schemas with the whole assemblage orchestrated by cooperative computation into temporary ensembles according to the demands of a current task. But, **Clark** worries, how does the process of cooperative computation actually solve the recruitment problem? How are the right schemas assembled into the right temporary wholes at the right times? **Morasso** recalls (from *Organization,* p. 261) that Newell (1990) observed that as biological organisms became more and more complex as a result of natural selection, learning an exponentially increasing set of specialized mechanisms became inefficient and might have prompted the "Great Move" of evolution that established "a neutral, stable medium that is capable of registering variety and then composing whatever transformation are needed." **Morasso** suggests that neural self-organization should be able to build internal representations of the outside world in terms of maps, dynamic processes and sequences. We accept this, but not his caveat that this

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

would occur "independent of any functional schema," for in fact the cerebral cortex involves many specialized subregions linked in highly specific ways to specific thalamic nuclei and thence to other regions of the brain. The self-organization that proceeds within these regions is then constrained by the functional specificity implicit in the anatomy. We share with **Clark** a sense of how much remains to be done here, but feel that *Organization* has made a genuine contribution by providing in section 8.6 at least the framework for seeing how complex cognitive behavior builds upon the interactions of a set of complex, specialized neural structures, rather than being an unconstrained creation of "pure rationality" – or pure self-organization.

Our humanity both builds on and defies our animal past. Arbib (1985) quotes Shylock's impassioned defense of his humanity: "Hath not a Jew eyes? Hath not a Jew hands, organs, dimensions, senses, affections, passions? Fed with the same food, hurt with the same weapons, subject to the same diseases, healed by the same means, warmed and cooled by the same winter and summer, as a Christian is? If you prick us, do we not bleed? If you tickle us, do we not laugh? If you poison us, do we not die? And if you wrong us, shall we not revenge?" (Shakespeare, *The Merchant of Venice,* Act 3.) It is his "experience" that makes Shylock human. What is amazing about this quotation is how much of that experience Shylock shares with animals, but he is distinguished by his conscious knowledge of much of this experience. Theories of the mind that treat it as disembodied falsify the nature of our cognitive lives.

Indeed, **Grunewald** notes that much of the early work on artificial intelligence (AI) assumed that there is a representation of the sensory world that is multimodal, not tied to a specific sensory coordinate system, and from which information could be used to perform any movement, whereas neuroscience continues to amass compelling evidence against this notion of a single central representation. He not only notes the data on the "two visual pathways" (Goodale & Milner 1992; Ungerleider & Mishkin 1982) cited in *Organization,* but also notes evidence that there are also two auditory pathways (Romanski et al. 1999). Moreover (cf. *Organization,* sect. 8.4), several parietal areas are preferentially active before arm movements but not eye movements, and vice versa (Snyder et al. 1997). Schema theory explicitly rejects the notion of a central representation, instead stressing cooperative computation of multiple representations.

Bridgeman, too, discusses the two visual systems, but here emphasizing the need for psychophysical data. Bridgeman et al. (1997; Bridgeman 1999) show how psychophysical studies demonstrate independent representations of visual space in the two systems. In particular, they show that a symbolic message about which of two targets to jab can be communicated from the cognitive to the sensorimotor system without communicating the biases that they have demonstrated in the cognitive system. **Bridgeman** notes that *Organization* asserts that the two systems engage in direct cross talk (p. 244), and appears to see this as a criticism. However, our models (whether expressed as schema assemblages, neural networks, or dynamic systems) are both predictive and open to refinement in the light of new data. We thus see **Bridgeman's** data as a welcome input for our further development of models of visuomotor integration, rather than as in any way a reflection on our methodology.

R3. Neural modeling and dynamics

R3.1. How detailed a model of the neuron is relevant to neural organization?

For **Gouras,** the consideration of oscillations or the pharmacological ingredients that are liberated at synapses seem irrelevant. He sees the Hodgkin-Huxley equations for propagated axonal conduction as a highpoint in the mathematical description of neural function, but of little relevance to understanding how neural connectivity defines brain function. For him, McCulloch-Pitts models are just detailed enough for understanding the brain, and there is no need to use biophysically detailed models. We cannot fully accept this argument – though we would be ready to defend the merit of the McCulloch-Pitts model if somebody attacked it as completely irrelevant neurally! In sympathy with **Gouras,** we note that in modern biological research we often meet with the "tyranny of microscopic details," whereas the search for organizational principles is neglected. However, the Hodgkin-Huxley equations certainly are important in understanding the neural building blocks of the brain. We would argue that different problems call for different levels of detail in their analysis, and that models with intermediate complexity will have an increased role in modeling interacting brain regions. Population models, simplified network models (sect. 8.4.1. in *Organization*) and models with spiking neurons are specific examples of this approach. In Érdi's group (Barna et al. 1998; Bazsó et al. 1999; Gröbler et al. 1998), a scale-invariant theory (and software tool) was developed (see Précis), which gives the possibility of simulating the statistical behavior of large neural populations, and monitoring the behavior of an "average" single cell synchronously. There is a hope that activity propagation among neural centers may be realistically simulated, and schemas can be built from structure-based modeling. In other studies, detailed multicompartmental modeling led to the understanding of the dynamic emergence of schemas. These include the analysis of the central pattern generators (CPGs) for the lamprey spinal cord (Wallen et al. 1992) and leech heart beat (Nadim et al. 1995).

Networks of neurons organized by excitatory and inhibitory synapses are structural units subject to time-dependent inputs and they also emit output signals. From a functional point of view, a single neural network may be considered sometimes as a pattern generating and/or a pattern recognizing device. More often it is not a single network but a set of cooperating neural networks that forms the structural basis of pattern generation and recognition. (Pattern generating devices are functional elements of schemas.) Consequently, whereas the structural organization of many neural centers can be understood based on the concept of modular architectonics, that is, a network made up of repetitive modular elements, there is no one-to-one correspondence between structural and functional modules.

Depending on its structure, an autonomous neural network may or may not exhibit different qualitative dynamic behavior (convergence to equilibrium, oscillation, chaos). Some architectures show unconditional behavior, which means that the qualitative dynamics does not depend on the numerical values of synaptic strengths. The behavior of other networks can be switched from one dynamic regime to another by tuning the parameters of the network. One mechanism of tuning is synaptic plasticity, which may help

to switch the dynamics between the regimes (e.g., between different oscillatory modes, or oscillation and chaos, etc. . . .).

Turning to finer levels of detail, **Székely** notes the importance of studying synaptic triads. Attributing chaotic behavior to the synapsing elements, the synaptic triad acts as an input dependent ON gate in the visual system. He thinks the successful treatment of triads in model studies has a special significance because serial section electron microscopic studies reveal that the "at a distance" type synaptic triads are in great abundance in the spinal cord where the conventional "closely packed" triads are relatively rare (Székely 1989). He notes (PCC) the challenge of analyzing impulse transmission in the more complex synaptic islands.

R3.2. Clarifying the multiple functions of neurons and brain regions

Grunewald asserts that *Organization* suggests that the role of the posterior parietal cortex (PPC) is to provide a remapping mechanism, and then he stresses that many neurons in PPC code a variety of signals at different times including memory signals during visual memory saccades (Gnadt & Andersen 1988), or auditory signals in fixation or saccade tasks (Grunewald et al. 1999; Linden et al. 1999). However, this is consistent with our approach because, in section 8.4, we too stress the specialization of different subregions of PPC. The emphasis on remapping was for portions of PPC involved in saccadic eye movements. In the same spirit, **Grunewald** notes that the role of the frontal eye fields (FEF) cannot be restricted to saccade or memory functions alone, as it was in section 8.4.2 of *Organization,* because recent physiological experiments suggest that FEF is also involved in other functions (Bichot & Schall 1999; Thompson & Schall 1999). But this is fine with us. Each model is limited in scope. Our continuing challenge is to show that our models do explain some new data, whereas other data are accommodated by minor revisions of the model that preserve its explanations of datasets considered earlier. What is to be avoided at all costs is the ad hoc development of a new model for each and every dataset. As such modeling progresses, we will come to better understand the interplay between specificity and polyfunctionality exhibited in so much of the nervous system.

R3.3. In defense of dynamics

Edelman may be right when he says that "'dynamics' does not play too prominent a role in the book." However, he qualifies this fact as a fortunate one. We believe that it reflects the present status of research rather than the tendency of the future. We think that both structuralist "bottom up" and functionalist "top down" approaches should lead to models formulated in the language of dynamical system theory. Of course, we do not expect to find (or even to search for) the differential equations to be called "the brain equations." In accordance with what we wanted to express in *Organization,* Érdi (2000) argues that dynamical system theory offers conceptual and mathematical tools for describing the performance of neural systems at very different levels of neural organization. Neural rhythms, neural and mental development, and macroscopic brain phenomena monitored by brain mapping devices are different aspects of neural dynamics.

R3.4. Functional neuro-imaging

We certainly accept the critique (**Sporns** and **Gouras**) that neuroimaging is under-represented in the book. Indeed, the whole area developed very much during the writing of the book and thereafter. Here, we briefly summarize our views on bridging the gap between conventional neural modeling and data evaluation for the new brain imaging methods. PET and fMRI studies observe average activity data calculated for volume units (voxels). A voxel of the present techniques is large compared with even a cortical module. (A cortical module contains several thousand neurons.) Voxels showing correlations in blood flow during a specific cognitive task are assumed to be functionally connected. Interacting brain regions form a functional network. Structural modeling provides the numerical values of correlation coefficients between brain regions by assuming a linear relationship between these active areas. More precisely, although functional connectivity reflects temporal correlations between remote neurophysiological events, effective connectivity characterizes the influence one neural system exerts over another (Friston 1994). The present evaluation techniques are based on statistical analysis (such as correlation, regression, principal component, and independent component analysis). A more sophisticated, still static, technique, called "covariance structural equation modeling" was also introduced and applied for a number of specific cognitive tasks. This method combines anatomical and hemodynamic-metabolic data and was qualified (Horwitz et al. 1999) as system-level neural modeling. Arbib et al. (1995) introduced a large-scale neural modeling technique that connects neural activity to PET data. The key hypothesis to establishing a synthetic PET result (cited in the book) is that counts acquired in PET scans are correlated with the synaptic activity within a region. Simulated PET activity was computed by integrating the absolute value of the total synaptic activity of a neural subnetwork over the time course of the study within the different areas. To study saccade generation, for example, several regions, such as basal ganglia, frontal eye fields, posterior parietal cortex, visual cortex, and superior colliculus were modeled.

Sporns notes that *Organization* does not concern itself much with statistical approaches to neural activity based on information theory (Rieke et al. 1997), or with approaches based on functional connectivity, correlation or coherency analysis. By and large, there are two classes of problems related to modeling studies. First, the direct problem starts with the setting up of a model, and then studies the properties of the model. Second, the inverse problem starts from experimental data, and tries to find out the proper mechanisms and algorithms, which produce results best fitted to the data. Techniques of data evaluation, even in a broad sense, were outside the scope of the book. Nonetheless, we await with interest the outcome of Sporns's claim that analytical techniques based on information theory may prove useful in the analysis of multidimensional data sets from neurophysiology or neuroimaging.

Recently (see, e.g., the Fifth International Conference on Functional Mapping of the Human Brain, *http://www.apnet.com/hbm99/methphyscog.html*) there has been an increasing tendency both to improve the temporal resolution of the experimental devices, and to develop new techniques of evaluation. Among others the use of nonlinear dynamics was demonstrated in the human brain using PET

and transcranial magnetic stimulation (Meyer-Lindenberg et al. 1999). Unstable cortical states are accompanied by increased rCBF in a well-defined, discrete set of areas. This may reflect increased synaptic activity due to increased network firing rate near the transition point. There seem to be methods to detect cortical instabilities, which may lead to neuropsychiatric disorders, and analyze them by the methods of dynamic system theory. To connect neural mechanisms to dynamic cognitive networks is still a long way off. There is a need for mesoscopic level neural simulations to bridge the gap between microscopic single-cell activities and macroscopic brain states. A step in this direction is statistical population dynamics (Barna et al. 1998; Gröbler et al. 1998).

R3.5. Attention

The metaphor of a window of attention describes a space-based selection process (Crick 1984). **Heinke** argues that nowadays selective attention is not considered as a merely spatial selection. **Heinke** (PCC) offers the following example: If two identical items move in the same direction and the two items are separated by a static different item, the two moving items still can be selected together. This contradicts the view of selective attention operating as a focus of attention (FOA). Interesting to note, a similar issue was considered long ago in the modeling of stereovision. If grass is viewed through a picket fence, there is no reason the disparities of the fence should be related in any way to the disparities of the grass surface. Similarly, one may observe a pane of glass and the scene behind it. Prazdny (1985) offered an algorithm that successfully detects disparities generated by opaque as well as transparent surfaces. The principal disambiguation mechanism is facilitation due to disparity similarity; Prazdny argues that dissimilar disparities should not inhibit each other because, when there are transparent surfaces, features at a given disparity may be surrounded by a set of features corresponding to other surfaces, and thus at different disparities.

R3.6. Synthetic brain theory versus attractors

Sporns notes that models that incorporate diverse neurobiological knowledge, from realistic synaptic rules to patterns of interconnectivity that mimic those of real brain areas, are *synthetic,* in that they "put together" individual components across levels. We agree with Sporns about the necessity and possibility of constructing a synthetic brain theory. Typically, one approach (the bottom up) we adopted in the book is in accordance with its spirit. First, we defined structure, including network architecture, then activity patterns are determined, followed by the formation and modification of synaptic connectivity patterns. If everything goes well, overall behavioral patterns are constructed. A very clear illustration of this strategy in *Organization* is Chapter 5 on the olfactory system. However, the neural coding problem is under-represented in the book. At the time of writing of *Organization,* an extensive discussion started about the nature of the neural code. The specific question is how strong are the arguments for supporting the traditional (but not clearly defined) coding scheme based on the "mean firing rate," and what evidence supports the view that the neural code is buried in the temporal pattern of the spikes? To give some pointer to the broad literature we mention the two recent books by Rieke et al. (1997) and Maass and Bishop (1999).

As **Sporns** notes, such biologically constrained models are to be contrasted with "classical" connectionist networks with their highly stereotypic architectures and algorithmic learning procedures. He sees brain dynamics as characterized by dynamic transients, spontaneous activity and time-varying inputs, rather than by attractors selected by constant input patterns. Indeed, *Organization* offers a critique of the role of attractors in neural modeling in section 4.5.4. **Sporns** (PCC) adds that cortical dynamics is highly stochastic, perhaps chaotic, with seemingly highly irregular spike trains (Mainen & Sejnowski 1995; van Vreeswijk & Sompolinsky 1996) with on-going activity characterized by large fluctuations and persistent variability (Arieli et al. 1996). Another complication is that nervous systems often arrive at "decisions" very quickly, with some people arguing (e.g., Rieke et al. 1997) that very few spikes per neuron are needed to produce a result, perhaps only a single one. Perhaps transients are all there are. Indeed, models that take explicit account of continuous interaction with the environment are nonautonomous in the mathematical sense (Aradi et al. 1995; Érdi et al. 1992b). Such systems do not have attractors in the general case. Consequently, attractor neural network models cannot be considered as general frameworks of cortical models.

However, the role of attractors in the dynamic system theory of neural networks is not as limited as an emphasis on fixed points might suggest. The notion of chaotic itinerancy, as a universal dynamical concept in high-dimensional dynamical systems, has been suggested and applied to neural information processing by Tsuda (1991; 1992; 1996):

> The starting point is a dynamical system which can be characterized by the coexistence of attractors. A state can be a particular attractor. This attractor, however, may become unstable, while other attractors could preserve their stability, i.e., the instability can remain localized. The trajectory of the dynamical system moves along the unstable manifold, and approaches another attractor. If instability is strong enough, many chaotic modes appear and consequently the system goes toward a turbulent state, that is, a quite noisy macroscopic state. Then, even a "trace" of the original attractors disappears. If instability is, however, not so strong, an intermediate state between order and disorder can appear. The dynamics may be regarded as an itinerant process which ensures a transition among states which were at the beginning described as attractors but now are no longer attractors. In this case, a crucial characteristic is that a "trace" of attractors remains in spite of the generation of unstable directions in the neighborhood of attractors. (Courtesy of Tsuda; Érdi & Tsuda, in press.)

R3.7. Multi-rhythmicity

Borisyuk notes the occurrence of multi-frequency oscillation as a promising tool for neural information processing. Although we have not analyzed it in a very detailed way, at least in two cases we referred to this interesting concept. First, referring to the work of Jahnsen and Llinás (1984) we mentioned (p. 82) that multi-rhythmicity may occur even at the single-cell level. Second, in the network model of the olfactory bulb the numerical bifurcation analysis (Fig. 5.6) showed a periodic-doubling bifurcation structure from periodicity to chaos. This transition goes trivially through multifrequency oscillations. However, its eventual role in the

information processing was not mentioned, because it has not been clarified. Multi-frequency oscillation may have a very important role in hippocampal information processing. Two main, normally occurring, global hippocampal states are known: the rhythmic slow activity, called the theta rhythm with the associated gamma oscillation, and the irregular sharp waves (SPW) with the associated high frequency (ripple) oscillation (Buzsáki 1996). In a recent model of a network of interneurons we showed that synchronized gamma oscillation and theta rhythm can be generated together (Kiss et al., in press).

R3.8. The binding problem

Edelman comments that "invariant" pattern recognition is finessed into the "binding problem" and is offered a solution in the form of von der Malsburg's Dynamic Link Architecture (p. 102). He asserts that the model only becomes relevant if one accepts the prior assumption that binding is indeed a problem. **Edelman** (PCC) is now working on a model of structure representation that does not require traditional symbolic binding.

Borisyuk states that the feature binding problem can be described using multifrequency oscillations (Borisyuk et al. 1999) when feature binding is realized owing to synchronization of oscillators. But if we see a bookshelf with hundreds of books, we have the impression that they are clearly segmented from one another – certainly we can see 50 or 60 books as distinct objects in a single foveation. How can neural oscillators support this many different frequencies? Something else must be at work. **Borisyuk** (PCC) agrees that the range of possible frequencies is limited. He suggests that the same frequency might be used in several different subregions of the visual cortex but this seems to raise more questions than it answers.

R3.9. Cortical maps and architectonics

Edelman notes that when we discuss dynamical models of ocular dominance formation, nothing is said about why there are ocular dominance columns in the first place. In fact, current evidence suggests that they are not functional, but a product of early self-organization, after which functional effects (e.g., mediating stereo) become more important. Viewed this way, ocular dominance columns are a by-product of squeezing several dimensions into two. Frogs have complete crossing of optic fibers at the chiasm and do not exhibit ocular dominance columns in the tectum. Yet when a third eye is grafted into the tadpole, and the tectum of the frog shows retinal innervation by two eyes as a result, then eye-specific termination bands do indeed appear (Constantine-Paton & Law 1978). Indeed, **Morasso** notes that if cortical maps are trained to represent higher-dimensional manifolds, apparent fragmentation will be a side-effect of mapping an N-dimensional manifold into a two-dimensional substrate. Thus absence of somatotopy may be more apparent than real. However, this does not, as **Morasso** argues, contradict our opinion that the computational functions of associative cortical areas are not unitary (sect. 8.4). For example, Krubitzer (1998) stresses the strong role of activity-based reorganization of cortical maps, but also notes that in a monkey with an eye removed before the visual pathway was well established, the primary visual area and its inputs still formed but were smaller than normal (Rakic 1988). In other words, evolution lays down specialized areas in the brain that are malleable during the self-organization of development, but are not the mere expression of "mapping an N-dimensional manifold" into a uniform cerebral cortex.

Kötter notes that our discussion (sect. 8.2) of models generating modularity in visual cortex omits such other aspects of visual cortical organization as topological analyses of long-range connectivity (Young 1992; 1993). **Kötter** (PCC) suggests that new data to challenge Szentágothai's concepts came from studies of autoradiography of receptor distributions (cf. Geyer et al. 1998; Zilles & Clarke 1997). Such data could resolve the debate whether every layer V pyramidal neuron in the cerebral cortex can be regarded as the center of a virtual column or whether a far smaller number of columns exists. In the rodent barrel cortex, one can ask whether barrels are superordinate to pyramidal cell-centered columns or replace them. They certainly are subject to reorganization if vibrissae are removed. As a contribution to updating the debate, we went beyond *Organization* in the Précis by reviewing Somogyi et al.'s (1998) discussion of "Salient features of synaptic organization in the cerebral cortex" as well as recent developments related to modular architectonics.

R3.10. Chaotic dynamics

Borisyuk suggests that chaotic activity in the brain might be synchronized or partially synchronized and this is a possible basis for an information processing. **Borisyuk** (PCC) offers Makarenko and Llinás (1998) and Zaks et al. (1999) as relevant papers.

Gregson wants to get away from "single idealized neurons and [use] the properties of recursive pathways as the elements of modelling. Such pathways need augmentation by gating functions, whose dynamics have been studied from both physiological and mathematical perspectives." We are sympathetic with this approach, but are also concerned that our dynamic analyses do indeed make contact with the data of neurophysiology and neuroanatomy.

Gregson suggests that "in psychological dynamics, discrete time processes may be more plausible and tractable." First, we admittedly reviewed neurodynamic models, though fully aware of important problems and models at the psychological level (e.g., multistable perception and cognition; Stadler & Kruse 1994). There are many arguments for using continuous or discrete-time models. What is true is that the notion of "immediate next time" can be interpreted easily in the discrete case only. Second, as concerns the history of neurodynamics and neuropsychology, even the pioneers made different choices: "Where Rashevsky had assumed that the relevant mathematics was differential equations and the relevant conceptual tools of physics, McCulloch and Pitts assumed that the relevant mathematics was the Boolean and first-order logic found in Rudolph Carnap's system logic, inspired by Bertrand Russell and A. N. Whitehead" (Aizawa 1996). (The McCulloch-Pitts model is formulated in discrete-time.) Third, the tractability of an equation does not depend on its semantic connotation. When you solve an equation, it is equally difficult or easy if it is considered as a model of physical, chemical, biological, economical, or whatever phenomenon.

R4. Learning, development, and self-organization

R4.1. Monoaminergic systems, reinforcement, and varieties of learning

Zhadin argues that positive reinforcement (PR) and negative reinforcement (NR) must be obligatory elements in any model of learning, and that unconditioned inputs originating from unified centers of PR and NR must terminate on the overwhelming bulk of neurons to provide a wide variety of conditioning forms. He then observes that many Hebbian synapses are detected in the hippocampus, but they are seldom revealed in the mature neocortex, which is responsible for the highest mental activity, recognition, and complicated forms of memory and learning. He even asserts that monosynaptic LTP is most likely to have nothing to do with learning at all, having no association with the reinforcement process. He argues that the only possible source of the multiple synaptic input of reinforcement could be the monaminergic systems (MS): serotoninergic (SS) from the raphé nuclei, norepinephrinergic (NS) from the locus coeruleus, and dopaminergic (DS) from the substantia nigra. Gromova's (1980) lesion data suggest that that SS mediates PR, and NS mediates NR. (We are unable to adjudicate the apparent contradiction with the view of NS mediating PR, Olds & Olds 1963.) **Zhadin** modeled electrical activity of cortical neurons with prolonged action of PR or NR. With PR, the neuronal excitation tends to its highest level upon initial relatively high excitation or to complete inhibition upon initial low excitation with possible accidental transitions from one extreme level to another. With NR, it tends to some intermediate excitation level, regardless of its initial excitation. In our own group, we have recently analyzed related phenomena in striatum. To model the transient influences of dopamine on membrane properties of medium spiny neurons in striatal matrisomes, Suri et al. (2000) simulated the in vitro finding that activation of D1 dopamine receptors decreases firing evoked from the hyperpolarized resting potentials but increases firing evoked from elevated holding potentials (Hernandez-Lopez et al. 1997). Likewise, we model dopamine effects on corticostriatal transmission depend on the postsynaptic membrane potential (Cepeda & Levine 1998). The membrane potential is influenced by synaptic inputs, by dopamine levels, and by rhythmic fluctuations of about 1 Hz between a depolarized up-state and a hyperpolarized down-state (Stern et al. 1997).

Zhadin (PCC) stresses that learning must serve the goal of helping a living being survive, and asserts that in the light of all the foregoing, some modification of Hebbian learning rules seems to be necessary. However, without in any way diminishing the importance of **Zhadin's** study of the role of the monoaminergic systems in cerebral cortical plasticity, we are still left with interesting problems: When is a generic Olds-type signal an adequate basis for learning, and when are more specific signals needed (as seems to be the case for error-based learning in the cerebellum)? Going further, when are both relevant? For example, when do we learn something despite the pain, motivated by some distant goal? Indeed, the whole point of reinforcement learning (Sutton & Barto 1998; with roots going back to Samuel 1959) is that (as in playing a game of checkers) the primary reinforcement (e.g., whether we win or lose) is seldom available. We must continually build expectations of reward and build our decisions on this. In fact, evolutionary "learning" does this by building in "pain" and "pleasure" systems that give a "hard-wired" estimate of whether the current state of activity is or is not conducive to survival. These become the built-in "primary reinforcers," but further learning is required to determine whether a particular activity, neutral in itself, is more likely to lead to pleasure or pain. Midbrain dopamine neurons do not signal "primary reward," but are themselves shaped by learning. They are phasically activated by unpredicted rewards or by the first sensory event that allows the animal to predict the reward. They do not respond to predicted rewards (they have responded already to the prediction) and their activity is depressed when a predicted reward fails to occur (Schultz 1998). These features of dopamine neuron activity can be reproduced with the reward prediction error of temporal difference models (TD models) (Montague et al. 1996; Schultz et al. 1997; Sutton & Barto 1990).

Simulation studies with TD models demonstrate that a dopamine-like reward prediction error can serve as a powerful effective reinforcement signal for sensorimotor learning (Houk et al. 1995; Suri & Schultz 1998; 1999). In such models, the TD model was the "critic" and the model component that learns sensorimotor associations was the "actor." Suri et al. (2000) simulate dopamine neuron activity with an Extended TD model and examine the influence of this signal on medium spiny neurons in striatal matrisomes. This model includes transient membrane effects of dopamine, dopamine-dependent long-term adaptations of corticostriatal transmission, and rhythmic fluctuations of the membrane potential between an elevated "up-state" and a hyperpolarized "down-state." The most dominant activity in the striatal matrisomes elicits behaviors via projections from the basal ganglia to thalamus and cortex. To investigate possible functions of the simulated biological mechanisms, we tested the performance of several model variants that lack one of these mechanisms. These simulations show that adaptation of the dopamine-like signal is necessary for planning and for sensorimotor learning. Lack of dopamine-like novelty responses decreases the number of exploratory acts, which deteriorates planning capabilities. Sensorimotor learning requires dopamine-dependent, long-term adaptation of corticostriatal transmission. The model loses its planning capabilities if the dopamine-like signal is simulated with the original TD model. These simulation results suggest that striatal dopamine is important for sensorimotor learning, exploration, and planning.

But this is still very formal. In the end we must learn from ethology and seek to relate our studies to observation of the natural behaviors of animals (and ourselves). We must understand how animals that may have problems learning what seem to us very simple tasks in the laboratory, may nonetheless solve what seem to be far more complex problems when they are "natural." **Zhadin** (PCC) observes that his cat learns things very quickly when the cat considers them useful. For example, when looking for the way to the window through new positions of the blinds, he tries once, then sits, thinking for a time, and then goes through a new way at once without any hesitation. (Note the use of the words "considers" and "thinking.") Jean-Paul Joseph (personal communication) notes that in a delayed response task (spatial) with distracters, young monkeys can perfectly perform the task if the wells are baited with the reward while

they watch. Their working memory is good. By contrast, if the same monkeys have to perform the same task with buttons in a match-to-sample task (Barone & Joseph 1989), their performance is very poor. They do not understand the task when it becomes too abstract. A monkey can learn unnatural "spatial" or nonspatial (Push-Pull-Turn) sequences. When the monkey becomes familiar with the task, there is no doubt that it has an idea of the sequence. The problem is that the animal does not transfer (or at least not easily) this knowledge to a new environment or to a new setting. The "idea" of sequence remains attached to the display where the sequencing task has been learnt and performed.

R4.2. Self-organization

With this background, we may continue our response to **Morasso's** discussion of Newell's "Great Move" in evolution: there is an interplay between a complex array of "natural" schemas and general learning mechanisms underlying a vast range of behavior. It is still a great mystery why the human brain can support sequences and symbols at an abstract level denied to other species. Building on the work of many others, this is a matter of current concern to us (Arbib 2000; Rizzolatti & Arbib 1998).

Morasso notes that Doya (1999) has pointed out that the learning paradigm, rather than specific perceptual or motor functions, characterizes in a unique way the computational nature of cerebral cortex, cerebellum, and the basal ganglia: (1) an unsupervised self-organizing paradigm in the cerebral cortex, for building cortical maps; (2) a (self)-supervised paradigm in the cerebellar circuitry, for learning internal dynamical models; and (3) a reinforcement learning paradigm, in the basal ganglia, for learning sequential aspects of actions in complex tasks. We find it hard to see what Doya adds thereby to *Organization.* We present models for each of the three classes above, but also present models exploiting reinforcement learning in cerebral cortex. Our model of the role of basal ganglia in sequence learning (*Organization,* sect. 10.5; Dominey et al. 1995) has recently been extended to a "dual process model" (Dominey et al. 1998) and applied to infant linguistic performance (Dominey & Ramus 2000). Note, however, that it is a mistake to think that learning principles alone exhaust the significant differences in the cerebral cortex, cerebellum, and the basal ganglia since the regions differ in their cellular structure and connectivity, and have quite different relations with the sensory and motor peripheries, which differ dramatically from subregion to subregion.

Morasso's problem seems to be that he sees a sharp dichotomy where none exists. His approach to brain theory emphasizes self-organization, cortical maps, and motor control (Morasso & Sanguineti 1997), and we agree with the importance of all of these. But he then insists that the key lies in emphasizing the constraining role of the nonlinear dynamics of organism-environment interaction and the interactions among neural assemblies in the formation of cortical maps, to the exclusion of the study of schemas. Yes, much comes about by self-organization. Nonetheless, we may properly ask how, for example, our navigating integrates a cognitive map with visual and auditory cues, and how the specialization of the architecture and chemistry of different brain regions evolved to serve these subschemas of the overall behavior and their integration. We fail to see how progress can be made without a careful top-down analysis of the animal's range of behaviors to balance bottom-up studies of the dynamics of self-organization. It would be like a physicist who claims that it is a mistake to study thermodynamics, for surely all will emerge from a statistical analysis of a myriad of molecules. No. Just as statistical mechanics needs thermodynamics to frame its questions (and can then yield new insights), so does neurophysiology need schema theory as a language for behavioral analysis adapted to the distributed computational style of the brain to frame studies of how neural self-organization can yield the types of natural behavior discussed above.

We certainly agree with **Székely** that questions of how nervous structures evolve in phylogenesis and develop in ontogenesis are hard to approach in a framework based on precise specification of cell-to-cell interconnections among neurons forming reflex arcs. **Székely** observes that if the notion of genetically controlled cell-to-cell neuronal interconnections is abandoned, then self-organization may be the mechanism of structure formation in the brain. We almost agree. Although we happily abandon "genetically controlled cell-to-cell neuronal interconnections," we would still want to invoke species-specific genetical control of interconnections between specific nuclei, with the gross positioning of nuclei and the morphology and position of the cell types within them being also genetically controlled to a first approximation. Indeed, **Székely** himself asserts that in "the formation of modules in the basal ganglia . . . the peculiar form of the component neurons almost present themselves to get organised 'spontaneously' into a structure. . . . It seems that neuronal morphology plays a significant role in the organisation of neural structures." However, it is unclear that the presence of self-organization is relevant to **Székely's** desire to displace "the position of the reflex arc as the basic functional unit of nervous activities." The gross connectivity would favor or disfavor reflexes, whatever the self-organization of the "second order connectivity" achieved. However, we agree with **Székely** that the reflex arc is not the basic functional unit of nervous activities; our only disagreement is that we see the battle as already won! We cited Székely (1989) as crucial evidence for this in section 2.1.3. Other relevant data in *Organization* concerns the study of central pattern generators (CPGs), which displace reflexes even at the level of the spinal cord, as well as "higher-level" considerations of action-oriented perception, and so on.

R4.3. Synapse pruning during postnatal development

Hoffman's simulations demonstrate that selective elimination of connections enhances the computational capacity of networks capable of temporal processing. Early in development it is likely that synaptic connections are created more or less randomly, with subsequent selective elimination based on environmental experience as well as endogenous factors. Earlier computational studies on the Boltzmann machine (Barna & Érdi 1990) showed that "contrary to the naive expectation it turned out that sometimes performance might be improved by reducing the degree of connections."

It is more or less accepted that very different variables characterizing neural development (such as number and density of synapses, dendritic spines, axonal and dendritic arborization patterns, cortical volumes, etc.) exhibit a well-defined maximum in their time course. Specifically, Quartz

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

and Sejnowski (1987) suggested that directed dendritic development seems to be an important component in brain development and also establishing the representational properties of the cortex. (The development is called "directed" rather than "selected" because the emerging shape is strongly dependent on its input activity pattern.) It is a challenge to give a satisfactory explanation for the general occurrence of the overshooting phenomena during normal neural development. Based on kinetic analysis, Érdi (1984) suggested as a phenomenological explanation that the "genetic component" of the development shows a sharp maximum, and it decays without reinforcement; the environmental factors may support the survival of the intrinsically formed structure by some feed-forward mechanism.

R4.4. Disease

Where **Gouras** views the consideration of epilepsy in *Organization* as irrelevant – "That an epileptic focus can develop in certain areas, presumably due to some defect in negative feedback, also seems irrelevant. Modeling epilepsy or oscillations seems of secondary importance." – **Bookstein** observes that the neurosciences arose in neurology, and takes us to account for providing (almost) no explanations of disease states that follow from predictable behaviors of the models developed in *Organization:*

> [The book restricts its] subject-matter to normative (healthy) behaviors, and the more admirable of those behaviors at that – nothing here about rage, lust, or the hallucinations of schizophrenia – [and ignores] differences among the different nervous systems that are all evidently equally possible: the nervous systems that fill our psychiatric wards right alongside our neuroscience laboratory chairs.

Dominey suggests the importance of seeing how data from the treatment of Parkinson's disease by selective lesions and stimulation in the basal ganglia nuclei could provide important constraints for improved models of basal ganglia function. "What may now be required is an effort to begin to integrate the pieces (i.e., the models of cerebellum, basal ganglia and cortex, hippocampus, olfactory system, etc.) into a system model, with 'macaca computatrix' as the descendent of 'rana computatrix'." We believe that *Organization* provides a proper conceptual and mathematical framework for a "computational neurology," but side with **Bookstein** and **Dominey** in agreeing that further comparison of brains in sickness and in health will greatly strengthen our understanding of neural organization. We offer some observations for work in this direction.

R4.4.1. Epilepsy: Dynamic models of generation and control. Epilepsy is a typical example of a dynamical disease, that is, one that occurs in an intact physiological system yet leads to abnormal dynamics. Epilepsy itself is characterized by the occurrence of seizures (i.e., ictal activities). During epileptic seizures oscillatory activities emerge, which usually propagate through several distinct brain regions. The epileptic neural activities are generally displayed in the local field potentials measured by local electroencephalogram (EEG). Epileptic activity occurs in a population of neurons when the membrane potentials of the neurons are "abnormally" synchronized. Both experiments and theoretical studies suggest the existence of a general synchronization mechanism in the hippocampal CA3 region. Synaptic inhibition regulates the spread of firing of pyramidal neurons. Inhibition may be reduced by applying drugs to block (mostly) $GABA_A$ receptors. If inhibition falls below a critical level, complete synchrony occurs. Collective properties of networks of pyramidal cells modulated by inhibition have been studied successfully by Traub and Miles (1991). As we already know, a certain degree of synchrony is necessary for normal theta and SPW (sharp wave) behavior, and the transition between normal and abnormal degrees of synchrony is not clear. Rather arbitrarily, activity has been considered epileptic if more than 25% of the cells fire during 100 msec. In vitro models of epilepsy (Traub & Miles 1991) offer a means to study the cellular mechanisms of the different types of epileptic phenomena by combined physiological and simulation methods. Several in vitro models of seizures have been developed, including electrical stimulation, low calcium, low magnesium, and elevated potassium levels.

Dynamic system theory offers a conceptual and mathematical framework to study epileptogenesis (Barna et al. 1998; Lopes da Silva et al. 1994; Lytton et al. 1998; Traub & Miles 1991). Analytical studies based on bifurcation theory should clarify the possible operating modes of a given neural network. The balance between excitation and inhibition is certainly one important control parameter, and its change may imply transition between the regimes. Epileptic activities may be considered as chaotic processes (Babloyantz & Destexhe 1986). There has been some hope that techniques of controlling chaos may offer new therapeutic and diagnostic tools for controlling epileptic activities (Schiff et al. 1994).

R4.4.2. Alzheimer's disease. Two network models of pathogenesis of Alzheimer's disease have been given. A compensatory mechanism was suggested by Horn et al. (1993) who showed that deterioration of memory retrieval due to synaptic deletion can be much delayed by strengthening the remaining synaptic weights by a uniform compensatory factor. A different approach was given by Hasselmo (1994) by using the concept of "runaway synaptic modification." Runaway synaptic modification denotes a pathological exponential growth of synaptic connections that may occur due to interference by previously stored patterns during the storage of new patterns. Menschik and Finkel (1998) gave a model-framework for Alzheimer's disease in terms of neuromodulatory control of hippocampal function. They used a multicompartmental technique but their network architecture mimicked the hippocampus rather roughly, since the convergence and divergence numbers were not take into account.

R4.4.3. Schizophrenia. Schizophrenia has both positive and negative symptoms. One example of the former is hallucinations, and an example of the latter is poor speech content. Although the general paradigm for explaining neural diseases is correlated to diminished functions caused by structural lesions, Han et al. (1998) suggested that hippocampal lesions occasionally may enhance learning performance. In some sense the lack of competitive cue interaction can be explained by certain types of enhanced learning. The pathogenesis of schizophrenia is basically unknown. Several families of models toward the explanation of schizophrenia exist. **Hoffman** reports on simulations that may provide insights into the pathophysiology of schizophrenia. He argues that symptoms characteristic of schiz-

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

ophrenia, such as hallucinated and disorganized speech, can arise from overzealous pruning of cortical networks during adolescence. Hoffmann (1987; 1996) used the physicists' well-known Hopfield model, and showed how pathological alterations in the attractor neural network can lead to the formation of so-called parasitic attractors. Such attractors may be responsible for the emergence of schizophrenic delusions and hallucinations. These phenomena may correlate with overloading of memory capacity due to neurodegenerative changes. **Hoffman** (PCC) offers (Honer et al. 1999; Selemon & Goldman-Rakic 1999) as pertinent to the notion that if the pruning in association cortex that ordinarily occurs during adolescence does not get shut off or if connectivity at baseline in childhood is excessively sparse, then pathology should emerge which we see as schizophrenia.

The hippocampus, the amygdala, and the "limbic region" of the basal ganglia form a system which seems to be the structural basis of anxiety and schizophrenia. The hypothesis of Gray et al. (1991) is based on the neuropsychological assumption that the septo-hippocampal system has predictive and comparator functions. Friston (1998) reviewed the disconnection hypothesis of schizophrenia and has presented a mechanistic account of how dysfunctional integration among neuronal systems might arise. It was hypothesized that modulation of the asociative changes in synaptic weights implies abnormal interactions between certain areas, such as prefronto-temporal interactions. Clearly, much remains to be defined to extract a comprehensive understanding of schizophrenia from these diverse approaches.

R5. Postmodernism and reductionism

R5.1. Post-modernism and the brain

Keil & Davids brand our "reliance on terms like coding, transformation, and representation" as reprehensibly modernist but applaud our efforts "to introduce postmodern concepts such as chaos and self-organization to the study of neural organization." Moreover, **Keil & Davids** try to reduce schema theory to a mere "computer metaphor" by ignoring the whole discussion of *cooperative computation* whose job it is to move us beyond classical serial models of computation – so is schema theory modernist (computer metaphor) or postmodernist (a new computational paradigm)? We do not find the labels of modernism and postmodernism useful, at least in discussing neural organization. Indeed, Szentágothai and Érdi (1983), writing when Szentágothai was searching for a third option between eliminative reductionist materialism and the interactions dualism, stated in a footnote:

> It has become (again) an unfortunate practice to label anybody as an "xy-ist" on account of some theory he finds appropriate or applicable. If someone believes that Charles Darwin's theory of natural selection correctly explains certain events that life underwent in the history of [the] planet, he cannot (or ought not) by the same token be labeled a "Darwinist". This would put science on the same footing with religion or political credos that [belong] to different categories. No harm is done, of course, if this were used in a purely colloquial sense, like for example (to stay within neuroscience) if someone were labeled as a "neuronist" or a "reticularist", because these terms carried no philosophical labels.

Keil & Davids reject the notion of "coding" without any constructive suggestion as to what is to take its place. For example, one could paraphrase the work of Hubel and Wiesel (1962; 1967) by saying that they showed that cells of visual cortex "encode" edges in the visual scene. Most students of self-organization in the visual system would see one of their tasks to be to understand why the nervous system should organize itself to extract such features/constitute such a code, rather than rejecting the notion of coding. They would not find it useful to have such work categorized as repudiating Hubel and Wiesel's in favor of postmodernism. **Keil & Davids** state that we assume computation "to be independent of the structural organization and evolutionary development of the nervous system" whereas *Organization* explicitly shows how schemas are distributed across specific neural structures, and discusses evolution (admittedly briefly) in several places. We earlier suggested that **Morasso** sees a sharp dichotomy between schema theory and self-organization where none exists. **Keil & Davids** see such a dichotomy between modernist and postmodernist theories. We see such labeling as neither unambiguous nor as useful. For us, a model is to be understood by its efficacy in illuminating a variety of data. The labeling of the various concepts by "-isms" seems to us irrelevant to the process of model evaluation.

Keil & Davids state that

> Thoughts, emotions, ideas, beliefs, images and, actions are merely the neural traffic constantly being produced between the billions of neurons in the CNS. A postmodernist approach attempts to explain processes of brain and behavior within the boundaries of natural laws. Their delineation should not be screened by the unnecessary introduction of an additional layer of hypothetical constructs, which a modernist framework imposes. Lifting the screen could help us to fully understand brain, behavior and, ultimately, consciousness.

But this is to ignore at least three crucial questions: (1) We are far from knowing what constitutes a successful theory of consciousness. (2) What natural laws are they referring to, and why is it wrong to add an additional layer of hypothetical constructs? For example, are their natural laws those of fundamental physics? We have already seen that statistical mechanics needs hypothetical constructs both to "get started" and to bridge up to the level of thermodynamics. Why should neuroscience be different? (3) How do they decide which aspects of the neural traffic are characterized as thoughts, emotions, ideas, beliefs, images or actions, respectively? We still need to understand the role of different skills and thought processes on the one hand, and the role of different brain regions or neurons on the other – eye movements versus reaching versus dancing, cerebellum versus superior colliculus, and so on. We believe that the only reason that much of neuroscience manages without a computational layer of functional modeling is that most neuroscience research finesses the issue by having a very narrow focus. For example, if one talks of the role of superior colliculus in eye movements, the function "disappears into the woodwork." But such a strategy fails if one tries to integrate a multitude of functions in one's modeling.

Érdi (1993) proposed that neurodynamical system theory may be used to connect structural and functional aspects of neural organization. Systems with feedback connections and the systems of these connected loops can be understood based on the concepts of circular and network causality (Sattler 1986). Generalized causal systems, which are more general than systems characterized by "single cause-single effect" only, and which implement circular and

https://doi.org/10.1017/S0140525X00413361 Published online by Cambridge University Press

network causality, are proper frameworks for describing the self-organizing mechanisms of the nervous system. Such concepts as circular and network causality (and many others, such as chaos, unpredictability, information, emergence, complexity, etc.) challenged the mechanistic paradigms used to describe the behavior of "simple systems." Note, however, that circular causality and information were already key concepts of the cybernetics of the 1940s. In any case, brain theory and evolutionary theory are the prototypes of the "science of complexity" irrespective of "-ism."

Érdi (1996) suggested that hermeneutics, the "art of interpretation," which is neither a priori monist nor dualist, can be applied to the study of the brain. On one side the brain is an "object" of interpretation, on the other side it is itself an interpreter: the brain is a hermeneutic device. He argued that there seems to be a convergence between the "device approach" and the "philosophical approach" to the brain. Systems exhibiting "high" structural complexity and "high" dynamic complexity" (e.g., but not exclusively, chaos) may be candidates for being hermeneutic devices, because they are both object or subject of interpretation and interpreting agents. (It should be recalled, however, that even simple systems may lead to complex dynamics, May 1976). So, the occurrence of chaos is not a sufficient condition for being a hermeneutic device. What can we say, if anything, that we think may be reconcilable with some views of postmodern philosophers? Both natural science as an "objective analyzer" and (post)modern art reiterate the old philosophical question: What is reality? The human brain is not only capable of perceiving what is called objective reality; it can also create new reality. It is a hermeneutic device.

R5.2. Reductionism

For Descartes, the properties of inanimate objects as well as of animal bodies can exclusively be derived from the arrangements of the constituting matter – physiology can be reduced to physics. Based on this approach, Descartes could be qualified as one of the founders of monistic-flavored reductionism. However, Descartes argued that two notions necessary to the operation of the "machine" cannot be described by mechanics: First, Almighty God created the world and gave the initial impetus; second, human thought cannot be "modeled" by automata. Descartes is qualified in this sense as the founder of interactionist dualism (or dualistic interactionism, if you like) emphasizing the interaction between that spatially extended body and a noncorporeal mind (Érdi 1988). The opposite view, monism, is that the human mind is indeed just an aspect of matter. One view of monism then holds that the "laws of the mind" or "the laws of the brain" can be reduced to those of physics. What, then, did we have in mind when we stated (*Organization,* p. *ix*) that

> much of modern neuroscience seems to us excessively reductionist, focusing on the study of ever smaller microsystems to the exclusion of an appreciation of their contribution to the behaving organism. We do not reject the data gained in this way but are concerned with restoring some equilibrium between systems neuroscience, cellular neuroscience, and molecular neuroscience.

This is not an in-principle rejection of reductionism. It is, rather, a rejection of the view that one need focus on only the finest details of the brain (e.g, the molecules and membranes of neurochemistry) and all secrets of the brain will be revealed. Rather, we argue that sense can be made of the fine details only if they are placed within the context of the study of organisms which are the embodiment of evolutionary forces shaped by complex historical accidents, interacting with a complex physical and social environment.

We reject the view that the lowest level of analysis of a system is the "true" level for scientific understanding. Instead, we adopt the notion of "two-way reduction" (Arbib & Hesse 1986) in which, for example, the study of personal experience and neuroscience can act reciprocally to enrich our understanding of both the mind and brain. This is in contrast to the view that psychology, linguistics, anthropology, and social sciences could be reduced to the scientific vocabulary of an existing science, such as neuroscience, in the way that chemical phenomena can be explained by the formalism of physics. Even in the latter case, note that we do *not* use physical terminology to discuss chemical processes in general; we only use it when seeking explanations of generically important chemical processes. Since the relationship between physics and chemistry is paradigmatic for reductionism, we make here a side remark. After the birth of quantum physics, chemistry seemed, briefly, to be reducible to (micro)physicics: "The underlying physical laws for the mathematical theory of large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws lead to equations much too complicated to be soluble." (Dirac 1929). As Golden (1969) later showed, the treatment of chemical reactions needs additional requirements even at the level of quantum statistical mechanics. In a celebrated book, Primas (1983) deeply analyzed why chemistry cannot be reduced to quantum mechanics.

Bookstein quotes the above excerpt from *Organization* and then states that "in eschewing any concern for disease states, the authors are inadvertently no less reductionist than the level-specific approaches they would supersede . . . [with] the reduction of the subject matter to normative (healthy) behaviors." But this is a confusion of reductionism in the sense explained above with "reduce" in the sense of "restricting one's attention." We have already expressed our sympathy with **Bookstein's** emphasis on the study of brains in disease as well as in health. **Bookstein** (PCC) explains that his use of the word "reductionism" comes from arguments in the history of biology, where (e.g., in systematics) the reduction of populations to type specimens is considered an explicit error under the heading of reductionism. But he then offers a bridge between the two usages, averring that reducing the study of a population to the description of a mean is exactly the same as reducing the behavior of water to water molecules, and so on: it is the same systematic omission of other levels of measurement.

Sporns quotes us as saying that a multilevel analysis of brain function "is a necessary complement to the reductionist program of empirical neuroscience" (p. 4). For him, reductionism is in perhaps terminal decline, as the complexity, interconnectedness, and the integrated function of biological systems are increasingly recognized as central and irreducible problems of modern biology. In this spirit, he sees *Organization* pointing the way to a cognitive neuroscience based on a computational and synthetic view of how neural structure and dynamics interrelate across multiple levels.

References

Letters "a" and "r" before authors' initials refer to target article and response, respectively.

Acsády, L., Halasy, K. & Freund, T. F. (1993) Calretinin is present in non-pyramidal cells of the rat hippocampus - III. Their inputs from the median raphe and medial septal. *Neuroscience* 52:829–41. [aMAA]

Adrian, E. D. (1942) Olfactory reactions in the brain of hedgehog. *Journal of Physiology* 10:459–73. [aMAA]

(1950) Sensory discrimination with some recent evidence from the olfactory organ. *British Medical Bulletin* 6:330–31. [aMAA]

Aguirre, L. A. & Billings, S. A. (1995) Identification of models for chaotic systems from noisy data: Implications for performance and nonlinear filtering. *Physica D* 85:239–58. [RAMG]

Aizawa, K. (1996) Some neural network theorizing before McCulloch: Nicolas Rashevsky's mathematical biophysics. In: *Brain processes, theories and models,* ed. R. Moreno-Diaz & J. Mira-Mira. MIT Press. [rMAA]

Allman, J. A. (1999) *Evolving brains.* Scientific American Library. [rMAA]

Almássy, N., Edelman, G. M. & Sporns, O. (1998) Behavioral constraints in the development of neuronal properties: A cortical model embedded in a real world device. *Cerebral Cortex* 8:346–61. [rMAA, OS]

Amari, S. (1974) A method of statistical neurodynamics. *Kybernetik* 14:201–15. [aMAA]

Amos, A. J. (2000) Information processing in the frontal cortex and basal ganglia. *Journal of Cognitive Neuroscience,* (in press). [AJA]

Andersen, P., Bliss, T. V. P. & Skrede, K. K. (1971) Lamellar organization of hippocampal excitatory pathways. *Experimental Brain Research* 113:222–38. [aMAA]

Anderson, J. R. (1983) *The architecture of cognition.* Harvard University Press. [arMAA]

(1993) *Rules of the mind.* Erlbaum. [rMAA]

Anokhin, P. K. (1974) *Biology and neurophysiology of conditioned reflexes and their role in adaptive behaviour.* Pergamon Press. [rMAA, PGM]

Aradi, I., Barna, G., Érdi, P. & Gröbler, T. (1995) Chaos and learning in the olfactory bulb. *International Journal of Intelligent Systems* 10:89–117. [arMAA]

Aradi, I. & Érdi, P. (1996) Multicompartmental modeling of neural circuits in the olfactory bulb. *International Journal of Neural Systems* 7:519–27. [aMAA]

Arbib, M. A. (1984) From synergies and embryos to motor schemas. In: *Human motor actions – Bernstein revisited,* ed. H. T. A. Whiting. Elsevier. [rMAA]

(1985) *In search of the person: Philosophical explorations in cognitive science.* University of Massachusetts Press. [arMAA]

(1989) *The metaphorical brain 2: Neural networks and beyond.* Wiley-Interscience. [arMAA]

(1992) Schema theory. In: *The encyclopedia of artificial intelligence,* ed. S. Shapiro. Wiley. [AC]

(1997) From visual affordances in monkey parietal cortex to hippocampo-parietal interactions underlying rat navigation. *Philosophical Transactions of the Royal Society of London B* 352:1429–36. [rMAA]

(2000) The mirror system, imitation, and the evolution of language. In: *Imitation in animals and artifacts,* ed. C. Nehaniv & K. Dautenhahn. MIT Press. (in press) [rMAA]

Arbib, M. A., Bischoff, A., Fagg, A. H. & Grafton, S. T. (1994) Synthetic PET: Analyzing large-scale properties of neural networks. *Human Brain Mapping* 2:225–33. [arMAA]

Arbib, M. A., Érdi, P. & Szentágothai, J. (1998) *Neural organization: Structure, function and dynamics.* MIT Press. [rMAA]

Arbib, M. A. & Grethe, J., eds. with the Project Team of the University of Southern California Brain Project (2000) *Computing the brain: A guide to neuroinformatics.* Academic Press. [rMAA]

Arbib, M. A. & Hesse, M. B. (1986) *The construction of reality.* Cambridge University Press. [arMAA]

Arhem, P., Blomberg, C. & Liljenstrom, H. (1999) *Disorder versus order in brain function: Essays in theoretical neurobiology.* World Scientific. [RAMG]

Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. (1996) Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* 273:1868–71. [rMAA]

Babloyantz, A. & Destexhe, A. (1986) Low-dimensional chaos in an instance of epilepsy. *Proceedings of the National Academy of Sciences USA* 83:3513–17. [rMAA]

Barna, G. & Érdi, P. (1990) On the neural connectance-performance relationship. In: *Cybernetics and systems 90,* ed. R. Trappl. World Science. [rMAA]

Barna, G., Gröbler, T. & Érdi, P. (1998) Statistical model of the hippocampal CA3 region, II. The population framework: Model of rhythmic activity in the CA3 slice. *Biological Cybernetics* 79:309–21. [arMAA]

Barnes-Holmes, D., Healy, O. & Hayes, S. C. (2000) Relational frame theory and the relational evaluation procedure: Approaching human language as derived relational responding. In: *Issues in experimental and applied analyses of human behavior,* ed. J. C. Leslie & D. Blackman. Context Press. [JCL]

Barone, P. & Joseph, J.-P. (1989) Prefrontal cortex and spatial sequencing in macaque monkey. *Experimental Brain Research* 78:447–64. [rMAA]

Batardiére, A., Barone, P., Dehay, C. & Kennedy, H. (1998) Area-specific laminar distribution of cortical feedback neurons projecting to cat area 17: Quantitative analysis in the adult and during ontogeny. *Journal of Comparative Neurology* 396:493–510. [RK]

Baylis, G. & Driver, J. (1993) Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance* 19(3):451–70. [DH]

Bazsó, F., Szalisznyó, K., Payrits, S. & Érdi, P. (1999) A statistical approach to neural population dynamics: Theory, algorithms, simulations. *Neurocomputing* 26–27:329–34. [rMAA]

Beaudet, A. & Descarries, L. (1978) The monoamine innervation of rat cerebral cortex: Synaptic and nonsynaptic axon terminals. *Neuroscience* 3:851–60. [MNZ]

Bernstein, N. A. (1957) *The coordination and regulation of movement.* Pergamon Press. [PGM]

Bhalla, H. S. & Bower, J. (1993) Exploring parameter space in detailed single neuron models: Simulations of the mitral and granule cells of the olfactory bulb. *Journal of Neurophysiology* 69:1948–65. [aMAA]

Bichot, N. P. & Schall, J. D. (1999) Effects of similarity and history on neural mechanisms of visual selection. *Natural Neuroscience* 2:549–54. [rMAA, AG]

Biedenbach, M. A. (1966) Effects of anesthetics and cholinergic drugs on prepyriform electrical activity in cats. *Experimental Neurology* 16:464–79. [aMAA]

Bienenstock, E. & Geman, S. (1995) Compositionality in neural systems. In: *The handbook of brain theory and neural networks,* ed. M. A. Arbib. MIT Press. [rMAA, SE]

Bizzi, E., Hogan, N., Mussa Ivaldi, F. A. & Giszter, S. F. (1992) Does the nervous system use equilibrium-point control to guide single and multiple movements? *Behavioral and Brain Sciences* 15:603–13. [PGM]

Bliss, T. V. P. & Collingridge, G. L. (1993) A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature* 361:31–39. [aMAA]

Bliss, T. V. P. & Lymo, T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of perforant path. *Journal of Physiology* 232:331–56. [aMAA]

Borisyuk, G., Borisyuk, R., Kazanovich, Y. & Strong, G. (1999) Oscillatory neural networks: Modeling binding and attention by synchronization of neural activity. In: *Oscillations in neural systems,* ed. V. Brown, D. Levine & T. Shirey. Erlbaum. [RB, rMAA]

Braitenberg, V., Heck, D. & Sultan, F. (1997) The detection and generation of sequences as a key to cerebellar function: Experiments and theory. *Behavioral and Brain Sciences* 20:229–45. [PGM]

Bridgeman, B. (1999) Separate representations of visual space for perception and visually guided behavior. In: *Cognitive contributions to the perception of spatial and temporal events,* ed. G. Aschersleben, T. Bachmann & J. Müsseler. Elsevier. [rMAA, BB]

Bridgeman, B., Peery, S. & Anand, S. (1997) Interaction of cognitive and sensorimotor maps of visual space. *Perception and Psychophysics* 59:456–69. [rMAA, BB]

Britten, K. H. (1998) Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Visual Neuroscience* 15:553–58. [aMAA]

Brothers, L. (1997) *Friday's footprint: How society shapes the human mind.* Oxford University Press. [rMAA]

Brooks, R. A. (1989) A robot that walks: Emergent behaviors from a carefully evolved network. *Neural Computation* 1:253–62. [SE]

Brothers, L. & Ring, B. (1992) A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience* 4:107–18. [aMAA]

(1993) Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behavioral Brain Research* 57:53–61. [aMAA]

Buhlmann, P. (1999) Dynamic adaptive partitioning for nonlinear time series. *Biometrika* 86:555–71. [RAMG]

Burgess, N., Recce, M. & O'Keefe, J. (1994) A model of hippocampal function. *Neural Networks* 7:1065–81. [aMAA]

Burgess, N. & O'Keefe, J. (1996) Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus* 6:749–62. [aMAA]

Buzsáki, G. (1989) Two-stage model of memory trace formation: A role for "noisy" brain states. *Neuroscience* 31:551–70. [aMAA]

(1996) The hippocampo-neocortical dialogue. *Cerebral Cortex* 6:81–92. [rMAA]

Buzsáki, G., Bragin, A., Chrobak, J. J., Nádasdy, Z., Sík, A., Hsu, M. & Yiinen, A. (1994) Oscillatory and intermittent synchrony in the hippocampus: Relevance

to memory trace formation. In: *Temporal coding in the brain,* ed. G. Buzsáki, R. Llinás, W. Singer, A. Berthoz & Y. Chrisen. Springer-Verlag. [aMAA]
Castiello, U., Paulignan, Y. & Jeannerod, M. (1991) Temporal dissociation of motor responses and subjective awareness: A study in normal subjects. *Brain* 114:2639–55. [aMAA, BB]
Cepeda, C. & Levine, M. S. (1998) Dopamine and N-methyl-D-aspartate receptor interactions in the neostriatum. *Developmental Neuroscience* 20:1–18. [rMAA]
Chapeau-Blondeau, F. & Chauvet, G. (1992) Stable, oscillatory and chaotic regimes in the dynamics of small neural networks with delay. *Neural Networks* 5:735–43. [RAMG]
Chiel, H. J. & Beer, R. D. (1997) The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neuroscience* 20:553–57. [rMAA]
Chomsky, N. (1965) *Aspects of the theory of syntax.* MIT Press. [aMAA]
Christiansen, F., Cvitanovi, P. & Putkaradze, V. (1997) Spatiotemporal chaos in terms of unstable recurrent patterns. *Nonlinearity* 10:55–70. [RAMG]
Chrobak, J. J. & Buzsáki, G. (1996) High-frequency oscillations in the output networks of the hippocampus-entorhinal axis of the freely-behaving rat. *Journal of Neuroscience* 16:3056–66. [aMAA]
Chugani, H. T., Phelps, M. E. & Mazziotta, I. C. (1987) Positron emission tomography study of human brain functional development. *Annual of Neurology* 22:487–97. [REH]
Clark, A. (1997) *Being there: Putting brain, body and world together again.* MIT Press. [AC]
(1999) Vision and visuomotor action. *Journal of Consciousness Studies* 6(11–12):1–18. [AC]
Collins, A. M. & Loftus, E. F. (1975) A spreading activation theory of cognitive processing. *Psychological Review* 82:407–42. [aMAA]
Collins, A. M. & Quillian, M. R. (1969a) A spreading-activation theory of semantic processing. *Psychological Review* 82:45–73. [aMAA]
(1969b) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8:240–48. [rMAA]
Constantine-Paton, M. & Law, M. I. (1978) Eye-specific termination bands in the tecta of three-eyed frogs. *Science* 202:639–41. [rMAA]
Contopoulos, G. (1998) Dynamical spectra and the onset of chaos. In: *Nonlinear dynamics and chaos: A festschrift in honor of George Contopoulos,* ed. J. R. Buchler, S. T. Gottesman & H. E. Kandrup. New York Academy of Sciences. [RAMG]
Crannell, A. (1995) The role of transitivity in Devaney's definition of chaos. *American Mathematical Monthly* 102:788–93. [RAMG]
Crick, F. (1984) Function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences USA* 81:4586–90. [rMAA]
Dehaene, S. (1993) Temporal oscillations in human perception. *Psychological Sciences* 4:264–70. [RAMG]
Dell, G. S. (1986) A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93:283–321. [rMAA]
Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992) Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91:176–80. [aMAA]
Dirac, P. A. M. (1929) Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London A* 123:243–65. [rMAA]
Dominey, P. F., Arbib, M. A. & Joseph, J.-P. (1995) A model of corticostriatal plasticity for learning associations and sequences. *Journal of Cognitive Neuroscience* 7:311–36. [rMAA]
Dominey, P. F., Lelekov, T., Ventre-Dominey, J. & Jeannerod, M. (1998) Dissociable processes for learning the surface and abstract structure sensorimotor sequences. *Journal of Cognitive Neuroscience* 10(6):734–51. [rMAA]
Dominey, P. F. & Ramus, F. (2000) Neural network processing of natural language: I. Sensitivity to serial, temporal and abstract structure of language in the infant. *Language and Cognitive Processes.* (in press). [rMAA]
Doya, K. (1999) What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex? *Neural Networks* 12:961–74. [rMAA, PGM]
Duchamp-Viret, P., Duchamp, A. & Chaput, M. (1993) GABAergic control of odor-induced activity in the frog olfactory bulb: Electrophysiological study with picrotoxin and bicuculline. *Neuroscience* 53:111–20. [aMAA]
Duncan, J., Humphreys, G. & Ward, R. (1997) Competitive brain activity in visual attention. *Current Opinion in Neurobiology* 7:255–61. [DH]
Eccles, J. C., Ito, M. & Szentágothai, J. (1967) *The cerebellum as a neuronal machine.* Springer-Verlag. [aMAA]
Edelman, S. (1999) *Representation and recognition in vision.* MIT Press. [SE]
Eichenbaum, H., Otto, T. & Cohen, N. J. (1994) Two functional components of the hippocampal memory system. *Behavioral and Brain Sciences* 17:449–517. [SE]
Ellis, A. (1996) *Human cognitive neuropsychology.* Psychology Press. [rMAA]
Elman, J. L. (1990) Finding structure in time. *Cognitive Science* 14:179–211. [REH]
Érdi, P. (1984) System-theoretical approach to the neural organization: Feed-forward control of the ontogenetic development. In: *Cybernetics and systems research, vol. 2,* ed. R. Trappl. Elsevier. [rMAA]
(1988) From brain theory to future generation computer systems. In: *Nature, cognition and system I,* ed. M. E. Carvallo. Kluwer. [rMAA]
(1993) Neurodynamic system theory. *Theoretical Medicine* 14:137–52. [rMAA]
(1996) The brain as a hermeneutic device. *BioSystems* 38:179–89. [arMAA]
(1998) Structure-bases modeling of schemas. *Artificial Intelligence* 101:341–43. [rMAA]
(2000) On the 'Dynamic Brain' metaphor. *Brain and Mind* 1:119–45. [rMAA]
Érdi, P., Aradi, I. & Grôbler, T. (1997) Rhythmogenesis in single cells and population models: Olfactory bulb and hippocampus. *BioSystems* 40:45–53. [aMAA]
Érdi, P., Grôbler, T., Barna, G. & Kaski, K. (1993) Dynamics of the olfactory bulb: Bifurcations, learning, and memory. *Biological Cybernetics* 69:57–66. [aMAA]
Érdi, P., Grôbler, T. & Marton, P. (1992a) On the double architecture of the semantic memory. In: *Nature, cognition and system II,* ed. M. E. Carvallo. Kluwer. [rMAA]
Érdi, P., Gröbler, T. & Toth, J. (1992b) On the classification of some classification problems. *International Symposium on Information Physics, Kyushu Institute of Technology, Iizuka,* 110–17. [rMAA]
Érdi, P., Kepecs, Á., Lengyel, M., Obermayer, K. & Szatmáry, Z. (1998) Dynamics of the hippocampus: Multiple strategies. Paper presented at the International Conference on Neural Information Processing, Kitakyushu, October 21–23, 1998. [aMAA]
Érdi, P. & Tsuda, I. (in press) Hermeneutic approach to the brain: Process versus device? *Theoria et Historia Scientific.* [rMAA]
Fagg, A. H. & Arbib, M. A. (1998) Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 11:1277–303. [aMAA]
Feldman, A. G. & Levin, M. F. (1995) The origin and use of positional frames of references in motor control. *Behavioral and Brain Sciences* 18:723–45. [PGM]
Fleck, L. (1979) *Genesis and development of a scientific fact.* [Original work published in 1936]. University of Chicago Press. [FLB]
Fodor, J. (1983) *The modularity of mind.* MIT Press. [AC]
Freedman, R., Hoffer, R. J., Woodward, D. J. & Puro, D. (1977) Interaction of norepinephrine with cerebellar activity evoked by mossy and climbing fibers. *Experimental Neurology* 55:269–88. [MNZ]
Freeman, W. J. (1975) *Mass action in the nervous system.* Academic Press. [aMAA]
(1978) Spatial properties of an EEG event in the olfactory bulb and cortex. *Electroencephalography and Clinical Neurophysiology* 44:585–605. [aMAA]
(1991) Nonlinear dynamics in olfactory information processing. In: *Olfaction. A model system for computational neuroscience,* ed. J. L. Davis & H. Eichenbaum. MIT Press. [aMAA]
Freeman, W. J. & Barrie, J. M. (1994) Chaotic oscillations and the genesis of meaning in cerebral cortex. In: *Temporal coding in the brain,* ed. G. Buzsaki, R. Llinas, W. Singer, A. Berthoz & Y. Christen. Springer-Verlag. [RB]
Freeman, W. J. & Schneider, W. (1982) Changes in spatial patterns of rabbit olfactory EEG with conditioning to odors. *Psychophysiology* 19:44–56. [aMAA]
Freund, T. F. & Antal, M. (1988a) Septal GABAergic control of the hippocampus: A novel mechanism. *Nature* 336:170–73. [aMAA]
Freund, T. F. & Buzsáki, G. (1996) Interneurons of the hippocampus. *Hippocampus* 6:347–470. [aMAA]
Freund, T. F., Gulyás, A. I., Acsádi, L., Görcs, T. & Tóth, K. (1990) Serotonergic control of the hippocampus via local inhibitory interneurons. *Proceedings of the National Academy of Sciences USA* 87:8501–505. [aMAA]
Friston, K. J. (1994) Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping* 2:56–78. [rMAA]
(1998) The disconnection hypothesis. *Schizophrenia Research* 30:115–25. [rMAA]
Fulton, J. F., ed. (1938) *Physiology of the nervous system.* Oxford University Press. [aMAA]
Geissler, H.-G., Schebera, F.-U. & Kompass, R. (1999) Ultra-precise quantal timing: Evidence from simultaneity thresholds in long-range apparent movement. *Perception and Psychophysics* 61:707–26. [RAMG]
Gerstner, W. & Abbott, L. F. (1997) Learning navigational maps through potentiation and modulation of hippocampal place cells. *Journal of Computational Neuroscience* 4:79–94. [aMAA]
Geyer, S. Matelli, M., Luppino, G., Schleicher, A., Jansen, Y., Palomero-Gallagher, N., Zilles, K. (1998) Receptor autoradiography mapping of mesial motor and premotor cortex of the macaque monkey. *Journal of Computational Neuroscience* 397:231–50. [rMAA]
Gibson, J. J. (1950) *The perception of the visual world.* Houghton-Mifflin. [PGM]

(1966) *The senses considered as perceptual systems.* Allen and Unwin. [arMAA]
Gilbert, C. D. (1983) Microcircuitry of the visual cortex. *Annual Review of Neuroscience* 6:217–47. [SE]
Giles, C. L. & Omlin, C. W. (1994) Pruning recurrent neural networks for improved generalized performance. *IEEE Transactions on Neural Networks* 5:848–51. [REH]
Glass, L. (1995) Chaos in neural systems. In: *The handbook of brain theory and neural networks,* ed. M. A. Arbib. Bradford Books/MIT Press. [RB]
Gnadt, J. W. & Andersen, R. A. (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research* 70:216–20. [rMAA, AG]
Golden, S. (1969) *Quantum statistical foundation of chemical kinetics.* Clarendon. [rMAA]
Goodale, M. A. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neuroscience* 15:20–25. [arMAA, AG]
Gould, E., Reeves, A. J., Graziano, M. S. A. & Gross, C. G. (1999) Neurogenesis in the neocortex of adult primates. *Science* 286:548–52. [AJA]
Grafton, S. T., Fagg, A. H. & Arbib, M. A. (1998) Dorsal premotor cortex and conditional movement selection: A PET functional mapping study. *Journal of Neurophysiology* 79:1092–97. [aMAA]
Gray, C. M. & McCormick, D. A. (1996) Chattering cells: Superficial pyramidal neurons contributing to the generation of synchronous oscillations in the visual cortex. *Science* 274:109–13. [aMAA]
Gray, J. A., Feldon, J. N. P., Rawlins, D. R., Hemsley, D. R. & Smith, A. D. (1991) The neuropsychology of schizophrenia. *Behavioral and Brain Sciences* 24:1–20. [rMAA]
Gregson, R. A. M. (1988) *Nonlinear psychophysical dynamics.* Erlbaum. [RAMG]
(1995) *Cascades and fields in perceptual psychophysics.* World Scientific. [RAMG]
(1999) Narrow parameter windows and analogues of contextual noise in nonlinear psychophysics. *Nonlinear Dynamics: Psychology and Life Sciences* 3:235–58. [RAMG]
Grôbler, T., Barna, G. & Érdi, P. (1998) Statistical model of the hippocampal CA3 region, I. The single-cell module: Bursting model of the pyramidal cell. *Biological Cybernetics* 79:301–308. [arMAA]
Grôbler, T., Marton, P. & Érdi, P. (1991) On the dynamic organization of memory. A mathematical model of associative free recall. *Biological Cybernetics* 65:73–79. [aMAA]
Gromova, E. A. (1980) Emotional memory and its mechanisms. *Nauka* (in Russian). [rMAA, MNZ]
Grunewald, A. (1999) Neurophysiology indicates cognitive penetration of the visual system. *Behavioral and Brain Sciences* 22:379–80. [rMAA, AG]
Grunewald, A., Linden, J. F. & Andersen, R. A. (1999) Responses to auditory stimuli in macaque lateral intraparietal area. I. Effects of training. *Journal of Neurophysiology* 82:330–42. [rMAA, AG]
Guazzelli, A., Corbacho, F. J., Bota, M. & Arbib, M. A. (1998) Affordances, motivation, and the World Graph Theory. *Adaptive Behavior* 6:435–71. [arMAA]
Gulyás, A. I., Görcs, T. J. & Freund, T. F. (1990) Innervation of different peptide-containing neurons in the hippocampus by GABAergic septal afferents. *Neuroscience* 37:31–44. [aMAA]
Han, J. S., Gallagher, M. & Holland, P. (1998) Hippocampal lesions enhance configural learning and the hippocampal system. *Hippocampus* 8:87–108. [rMAA]
Handford, C., Davids, K., Bennett, S. & Button, C. (1997) Skill acquisition in sport: Some applications of an evolving practice ecology. *Journal of Sports Sciences* 15:621–40. [rMAA]
Hasselmo, M. (1994) Runaway synaptic modification in models of cortex: Implications for Alzheimer's disease. *Neural Networks* 7(1):13–40. [rMAA]
Hebb, D. O. (1949) *The organization of behavior.* Wiley. [JCL]
Heinke, D. & Humphreys, G. W. (1998) Modelling emergent attentional properties. In: *Connectionist models in cognitive neuroscience: The 5th neural computation and psychology workshop,* ed. D. Heinke, G. Humphreys & A. Olson. Springer-Verlag. [DH]
Hernandez-Lopez, S., Bargas, J., Surmeier, D. J., Reyes, A. & Galarraga, E. (1997) D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *Journal of Neuroscience* 17:3334–42. [rMAA]
Hinton, G. E. & Shallice, T. (1991) Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review* 98:74–95. [aMAA]
Hoffman, J. E. (1998) Visual attention and eye movements. In: *Attention,* ed. H. Pashler. Psychology Press. [DH]
Hoffmann, R. E. (1987) Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Archives of General Psychiatry* 44:178–88. [rMAA]
(1996) Neural networks, cortical connectivity and schizophrenia psychosis. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Rupin & R. S. Berndt. World Scientific. [rMAA]
Hoffman, R. E. & McGlashan, T. (1993) Parallel distributed processing and the emergence of schizophrenic symptoms. *Schizophrenia Bulletin* 19:119–40. [REH]
(1997) Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. *American Journal of Psychiatry* 154:1683–89. [REH]
Holst, E. & von Mittelstaedt, H. (1950) Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften* 37:464–76. [PGM]
Honer, W. G., Falkai, P., Chen, C., Arango, V., Mann, J. J. & Dwork, A. J. (1999) Synaptic and plasticity-associated proteins in anterior frontal cortex in severe mental illness. *Neuroscience* 91(4):1247–55. [rMAA]
Horn, D., Ruppin, E., Usher, M. & Herrmann, M. (1993) Neural network modeling of memory deterioration in Alzheimer's disease. *Neural Computation* 5:736–49. [rMAA]
Horwitz, B., Tagamets, M.-A. & McIntosh, R. (1999) Neural modeling, functional brain imaging, and cognition. *Trends in Cognitive Science* 3:91–98. [arMAA]
Houk, J. O., Adams, J. L. & Barto, A. G. (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia,* ed. J. C. Houk, J. L. Davis & D. G. Beiser. MIT Press. [rMAA]
Hubel, D. H. & Wiesel, T. N. (1959) Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology* 1481:574–91. [arMAA]
Hubel, D. H. & Wiesel, T. N. (1962) Receptive fields, binocular and functional architecture in the cat's visual cortex. *Journal of Physiology* 160:106–54. [rMAA]
Hubel, D. H. & Wiesel, T. N. (1977) Functional architecture of macaque monkey cortex. *Proceedings of the Royal Society of London B,* 198:1–59. [rMAA]
Huttenlocher, P. R. (1979) Synaptic density in the human frontal cortex developmental changes and effects of aging. *Brain Research* 163:195–205. [REH]
Huttenlocher, P. R. & Dabholkar, A. S. (1997) Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology* 387:167–78. [REH]
Inman, V. T., Ralston, H. J. & Todd, F. (1981) *Human walking.* Williams and Wilkins. [rMAA]
Jahnsen, H. & Llinás, R. (1984) Ionic basis for the electroresponsiveness and oscillatory properties of guinea-pig thalamic neurons in vitro. *Journal of Physiology (London)* 349:227–47. [rMAA]
Jeannerod, M. (1994) The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences* 17:187–245. [AC]
Jensen, O. & Lisman, J. E. (1996) Hippocampal CA3 region predicts memory sequences: Accounting for the past precession of place cells. *Learning and Memory* 3:279–87. [aMAA]
Joyce, J. (1937) *Ulysses.* Bodley Head. [SE]
Just, M. A., Carpenter, P. A. & Varma, S. (1999) Computational modeling of high-level cognition and brain function. *Human Brain Mapping* 8:128–36. [rMAA]
Kaas, J. H. (1993) Evolution of multiple areas and modules within neocortex. *Perspectives on Developmental Biology* 1:101–107. [rMAA]
Kazanovich, Y. B. & Borisyuk, R. M. (1999) Dynamics of neural activity with a central element. *Neural Networks* 12:149–61. [rMAA, RB]
Kiss, T., Orbán, G., Lengyel, M. & Érdi, P. (in press) Theta frequency resonance in CA3 interneurons. In: *Cybernetics and systems 2000,* ed. R. Trappl. World Science. [rMAA]
Koch, C. & Crick, F. (1994) Some further ideas regarding the neuronal basis of awareness. In: *Large-scale neuronal theories of the brain,* ed. C. Koch & J. L. Davis. MIT Press. [AJA]
Krubitzer, L. (1998) Constructing the neocortex: Influences on the pattern of organization in mammals. In: *Brain and mind: Evolutionary perspectives,* ed. M. S. Gazzaniga & J. S. Altman. HFSP. [rMAA]
LaBerge, D. (1998) Attention emphasis in visual orienting and resolving. In: *Visual attention,* ed. R. D. Wright. Oxford University Press. [DH]
Lepages, M., Habib, R. & Tulving, E. (1998) Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus* 8:313–22. [rMAA]
Leslie, J. (1996) *Principles of behavioral analysis.* Harwood Academic. [rMAA]
Leslie, J. C. & O'Reilly, M. (1999) *Behavior analysis: Foundations and applications to psychology.* Harwood Academic. [JCL]
Lev, D. L. & White, E. L. (1997) Organization of pyramidal cell apical dendrites and composition of dendritic clusters in the mouse: Emphasis on primary motor cortex. *European Journal of Neuroscience* 9:280–90. [aMAA]
Levine, D. S. & Prueitt, P. S. (1989) Modeling some effects of frontal lobe damage: Novelty and perseveration. *Neural Networks* 2:100–16. [AJA]
Li, Z. & Hopfield, J. J. (1989) Modeling the olfactory bulb and its neuronal oscillatory processing. *Biological Cybernetics* 61:379–92. [aMAA]

Libet, B., Kobayashi, H. & Tanaka, T. (1975) Synaptic coupling into the production and storage of a neuronal memory trace. *Nature* 258:155–57. [MNZ]
Lieblich, I. & Arbib, M. A. (1982) Multiple representations of space underlying behavior. *Behavioral and Brain Sciences* 5:627–59. [aMAA]
Linden, J. F., Grunewald, A. & Andersen, R. A. (1999) Responses to auditory stimuli in macaque lateral intraparietal area. II. Behavioral modulation. *Journal of Neurophysiology* 82:343–58. [rMAA, AG]
Lopes da Silva, F. H., Pijn, J. P. & Wadman, W. J. (1994) Dynamics of local neuronal networks: Control parameters and state bifurcations in epileptogenesis. *Progress in Brain Research* 102:359–70. [rMAA]
Lorente de Nó, R. (1938) Analysis of the activity of chain of internuncial neurones. *Journal of Neurophysiology* 1:207–44. [aMAA]
Luria, A. R. (1973) *The working brain.* Penguin Books. [rMAA]
Lytton, W. W., Hellman, K. M. & Sutula, T. P. (1998) Computer models of hippocampal circuit changes of the kindling model of epilepsy. *Artificial Intelligence Medicine* 13:81–97. [rMAA]
Maass, W. & Bishop, C. M., eds. (1999) *Pulsed neural networks.* MIT Press. [rMAA]
Mainen, Z. F. & Sejnowski, T. J. (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–506. [rMAA]
Makarenko, V. & Llinás, R. (1998) Experimentally determined chaotic phase synchronization in a neuronal system. *Proceedings of the National Academy of Sciences USA* 95:15747–52. [rMAA]
Mamedov, Z. G. (1987) Changes in the activity of cortical neurons under the influence of biogenic amines. *Neuroscience and Behavioral Physiology* 17:160–67. [MNZ]
Marr, D. (1981) Artificial intelligence: A personal view. In: *Mind design,* ed. J. Haugeland. MIT Press. [rMAA, SE]
(1982) *Vision: A computational investigation into the human representation and processing of visual information.* W. H. Freeman. [rMAA, AG, DH, PGM]
Mata, M., Fink, D. F., Gainer, H., Smith, C. R., Davidson, L., Savak, H., Schwartz, W. J. & Sokoloff, L. (1980) Metabolism in rat posterior pituitary primary reflects sodium pump activity. *Journal of Neurochemistry* 34:213–15. [REH]
Matelli, M., Luppino, G. & Rizzolatti, G. (1985) Patterns of cytochrome oxidase activity in the frontal agranular cortex of macaque monkey. *Behavioral Brain Research* 18:125–37. [aMAA]
Matesz, C., Birinyi, A., Kothalwala, D. S. & Székely, G. (1995) Investigation of the dendritic geometry of brain stem motoneurons with different functions using multivariant statistical techniques in the frog. *Neuroscience* 65:1129–44. [GS]
Maunsell, J. (1995) The brain's visual world: Representation of visual targets in cerebral cortex. *Science* 270:764–69. [SE]
May, R. M. (1976) Simple mathematical models with very complicated dynamics. *Nature* 261:459–67. [rMAA]
McCarthy, R. A. & Warrington, E. K. (1990) *Cognitive neuropsychology: A clinical introduction.* Academic Press. [rMAA]
McCulloch, W. S. & Pitts, W. H. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5:115–33. [BB]
McGlashan, T. H. & Hoffman, R. E. (submitted) Schizophrenia as a disorder of developmentally reduced synaptic connectivity. [REH]
Menschik, E. D. & Finkel, L. H. (1998) Neuromodulatory control of hippocampal function: Towards a model of Alzheimer's disease. *Artificial Intelligence in Medicine* 13:99–121. [rMAA]
Meyer-Lindenberg, A., Ziemann, U., Hajak, G., Cohen, L. & Berman, K. (1999) A direct demonstration of nonlinear dynamics in the human brain using PET and transcranial magnetic stimulation (TMS). In: *Fifth International Conference on Functional Mapping of the Human Brain, Poster 287.* http://www.apnet.com/hbm99/methphyscog.htm [rMAA]
Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action.* Oxford University Press. [BB, AC]
Milnor, J. (1992) Remarks on iterated cubic maps. *Experimental Mathematics* 1:5–24. [RAMG]
Mishkin, M., Ungerleider, L. G. & Macko, K. A. (1983) Object vision and spatial vision: Two cortical pathways. *Trends in Neuroscience* 6:414–17. [aMAA]
Mishkin, M., Varga-Khadem, F. & Gaian, N. (1998) Amnesia and the organization of the hippocampal system. *Hippocampus* 8:212–16. [aMAA]
Molchanov, A. M. (1967) A possible role of oscillatory processes in evolution. In: *Proceedings of the First Symposium on Oscillatory Processes in Biological and Chemical Systems, Pushchino, USSR, March 1966.* Nauka. (in Russian). [RB]
Montague, P. R., Dayan, P. & Sejnowski, T. J. (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience* 16:1936–47. [rMAA]
Morasso, P., Baratto, L., Capra, R. & Spada, G. (1999) Internal models in the control of posture. *Neural Networks* 12:11173–80. [rMAA]
Morasso, P. & Sanguineti, V. (1997) *Self-organization, cortical maps and motor control.* North Holland. [rMAA, PGM]
Morris, R. G. M. (1984) Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods* 11:47–60. [aMAA]
Mountcastle, V. B. (1957) Modalities and typographic properties of single neurones of the cat's sensory cortex. *Journal of Neurophysiology* 20:408–34. [aMAA]
(1997) The columnar organization of the neocortex. *Brain* 120:701–22. [aMAA]
Muakkassa, K. F. & Strick, P. L. (1979) Frontal lobe inputs to primate motor cortex: Evidence for four somatotopically organized "premotor" areas. *Brain Research* 177:176–82. [aMAA]
Murphy, K. M., Jones, D. G., Fensternmaker, S. B., Pegado, V. D., Kiorpes, L. & Movshon, J. A. (1998) Spacing of cytochrome oxidase blobs in visual cortex of normal and strabismic monkeys. *Cerebral Cortex* 8:237–44. [aMAA]
Nadim, F., Olsen, O. H., De Schutter, E. & Calabrese, R. L. (1995) Modeling the leech heartbeat elemental oscillator I. Interaction of intrinsic and synaptic currents. *Journal of Computational Neuroscience* 2:215–35. [rMAA]
Newell, A. (1990) *Unified theories of cognition.* Harvard University Press. [rMAA, PGM]
Nicoll, R. & Jhr, C. E. (1982) Self-excitation of olfactory bulb neurones. *Nature* 196:441–44. [aMAA]
Nowycky, M. C., Mori, K. & Shepherd, G. M. (1981) GABAergic mechanisms of dendrodentritic synapses in isolated turtle olfactory bulb. *Journal of Neurophysiology* 46:639–48. [aMAA]
O'Keefe, J. & Nadel, L. (1978) *The hippocampus as a cognitive map.* Clarendon Press. [aMAA]
O'Keefe, J. & Recce, M. (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–30. [aMAA]
Olds, M. E. & Olds, J. (1963) Approach-avoidance analysis of rat diencephalon. *Journal of Comparative Neurology* 120:259–95. [rMAA, MNZ]
Ottoson, D. (1959) Studies of slow potentials in the rabbit's olfactory bulb and nasal mucosa. *Acta Physiology Scandanavia* 47:136–48. [aMAA]
Parkin, A. J. (1999) *Explorations in cognitive neuropsychology.* Psychology Press. [rMAA]
Pavlides, C. & Winson, J. (1989) Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *Journal of Neuroscience* 9:2907–18. [aMAA]
Piaget, J. (1963) *The origin of intelligence in children.* Norton Press. [PGM]
Pickering, J. (1997) Beyond cognitivism: Mutualism and postmodern psychology. In: *Brain, mind and physics,* ed. P. Pylkkanen, P. Pylkko & A. Hautamaki. IOS Press. [DK]
Pinker, S. & Bloom, P. (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–84. [aMAA]
Plaut, D. C. & Shallice, T. (1993) Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology* 10:377–500. [aMAA]
Prager, J. M. & Arbib, M. A. (1982) Computing the optic flow: The MATCH algorithm and prediction. *Vision, Graphics and Image Processing* 24:271–304. [rMAA]
Prazdny, K. (1985) Detection of binocular disparities. *Biological Cybernetics* 52:387–95. [rMAA]
Primas, H. (1983) *Chemistry, quantum mechanics, and reductionism.* Springer-Verlag. [rMAA]
Quartz, S. R. & Sejnowski, T. J. (1997) The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences* 20:537–96. [arMAA]
Quillian, M. R. (1968) Semantic memory. In: *Semantic information processing,* ed. M. L. Minsky. MIT Press. [aMAA]
Rakic, P. (1988) Specification of cerebral cortical area. *Science* 241:170–76. [rMAA]
Rakic, P. Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N. & Goldman-Rakic, P. S. (1986) Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science* 232:232–35. [REH]
Redish, A. D. (1999) *Beyond the cognitive map. From place cells to episodic memory.* MIT Press. [rMAA]
Reeke, G. N., Jr., Sporns, O. & Edelman, G. M. (1990) Synthetic neural modeling: The "Darwin" series of recognition automata. *Proceedings of the IEEE* 78:1498–530. [OS]
Rieke, F., Warland, D., de Ruyter van Steveninck, R. & Bialek, W. (1997) *Spikes: Exploring the neural code.* MIT Press. [rMAA, OS]
Rizzolatti, G. & Arbib, M. A. (1998) Language within our grasp. *Trends in Neuroscience* 21(5):188–94. [arMAA]
Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. & Matelli, M. (1988) Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research* 71:491–507. [aMAA]
Rizzolatti, G., Fadiga, L., Gallese, V. & Fogaassi, L. (1996a) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3:131–41. [aMAA]
Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Perani, D. & Fazio, F. (1996b)

Localization of grasp representations in humans by positron emission tomography: 1. Observation versus execution. *Experimental Brain Research* 111:246–52. [aMAA]

Rizzolatti, G., Luppino, G. & Matelli, M. (1998) The organization of the cortical motor system: New concepts. *Electroencephalography and Clinical Neurophysiology* 106:283–96. [PGM]

Roland, P. (1993) *Brain activation.* Wiley-Liss. [REH]

Rolls, E. T. & O'Mara, S. (1993) Neurophysiological and theoretical analysis of how the primate hippocampus functions in memory. In: *Brain mechanisms of perception and memory: From neuron to behavior,* ed. T. Ono, L. R. Squire, M. E. Raichle, D. T. Perrett & M. Fukuda. Oxford University Press. [rMAA]

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S. & Rauschecker, J. P. (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Natural Neuroscience* 2:1131–36. [AG, rMAA]

Rozin, P. (1976) The evolution of intelligence and access to the cognitive unconscious. *Progress in Psychobiology, Physiology and Psychology* 6:245–80. [aMAA]

Rumelhart, D. E. & McClelland, J. L. (1982) An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review* 89:60–94. [rMAA]

(1986) On learning the past tenses of English verbs. In: *Parallel distributed processing,* ed. D. E. Rumelhart & J. L. McClelland. MIT Press. [rMAA]

Sakata, H., Taira, M., Murata, A. & Mine, S. (1992) Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex* 5:429–38. [aMAA]

Samuel, A. L. (1959) Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3:210–29. [rMAA]

Sattler, R. (1986) *Biophilosophy. Analytic and holistic perspectives.* Springer-Verlag. [rMAA]

Schacter, D. L. & Wagner, A. D. (1999) Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* 9:7–24. [rMAA]

Scheibel, M. E. & Scheibel, A. B. (1958) Structural substrates for integrative patterns in the brain stem reticular core. In: *Reticular formation of the brain,* ed. H. H. Jasper, L. P. Proctor, R. S. Knighton, W. C. Noskay & R. T. Costello. Little, Brown. [aMAA]

Schiff, S. J., Jergens, K., Duong, D. H., Chang, T., Spano, M. L. & Ditto, W. L. (1994) Controlling chaos in the brain. *Nature* 370:615–20. [rMAA]

Schultz, W. (1998) Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* 80:1–27. [rMAA]

Schultz, W., Dayan, P. & Montague, P. R. (1997) A neural substrate of prediction and reward. *Science* 275:1593–99. [rMAA]

Schweighofer, N., Arbib, M. A. & Kawato, M. (1998a) Role of the cerebellum in reaching quickly and accurately: I. A functional anatomical model of dynamics control. *European Journal of Neuroscience* 10:86–94. [rMAA]

Schweighofer, N., Spoelstra, J., Arbib, M. A. & Kawato, M. (1998b) Role of the cerebellum in reaching quickly and accurately: II. A detailed model of the intermediate cerebellum. *European Journal of Neuroscience* 10:95–105. [rMAA]

Sechenov, I. M. (1863) *Reflexes of the brain* (In Russian. English translation by S. Belsky). MIT Press, 1965. [rMAA]

Selemon, I. D. & Goldman-Rakic, P. S. (1999) The reduced neuropil hypothesis: A circuit based model of schizophrenia. *Biological Psychiatry* 45(1):17–25. [rMAA]

Sherrington, C. S. (1906) *The integrative action of the nervous system.* Yale University Press. [aMAA]

Siekmeier, P. J. & Hoffman, R. E. (submitted) Enhanced semantic priming in schizophrenia: A computer model based on excessive pruning of local connections in association cortex. [REH]

Skarda, C. A. & Freeman, W. J. (1987) How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences* 10:161–95. [aMAA]

Snyder, L. H., Batista, A. P. & Andersen, R. A. (1997) Coding of intention in the posterior parietal cortex. *Nature* 386:167–70. [AG]

Somogyi, P. (1977) A specific axonal interneuron in the visual cortex of the rat. *Brain Research* 136:345–50. [aMAA]

Somogyi, P. & Cowey, A. (1981) Combined Golgi and electron microscopic study on the synapses formed by double bouquet cells in the visual cortex of cat and monkey. *Journal of Comparative Neurology* 195:547–66. [aMAA]

Somogyi, P.,Hodgson, A. J. & Smith, A. D. (1979) An approach to tracing neuron networks in the cerebral cortex and basal ganglia. Combination of Golgi staining, retrograde transport horseradish peroxidase and anterograde degeneration of synaptic boutons in the same material. *Neuroscience* 4:1805–52. [aMAA]

Somogyi, P., Tamás, G., Lujan, R. & Buhl, E. H. (1998) Salient features of synaptic organization in the cerebral cortex. *Brain Research Review* 26:113–35. [arMAA]

Sperry, R. W. (1969) A modified concept of consciousness. *Psychological Review* 76:195–206. [aMAA]

Sporns, O., Tononi, G. & Edelman, G. M. (2000) Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex* 10:127–41. [OS]

Squire, L. R. & Zola, S. M. (1998) Episodic memory, semantic memory, and amnesia. *Hippocampus* 8:205–11. [aMAA]

Staddon, J. E. R. (1999) Theoretical behaviorism. In: *Handbook of behaviorism,* ed. W. O'Donohue & R. Kitchener. Academic Press. [AJA]

Stadler, M. P. & Kruse, P., eds (1994) *Multistability in cognition.* Springer Series in Synergetics. Springer. [rMAA]

Stein, L., Xue, B. G. & Belluzzi, J. D. (1993) A cellular analog of operant conditioning. *Journal of the Experimental Analysis of Behavior* 60:41–53. [rMAA]

Stern, C. E. & Hasselmo, M. E. (1999) Bridging the gap: Integrating cellular and magnetic resonance imaging studies of the hippocampus. *Hippocampus* 9:45–53. [aMAA]

Stern, E. A., Kincaid, A. E. & Wilson, C. J. (1997) Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons *in vivo. Journal of Neurophysiology* 77:1697–715. [rMAA]

Stuss, D. T., Eskes, G. A. & Fournet, N. (1994) Experimental neuropsychological studies of frontal lobe functions. In: *Handbook of neuropsychology, vol. 9,* ed. F. Boller & J. Grafman. Elsevier. [AJA]

Sudakov, K. V. (1997) The theory of functional systems: General postulates and principles of dynamic organization (dedicated to the Anokhin Centenary). *Integrative Physiological and Behavioral Science* 32:392–414. [RB, rMAA]

Suri, R. E., Bargas, J. & Arbib, M. A. (submitted) Modeling functions of striatal dopamine modulation in learning and planning. [rMAA]

Suri, R. E. & Schultz, W. (1998) Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research* 121:350–54. [rMAA]

(1999) A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91:871–90. [rMAA]

Sutton, R. S. & Barto, A. G. (1990) Time derivative models of Pavlovian reinforcement. In: *Learning and computational neuroscience: Foundations of adaptive networks,* ed. M. Gabriel & J. Moore. MIT Press. [rMAA]

(1998) *Reinforcement learning: An introduction.* MIT Press. [rMAA]

Swindale, N. V. (1990) Is the cerebral cortex modular? *Trends in Neurosciences* 13:487–92. [aMAA]

(1998) Cortical organization: Modules, polymaps and mosaic. *Current Biology* 8:R270–73. [aMAA]

Szatmáry, Z., Lengyel, M., Érdi, P. & Obermayer, K. (submitted) A model of place field formation based on recurring consecutiveness of perceptual patterns. [aMAA]

Székely, G. (1976) The morphology of motoneurons and dorsal root fibers in the frog spinal cord. *Brain Research* 103:275–90. [GS]

(1989) Ontogeny and morphology of neural structures controlling tetrapod locomotion. In: *Complex organismal functions: Integration and evolution in vertebrates,* ed. P. B. Wake & G. Roth. Wiley. [rMAA]

Székely, G. & Czéh, G. (1971) Activity of spinal cord fragments and limbs deplanted in the dorsal fin of Urodele larvae. *Acta Physiologia Academy of Sciences Hungary* 40:303–12. [aMAA]

Székely, G., Nagy, I., Wolf, E. & Nagy, P. (1989) Spatial distribution of pre- and post- synaptic sites of axon terminals in the dorsal horn of the frog spinal cord. *Neuroscience* 29:175–88. [GS]

Szentágothai, J. (1967) The anatomy of complex integration units in the nervous system. In: *Recent development of neurobiology in Hungary I. Results in neuroanatomy, neuropharmacology and neurophysiology,* ed. K. Lissák. Akadémiai Kiadó. [aMAA]

(1969) Architecture of the cerebral cortex. In: *Basic mechanisms of the epilepsics,* ed. H. H. I. Jasper, A. A. Ward, Jr. & A. Pope. Little, Brown. [aMAA]

(1978) Specificity versus (quasi-) randomness revisited. *Acta Morphologica Hungary* 38:159–67. [GS]

(1979) *Unified brain theory: Utopy or reality?* Magyar Tudomány (in Hungarian). [rMAA]

(1982) Too "much" and too "soon." *Acta Biologica Academy of Sciences Hungary* 33:107–26. [aMAA]

(1983) The modular architectonic principle of neural centers. *Review of Physiology, Biochemistry and Pharmacology* 98:11–61. [aMAA]

(1984) Downward causation? *Annual Review of Neuroscience* 7:1–11. [aMAA]

(1993) Self organization: The basic principle of neural function. *Theoretical Medicine* 14:101–16. [aMAA]

Szentágothai, J. & Arbib, M. A. (1975) *Conceptual models of neural organization.* MIT Press. [aMAA]

Szentágothai, J. & Érdi, P. (1983) *Outline of a general brain theory.* Central Research Institute for Physics of the Hungarian Academy of Science KFKi-1983/117. [rMAA]

(1989) Self-organization in the nervous system. *Journal of Social Biological Structures* 12:367–84. [aMAA]
Taira, M., Mine, S., Georgopoulos, A. P., Murata, A. & Sakata, H. (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research* 83:29–36. [aMAA]
Thelen, E., Schöner, G., Scheier, C. & Smith, L. B. (2000) The dynamics of embodiment: A field theory of infant perserverative reaching. *Behavioral and Brain Sciences.* (in press). [rMAA, OS]
Thelen, E. & Smith, L. B. (1994) *A dynamic systems approach to the development of cognition and action.* MIT Press. [rMAA, AC]
Thompson, K. G. & Schall, J. D. (1999) The detection of visual signals by macaque frontal eye field during masking [see comments]. *Natural Neuroscience* 2:283–88. [rMAA, AG]
Tolman, E. C. (1932) *Purposive behavior in animals and man.* Century. [aMAA]
Tononi, G., Edelman, G. M. & Sporns, O. (1998) Complexity and coherency: Integrating information in the brain. *Trends in Cognitive Sciences* 2:474–84. [OS]
Traub, R. D., Jefferys, J. G. & Whittington, M. A. (1997) Simulation of gamma rhythms in networks of interneurons and pyramidal cells. *Journal of Computational Neuroscience* 4:141–50. [aMAA]
(1999) *Fast oscillations in cortical circuits.* MIT Press. [rMAA]
Traub, R. D. & Miles, R. (1991) *Neuronal networks of the hippocampus.* Cambridge University Press. [arMAA]
(1992) Modeling hippocampal circuitry using data from whole cell patch clamp and dual intracellular recordings in vitro. *Seminars in the Neurosciences* 4:27–36. [aMAA]
Traub, R. D., Miles, R., Muller, R. U. & Gulyás, A. I. (1992) Functional organization of the hippocampal CA3 region: Implications for epilepsy, brain waves and spatial behavior. *Network* 3:465–88. [aMAA]
Traub, R. D., Wong, K. S. & Miles, R. (1987) In vitro models of epilepsy. In: *Neurotransmitters and epilepsy,* ed. P. C. Jobe & H. E. Laird, II. The Humana Press. [aMAA]
Treves, A. & Rolls, E. T. (1994) Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4:374–91. [rMAA]
Tsuda, I. (1991) Chaotic itineracy as a dynamical basis of hermeneutics in brain and mind. *World Futures* 32:167–84. [rMAA]
(1992) Dynamic link of memory - chaotic memory map in nonequilibrium neural networks. *Neural Networks* 5:313–26. [rMAA]
(1996) A new type of self-organization associated with chaotic dynamics in neural networks. *International Journal of Neural Systems* 7:451–59. [rMAA]
Tulving, E. & Markowitsch, H. J. (1998) Episodic and declarative memory: Role of the hippocampus. *Hippocampus* 8:198–204. [aMAA]
Turing, A. M. (1936) On computable numbers with an application to the entscheidungs problem. *Proceedings of the London Mathematical Society (Series 2)* 42:230–65. [BB]
Turner, D. A., Buhl, E. H., Hailer, N. P. & Nitsch, R. (1998) Morphological features of the entorhinal-hippocampal connection. *Progress in Neurobiology* 55:537–62. [aMAA]
Ungerleider, L. G. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior,* ed. D. J. Ingle, M. A. Goodale & R. J. W. Mansfield. MIT Press. [arMAA, AG]
Ungerstedt, U. (1971) Stereotaxic mapping of the monoamine pathway in the rat brain. *Acta Physiologia Scandinavica,* Suppl. 367:1–48. [MNZ]
Van Vreeswijk, C. & Sompolinsky, H. (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274:1724–26. [rMAA]
Varga-Khadem, F., Gadian, D. G., Watkin, K. E., Conelly, A., Van Paesschen, W. & Mishkin, M. (1997) Differential effects of early hippocampal pathology, on episodic and semantic memory. *Science* 277:376–80. [aMAA]
Ventriglia, F. (1974) Kinetic approach to neural systems, I. *Bulletin of Mathematical Biology* 36:534–44. [aMAA]
(1994) Toward a kinetic theory of cortical-like neural field. In: *Neural modeling and neural networks,* ed. F. Ventriglia. Pergamon Press. [aMAA]
Von der Malsburg, C. (1985) Nervous structures with dynamical links. *Berliner Bunsen-Gesellschaft für Physicalische Chemie* 89:700–709. [SE]
Vygotsky, L. S. (1962) *Thought and language* (trans. from Russian original of 1934). MIT Press. [rMAA]
Wallen, P., Ekeberg, O., Lansner, A., Brodin, L., Traven, H. & Grillner, S. (1992) A computer based model for realistic simulations of neural networks. II. The segmental network generating locomotor rhythmicity in the lamprey. *Journal of Neurophysiology* 68:1939–50. [rMAA]
Wallenstein, G. V. & Hasselmo, M. E. (1997) GABAergic modulation of hippocampal activity: Sequence learning, place field development, and the phase procession effect. *Journal of Neurophysiology* 78:393–408. [aMAA]
Wang, X. J. (1999) Fast burst firing and short-term synaptic plasticity: A model of neocortical chattering neurons. *Neuroscience* 89:347–62. [aMAA]
Wang, X. J. & Buzsáki, G. (1996) Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of Neuroscience* 16:6402–13. [aMAA]
Weitzenfeld, A., Arbib, M. A. & Alexander, A. (2000) *Neural simulation language.* MIT Press. [RK]
Whittington, M. A., Traub, R. D. & Jefferys, J. (1995) Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature* 370:612–15. [aMAA]
Wilkins, K. & Wakefield, J. (1995) Brain evolution and neurolinguistic preconditions. *Behavioral and Brain Sciences* 18:161–226. [aMAA]
Wilson, C. J. (1998) Basal ganglia. In: *The synaptic organization of the brain,* ed. G. M. Shepherd. Oxford University Press. [RK]
Wilson, H. R. & Cowan, J. (1973) A mathematical theory of the functional dynamics of cortical and thalamic neurons tissue. *Kybernetik* 13:55–80. [aMAA]
Wilson, M. A. & McNaughton, B. L. (1994) Reactivation of hippocampal ensemble memories during sleep. *Science* 265:676–79. [aMAA]
Young, M. P. (1992) Objective analysis of the topological organization of the primate visual system. *Nature* 358:152–54. [rMAA, RK]
(1993) The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society of London B* 252:13–18. [rMAA, RK]
Záborszky, L., Palkovits, M. & Flerkó, B. (1992) A life-time adventure with the brain. An appreciation of his eightieth birthday. *Journal of Computational Neurology* 326:1–6. [aMAA]
Zaks, M. A., Park, E. H., Rosenblum, M. G. & Kurths, J. (1999) Alternating locking ratios in imperfect phase synchronization. *Physical Review Letters* 82:4228–31. [rMAA]
Zhadin, M. N. (1977) Model of conditioned reflex formation and analysis of functional significance of electrophysiological correlates of learning. *Journal of Higher Nervous Activity* 27:949–56. (in Russian). [MNZ]
(1987) Electrophysiological manifestations of the monoaminergic system's effects on the cerebral cortex. *Neuroscience and Behavioral Physiology* 17:152–60. [MNZ]
(1991) Biophysical mechanisms of the EEG formation. In: *Mathematical approaches to brain functioning diagnostics,* ed. I. Dvorak & A. Holden. Manchester University Press. [MNZ]
(1993) Possible mechanism of the action of biogenic amines on the activity of cortical neurones. *Biophysics* 38:353–58. [MNZ]
Zhadin, M. N. & Bakharev, B. V. (1987) Model of variations in the level of cortical neuron excitation at increased biogenic amine concentration. *Studia Biophysica* 121:81–88. [MNZ]
Zhadin, M. N. & Karpuk, N. N. (1996) Influence of serotonin on cross-correlation in neuronal activity in surviving slices of the cerebral cortex. *Journal of Higher Nervous Activity* 46:547–51. (in Russian). [MNZ]
Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. (1998) Interpreting neuronal population activity by reconstruction: Unified framework with application to hippocampal place cells. *Journal of Neurophysiology* 79:1017–44. [aMAA]
Zilles, K. & Clarke, S. (1997) Architecture, connectivity, and transmitter receptors of human extrastriate visual cortex. In: *Cerebral cortex, vol. 12,* ed. K. Rockland. Plenum Press. [rMAA]