

## **Meanings-based Networks: A New Learning Paradigm for ART Network Systems Models**

An LCNTR Tech Brief Prepared by

Richard B. Wells

May, 2007

### **I. The Problem of the Neural Code**

This tech brief concerns the problem of the neural code and a new approach to the neural coding problem. Briefly stated, the problem of the neural code is this: What is the organization of the system or systems of information coding in the brain? The problem was first raised in 1956 by John von Neumann and to this day has remained one of the outstanding unsolved problems of theoretical neuroscience. Von Neumann saw the neural coding problem as analogous to a “brain language.”

As pointed out, the nervous system is based on two types of communications: those which do not involve arithmetical formalisms, and those which do, i.e. communications of orders (logical ones) and communications of numbers (arithmetical ones). The former may be described as languages proper, the latter as mathematics.

It is only proper to realize that language is largely a historical accident. . . Just as languages like Greek or Sanskrit are historical facts and not logical necessities, it is only reasonable to assume that logics and mathematics are similarly historical, accidental forms of expression. They may have essential variants, i.e. they may exist in other forms than the ones to which we are accustomed. Indeed, the nature of the central nervous system and of the message systems that it transmits indicate positively that this is so. We have now accumulated sufficient evidence to see that whatever language the central nervous system is using, it is characterized by less logical and arithmetical depth than what we are normally used to (Neumann, 2000).

Since von Neumann wrote these words in 1956, theoretical neuroscience has examined only a few hypotheses regarding what form this “brain language” might take. The oldest one, proposed by von Neumann himself, is the *firing rate model*. In this model, it is speculated that information is conveyed by being encoded in the rate at which action potentials are generated by neurons. This hypothesis is closely tied to an analogy with man-made data communication systems, especially those that use frequency-shift-keying modulation methods. This model works well enough in some cases, such as in the transmission of sensory information from nerve endings in the peripheral nervous system, but there is now a large body of evidence indicating that this model is not adequate for explaining a great deal of what appears to be going on in the central systems.

A second hypothesis, hinted at by von Neumann but not developed until the work of von der Malsburg in 1982, holds that information is conveyed in the brain through correlations of neural firing patterns. This hypothesis is known as the *correlation theory of brain function* (Malsburg,

1994). Malsburg proposed this model largely in order to deal with certain mathematical difficulties the firing rate model is incapable of handling, but it received some later and indirect support from a model developed independently by Damasio, which holds that entities and events are represented in the brain by time-locked, synchronous neural firing patterns (Damasio, 1989a, 1989b). This model has since become known as the *convergence zone* hypothesis (Damasio, 1994, 1999).

Traditionally, theoretical neuroscience has proceeded by adopting the firing rate model or the correlation model as a starting point and then proposed neural network models based on this presupposition. These researches have had numerous successes (otherwise these approaches would have been abandoned long ago), but have run into unsolved difficulties as well. It is not unfair to say that in many ways the tradition of presupposing a signal representation for the neural code and then building up a theory from such a supposition is like trying to understand mountains by examining them rock by rock and without recourse to the science of geology. Given enough rock gathering and enough patience, one would expect to eventually understand a particular mountain, and it is certainly easy to recognize a rock when you see one. But it is better by far to try to see the whole mountain first and develop a theory from that starting point.

The neuroscience equivalent to viewing the whole of a mountain is called psychology. From psychological studies we do not obtain information regarding neuronal signals or neural anatomy and physiology. What we obtain is the gross picture of the overall end result as this is exhibited through the observable behaviors and self-reports of the human subject. Within psychology, the subdiscipline most closely linked to adaptive and learning phenomena is *developmental* psychology. The person generally regarded as the father of developmental psychology was the great Swiss psychologist Jean Piaget. Over his sixty years of experimental research, Piaget and his collaborators put together what is perhaps the most well-developed and self-consistent *system theory* to be found anywhere in psychology. Piaget's theory covers human psychological development from infancy to adulthood, and finds that the same fundamental psychological processes at work in the infant are also the processes most fundamentally underlying human intelligence at the adult level (Piaget, 1975), (Piaget et al., 1977).

One of Piaget's major findings was that logic and logical thinking in humans is the outgrowth of a primitive proto-logic operating in infancy at the sensorimotor level of development and before the development of language. This proto-logic is revealed in behavioral forms Piaget called the *logic of actions* and the *logic of meanings*. The vast corpus of Piaget's work documents in step-by-step detail the advancement from these most elementary forms of *practical* logics to the capacity for logico-mathematical thinking humans develop by age 15 years. Bärbel Inhelder,

Piaget's long-time collaborator, wrote in the preface to his last book,

Piaget intended to bring to light the very roots of logic by going back to implications between sensorimotor actions. Such a logic could only be a logic of meanings where implications are not restricted to statements: in the subject's view, every action or operation is endowed with meanings; therefore, one may deal with systems of implications among the meanings of actions, and then among the meanings of operations. Provided that the meaning of the actions and the causality of the actions are carefully distinguished, the subject's expectations and anticipations about the chaining of actions bear witness to the existence of early inferences. Hence a privileged form of inference is the action implication, which is an implication between the meanings of actions. Piaget was thus initiating research on a "protologic" in which forms and contents are less differentiated than in operatory systems. After his investigations of correspondences, elementary dialectics, and categorical logic, in which he brought out the elementary and formative stages of operations, Piaget was deepening his study of the modes of understanding that are used as cognitive tools well before the subject can thematize them (Piaget and Garcia, 1991, pp. vii-viii).

In simple terms, in the sensorimotor development stage of infancy the meaning of an action is what the action leads to, the meaning of an object is what can be done with the object. Piaget's finding implies that all the complex meanings we understand as adults have their point of origin in the extremely simple and practical logic of meanings supported by the infant's elemental logic of actions. His findings are based on observable behaviors in young children, which in the sensorimotor stage of development can be classified into three basic transformational structures.

To express the same idea in still another way, I think that human knowledge is essentially active. To know is to assimilate reality into systems of transformations. To know is to transform reality in order to understand how a certain state is brought about. By virtue of this point of view, I find myself opposed to the view of knowledge as a copy, a passive copy, of reality. In point of fact, this notion is based on a vicious circle: in order to make a copy we have to know the model that we are copying, but according to this theory of knowledge the only way to know the model is by copying it, until we are caught in a circle, unable ever to know whether our copy of the model is like the model or not. To my way of thinking, knowing an object does not mean copying it – it means acting on it. It means constructing systems of transformations that correspond, more or less adequately, to reality. . . . The transformational structures of which knowledge consists are not copies of the transformations in reality; they are simply possible isomorphic models among which experience can enable us to choose. Knowledge, then, is a system of transformations that become progressively adequate (Piaget, 1970, pg. 15).

In the early 1960s, after his discovery that the sensorimotor actions and behaviors that effect these transformations fit into a three-fold system of classification, Piaget was astounded to learn that these three types of structures corresponded neatly with the fundamental mathematical structures found by mathematicians – in particular, by an interesting group of young, mostly French mathematicians known collectively as the Bourbaki mathematicians – to be irreducible to one another and from which all the rest of mathematics can be constructed. These three mathematical structures, called *algebraic structure*, *topological structure*, and *order structure*, were called the Bourbaki mother structures.

A number of years ago I attended a conference outside Paris entitled "Mental Structures and

Mathematical Structures.” This conference brought together psychologists and mathematicians for discussion of these problems. For my part, my ignorance of mathematics then was even greater than what I admit to today. On the other hand, the mathematician Dieudonné, who was representing the Bourbaki mathematicians, totally mistrusted anything that had to do with psychology. Dieudonné gave a talk in which he described the three mother structures. Then I gave a talk in which I described the structures I had found in children’s thinking, and to the great astonishment of us both we saw that there was a very direct relationship between these three mathematical structures and the three structures of children’s operational thinking. We were, of course, impressed with each other, and Dieudonné went so far as to say to me: “This is the first time that I have taken psychology seriously. It may also be the last, but at any rate it’s the first.” (Piaget, 1970, pg. 26).

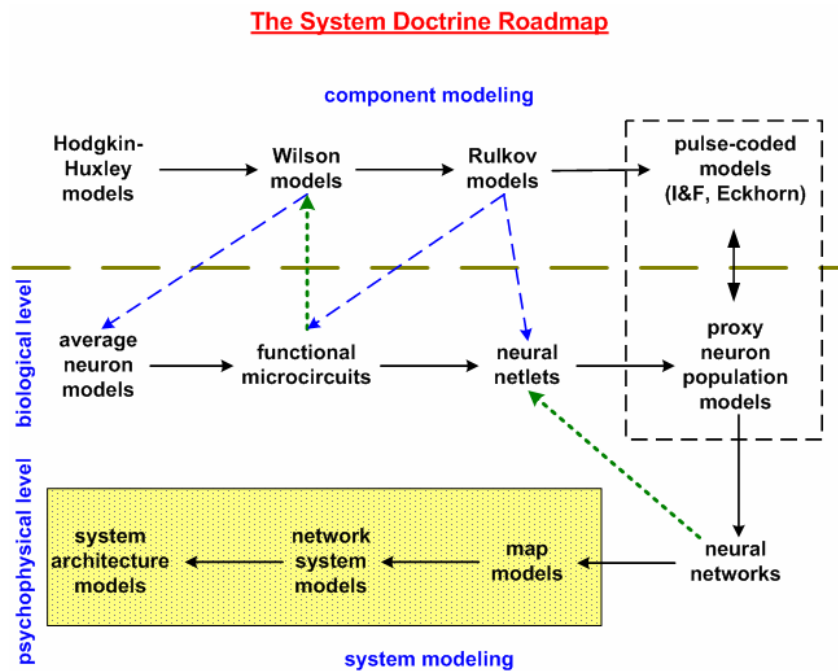
Was this mere coincidence or perhaps nothing more than another illustration of the power of mathematics to describe anything that *has* a formal description? Or is there something more to it than this? Philosophers and scientists alike have marveled for centuries over the ability of mathematics – which is clearly an invention of the human mind – to quantitatively describe nature – which is usually presumed to not be an invention of the human mind. Coincidence it may be, but consider Piaget’s finding that the copy of reality hypothesis is a provably false hypothesis. In this case, all the objects one comes to know must be *constructed* (since they cannot be the result of making a copy). The intriguing possibility this raises is that perhaps, since children’s knowledge is built up on the basis of Piaget’s practical sensorimotor transformational structures, the concordance of mathematics and the experimental sciences is due to one and the same root cause – namely, the intrinsic nature of how human beings come to know their world and everything in it. The concordance between mathematics and experimental science would then not be the product of mere accident; rather, it would be because mathematical knowledge and empirical knowledge are both weaved from the very same psychological cloth.

If this is true, then another implication is presented. It is a fundamental tenet of neuroscience that all behavior in animals – including man – is the product of brain structure and brain function. If the structure of human behavior is found to be isomorphic to the fundamental structures of mathematics – algebraic structure, topological structure, and order structure – then perhaps the structure and function of neural organization is also none other than Bourbaki mother structure. If this corollary implication is true, then perhaps the secret to unlocking the neural code lies not with searching for the brain’s “language” but, rather, by looking into the development of neural and neural-glial *structure*. The research question for computational neuroscience then becomes: ***What is functionally necessary if it is to be possible for the brain to develop algebraic, topological, and order structures in its neural network systems?*** Investigation of this question is the subject of this branch of research here at LCNTR. This is how we shall define the problem of the neural code.

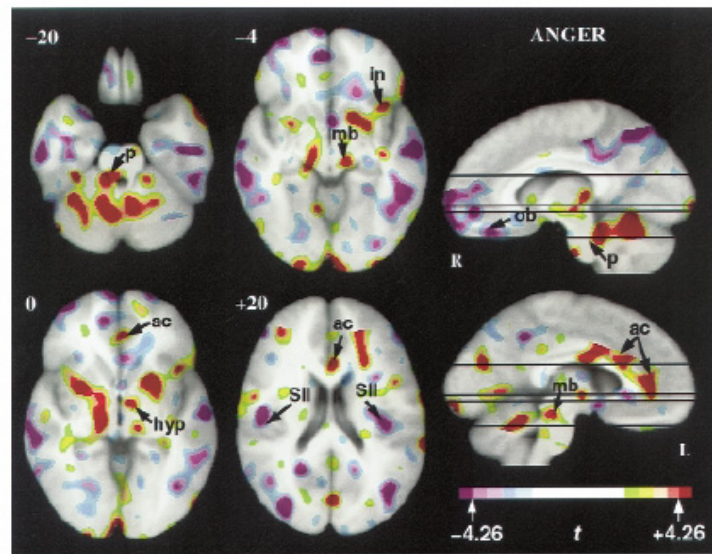
## II. Psychophysical Neural Network Modeling

It is not Piaget’s contention that sensorimotor correspondents to the mother structures exist full-grown and innate at the time of a child’s birth. Quite the opposite is the case. The newborn infant comes into the world with a set of basic and genetically-determined sensorimotor reflexes that do not yet rise to the level of, say, a full-blown algebraic structure. Although it is true that some innate capabilities – namely those relating to the child’s earliest perceptual representations of space – already show fundamental elements of topological representation (Piaget and Inhelder, 1967), the great majority of these structural transformations acting at the sensorimotor level must be built up during the stage of sensorimotor intelligence from birth to approximately age 2 years (Piaget, 1952, 1954). The extension of the mother structures to habits of thinking and reasoning follows their sensorimotor development over a period of years from age 2 years to approximately age 15 years (Inhelder and Piaget, 1964), (Inhelder et al., 1974).

Under the fundamental tenet of neuroscience, the neurological substrate for these behavioral abilities is modeled by mathematical neural network models. Such models quantitatively represent large numbers of neural networks and are functional in nature. To appreciate the physical scale of these models, it is helpful to consider the “system roadmap” shown in figure 1. Research in theoretical neuroscience can be broadly divided along the lines of component modeling, which runs from molecular and cell-level models to models of functional properties of



**Figure 1:** System roadmap depicting the span of levels of scientific reduction studied in computational neuroscience. The topic of this tech brief belongs to the system architecture model level but involves all three levels depicted in yellow background in this figure.

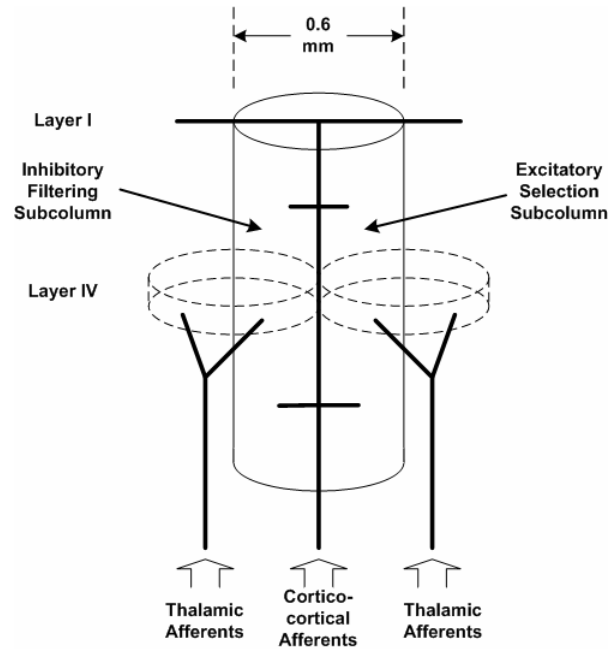


**Figure 2:** PET scans showing areas of significant activation (deep red) and significant inhibition (deep purple) in human subject during the experiencing of anger (Damasio et al., 2000).

interconnected groups of hundreds of neurons, and system modeling, which runs from statistical models of neuron populations close to the biological level to models of major brain systems with functions closely related to the psychophysical level of experimental research. The constituent elements of system architecture models, which is the topic of this tech brief, are models that take in the functional properties of assemblies of hundreds of thousands to millions of neurons.

Because system architecture models and the lower modeling levels from which they are constructed operate hand-in-glove with experimental psychology, it is important to appreciate how neuroscience makes the connection between brain and the objects of psychology. The method is based on drawing functional correspondences between brain structure actions and basic psychological entities such as emotion, perception, consciousness, and so on. Today this is done by using measurement tools such as positron emission tomography (PET) scans and functional magnetic resonance imaging (fMRI). Figure 2 illustrates a PET scan taken for a subject during the experiencing of anger. Different color levels in the scan correlate to levels of metabolic activity in the cells in various regions of the brain. Metabolic levels are known to be higher for neurons that are actively signaling than for neurons in the quiescent state. Thus, PET scans provide data on the *activity level* of neuronal signaling, and computational models at the map, network system, and system architecture modeling levels in figure 1 represent this activity level in their signal variables.

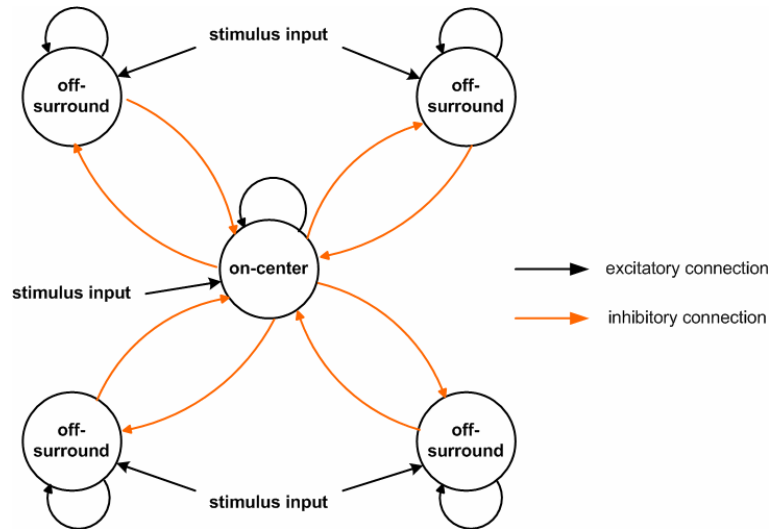
The organization of the various modules found in system architecture models is deduced from what is known of brain anatomy and intercellular connectivity. The mathematical descriptions for



**Figure 3:** Functional column structure proposed by Szentágothai.

these models must maintain a close connection with biological facts. The computational neuroscientist is not free to engage in arbitrary speculation concerning the structure of his or her models. The lowest-level constituent of a network architecture model is called a *map* model. A map node is an abstract representation of neuron cell assemblies comprised of thousands of neurons anatomically organized so as to form a *functional unit*. Biologists give such cell assemblies names like *functional columns*, *blobs*, *barrels*, and *nuclei*. Figure 3 illustrates a qualitative model of a functional column in the neocortex (Szentágothai, 1983). In human neocortex, one such functional column contains on the order of 17,000 neurons. Functional columns are closely interconnected to neighboring columns in the cortex and also connect to more distant columns and sub-cortical structures via long-distance projections through the white matter of the brain (Wells, 2005).

Map models are interconnected to make *network system models*. These models are functional representations of how map-level neuron assemblies affect one another due to the signals that pass between them. Network system models are deduced from gross anatomical properties of the synaptic interconnections among functional units. A variety of network system modeling theories have been proposed over the years by different theorists. The different modeling theories have differing degrees to which their “anatomy” and dynamical descriptions (“physiology”) are based on actual biological structure. The modeling theory that has most closely remained faithful to psychological *and* biological fact over the years was discovered and developed by Grossberg and goes by the name *adaptive resonance theory* or ART.



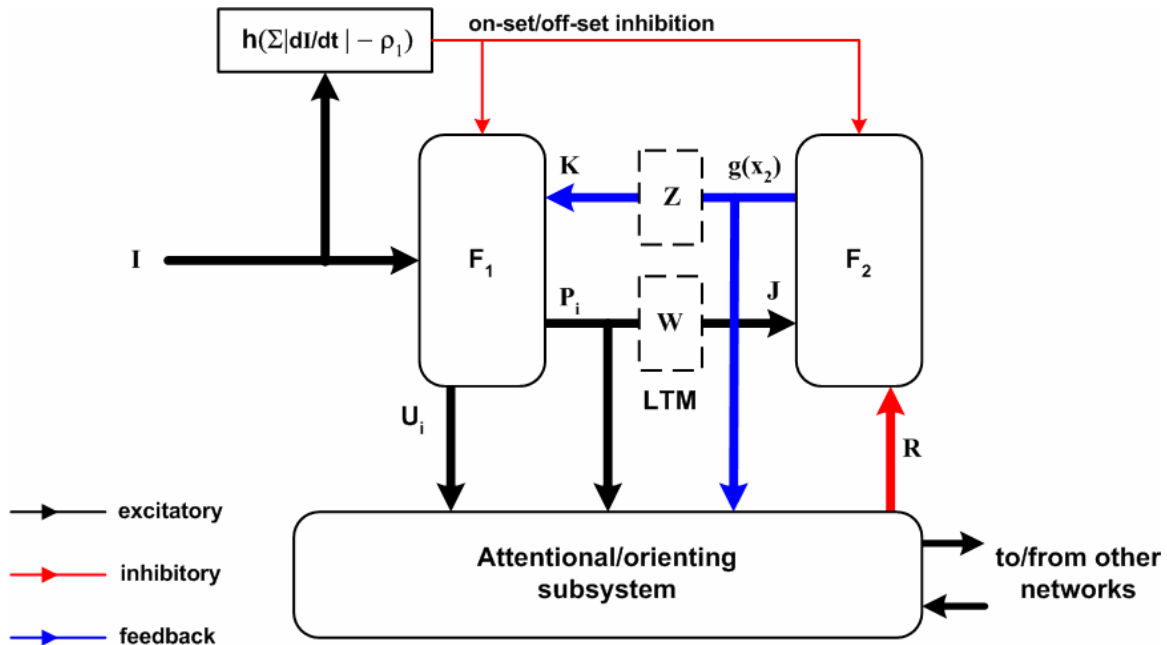
**Figure 4:** The on-center/off-surround anatomy. Each node in this diagram is a map model. Each map has excitatory self-feedback and projects inhibitory signals to the neighboring maps. The designation of a map as an on-center or an off-center map is relative, i.e. every map is “on-center” relative to itself and “off-center” relative to other maps.

ART networks are based on what is called an *adaptive resonator* network system model. Adaptive resonators, in turn, are based on a frequently-occurring biological structure known as the on-center/off-surround (OCOS) anatomy. Figure 4 illustrates the general concept of an OCOS anatomy. See the figure caption for the explanation of this structure. The mathematical expression for an adaptive resonator consists of a set of coupled, nonlinear first order differential equations. An ART resonator is so called because this system of dynamical equations undergoes a natural response leading to steady-state solutions where the pattern of state variables of each map in the network are said to form a *short term memory* (STM) pattern representing a “resonant state.” The ART resonator anatomy consists of two “fields” of interconnected maps, one of which,  $F_1$ , receives input signals (afferents) into the network system and the other of which,  $F_2$ , “classifies” or “categorizes” the afferent input pattern.

Figure 5 illustrates the basic layout of a typical ART network (Carpenter and Grossberg, 1987). In addition to the adaptive resonator (comprised of fields  $F_1$  and  $F_2$ ), the network system also contains additional map elements that comprise what is called the attentional/orienting subsystem of the model. As you might guess from this terminology, adaptive resonance theory grew out of neural network research originally aimed at understanding various psychological phenomena concerning classical psychological conditioning, perception, and behavior (Grossberg 1967, 1968, 1969a-c, 1970a-b, 1972a-b, 1973, 1975, 1976a-c), (Ellias and Grossberg, 1975), (Levine and Grossberg, 1976). A tutorial introduction to ART networks is given in (Wells, 2007, chapters 14-17).

ART networks solved, in part, one of the vexing problems in neural network theory. This was





**Figure 5:** A typical ART network structure. This particular structure is called an ART 2 network. The resonator is formed by the interconnection of the two network “fields”  $F_1$  and  $F_2$ . These fields are connected to one another through two matrices of connection weights,  $\mathbf{W}$  and  $\mathbf{Z}$ . The elements of  $\mathbf{W}$  are called the bottom-up weights; those of  $\mathbf{Z}$  are called the top-down weights. The two taken together are called the long-term memory (LTM) of the network. Vector signal  $\mathbf{K}$  is called the top-down expectation signal. Vector  $\mathbf{I}$  represents the afferent inputs to the network. The attentional/orienting subsystem is a third network subsystem that exerts a control function over the activities of the adaptive resonator. Adaptation in this network model means changing the values of the  $\mathbf{W}$  and  $\mathbf{Z}$  matrices in response to the signals and the state variables in the mathematical description of the network.

the problem of how to model neural network “learning” (in the mathematical rather than the psychological sense of that word) in a manner that was consistent with both psychological and neurological reality. The term “learning” when used in this sense denotes the adaptive determination of connection weight values between the various nodes in a network system model. Roughly speaking, the connection weights in the mathematical description of neural network systems mimic the biological phenomenon of cell-to-cell transmission of information at neuronal synapses. (The word “synapse” derives from a Greek word meaning “connection”). It has been known since the early 1970s that in many, many synapses the strength of connection between neurons is not constant. Rather, it undergoes activity-dependent changes leading to either greater or lesser conveyance of physical signals from one neuron to another. Tutorial overviews of this phenomenon are given in (Wells, 2003a-b) and in (Wells, 2007, chapter 11).

Prior to ART, neural network models employed various schemes for automatically adjusting the network weights. Grossberg was the first to point out that all these schemes suffered from two major flaws. First, these schemes employed neurologically unrealistic means for effecting the weight changes. Second, all these schemes suffered from a peculiar form of learning instability in which the learning of new weight connections by the network inevitably led to the “forgetting” of

previously learned recognition codes (Grossberg, 1987). ART was developed for the purpose of solving these problems in a biologically *and* psychologically realistic way.

### III. The Context Issue in ART Network Adaptation

Although ART and ART networks succeeded marvelously in meeting the research objectives pursued by Grossberg, there remain three troublesome and interrelated issues with the ART concept as it stands today. The first is what we might call the “open-mindedness” of ART learning. ART networks are in large measure classifier or categorization networks. An input pattern  $\mathbf{I}$  that is matched sufficiently closely by one of the LTM vectors  $W_j$  in the  $\mathbf{W}$  matrix will cause the activation of a corresponding node  $x_j$  in the  $F_2$  field and cause the deactivation of all other  $F_2$  nodes (this is called a 0-1 distribution in ART terminology). If no currently-established  $W_j$  vector provides a sufficiently close match to  $\mathbf{I}$ , the attentional/orienting subsystem will cause the network to commit one of its uncommitted  $W_j$  vectors to “learning” pattern  $\mathbf{I}$ . (Likewise, the corresponding uncommitted  $Z_j$  vector in the  $\mathbf{Z}$  matrix will also “learn”  $\mathbf{I}$ ). On the other hand, if  $\mathbf{I}$  is a sufficiently close match to some  $W_j$  vector, the network will adapt that vector in such a way that  $W_j$  approaches a statistical expected value representative of the features of the set of input vectors  $\{\mathbf{I}\}$  that satisfy the matching criterion (Carpenter and Grossberg, 1987), (Wells, 2007, chapter 17). Whether or not there is a “sufficiently close” match is determined by a parameter in the attentional/orienting subsystem called the network’s *vigilance parameter*. The only constraints imposed on the induction of the adaptation of  $\mathbf{W}$  and  $\mathbf{Z}$  are: (1) there must actually be an active input  $\mathbf{I}$  presented to the network; (2) the network itself must be in adaptive resonance; and (3) once a node  $x_j$  in the  $F_2$  field has been selected and adaptation begins, the network cannot switch to some other  $F_2$  node during adaptation. These constraints are built into the network and when they are all satisfied, the network will “learn”  $\mathbf{I}$  regardless of any other factor.

Now, this is precisely what an ART network is supposed to do. Considering only the network by itself, there is nothing wrong with any of this and a great deal about it that has demonstrated a close correspondence with psychological and psychophysical experimental findings. The “open-mindedness” of an ART network’s learning process only becomes an issue when it is joined to the second issue, which is the issue of *limited categorization capacity*. There is only a finite number of nodes in the  $F_2$  layer, and once they are all committed the network’s learning capacity is exhausted. The basic operation of an ART network is such that it will quickly use up this capacity as successive  $\mathbf{I}$  patterns are applied. The network does not discriminate, which is to say that it “neither knows nor cares” what any particular  $\mathbf{I}$  input might be; if the pattern is presented the network will learn it (if it still has uncommitted  $F_2$  nodes) and, after its capacity is exhausted

it will merely reject all subsequent mismatching patterns by producing resets in all the mismatched  $F_2$  nodes (resulting in an all-zeros  $F_2$  output). Put another way, the network does not attempt to “economize” its commitment of  $F_2$  nodes; it is like a novice poker player who stays in every hand all the way to the showdown and soon loses all his money. The only way to avoid this outcome and achieve some functional objective for the network is to *train* the network, either by selecting beforehand what patterns it is to learn and presenting only them until all  $F_2$  nodes are committed, or by augmenting the network with a “teacher” network. In the latter case, this multi-network network (the “learner” network plus the “teacher” network) is called an ARTMAP (Carpenter et al., 1991).

Resorting to either of these alternatives illustrates the third issue, namely that the ART network lacks what we might call a “sense of context.” If one pre-selects and presents patterns to be learned, the context is being supplied by external agency; if one augments the ART network with a “teacher network” then this is equivalent to building in what the rationalist philosophers of the 18th century called “innate ideas.” Neither recourse is acceptable as a brain model; neither recourse is consistent with psychological findings. Taking these three issues together, we arrive at the statement of the theoretical problem we wish to study: *how does one build “context” into a neural network system architecture without introducing unrealistic and ad hoc mechanisms?*

This problem is not bound up with any kind of shortcoming related to adaptive resonance theory itself. Grossberg (1975) very clearly pointed out that the basic ART network provides a functional rather than mechanistic description, and that the various mathematical properties of the network implicated – but did not claim to mechanistically model – the functions of numerous anatomical structures in the brain. He referred to the ART network’s functional mathematical descriptions as “mock” structures, e.g. mock-neocortex, mock-hypothalamus, mock-septum, and so on. It is worth keeping in mind that before specific working ART networks were first demonstrated, no one knew how to successfully model *any* of the psychological or psychophysical phenomena ART was developed to address. The ART network is a classifier and a very good one at that. Choosing *what* to classify is not part of its “job description.” Seen in this way, the problem stated above might be called a “meta-problem” in the sense that it is part and parcel of the larger context of network system architectures found at the psychophysical edge of the roadmap in figure 1. This brings us back around to the logic of meanings discussion in section I.

#### **IV. The Cortical-Thalamic Hypothesis of Meanings Context**

Viewed *very* abstractly, a classifier network meets the definition of an algebraic structure known as a *groupoid*. Formally, a groupoid is any mathematical structure consisting of a set of

inputs and an  $m$ -ary operation on this set possessing the property of *closure*. It takes quite a bit of abstract thinking to see the network of figure 5 in these terms; one must assign to every vector  $\mathbf{I}$  a unique subset of abstract symbols (such that each  $\mathbf{I}$  is regarded as consisting of  $m$  symbols drawn from a set  $S$ ) and likewise assign each possible categorization vector  $\mathbf{x}_2$  to *one* symbol in  $S$ . (In ART 2, these sets are finite). But once we have done so and the adaptation process has committed all the  $F_2$  nodes as categories, the network is an instantiation of a groupoid.

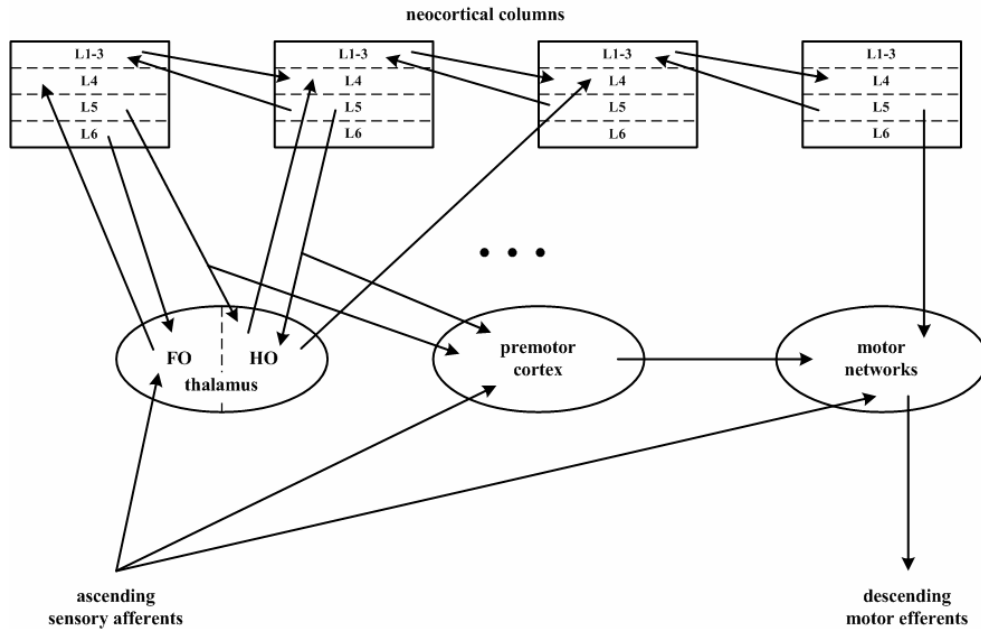
Piaget's theory holds that some groupoids are innate in the sense of already being present at birth (hereditary sensorimotor reflexes) and others develop soon after birth (first acquired habits). The theory further holds that the development of intelligence in the infant and child involves the successive structuring of ever-more complex structures, many of which correspond to the mathematical notions of *semigroups* (a set plus a closed, associative operation), *monoids* (a semigroup plus the existence of an identity element), and *groups* (a monoid plus the existence of inverses). Additional *algebraic structures* (a set with two or more operations on the set) likewise are constructed from these gradually and over time. Concurrently, topological structures and order structures are also found to develop. Piaget was able to present some concrete examples that illustrate these processes but, since he was not a mathematician, he did not develop a more general theory of this. This is not too surprising since the formulation of a general theoretical framework turns out to be rather complicated, as was shown and explained in (Wells, 2006). We will not attempt to go after a completely general theory in this tech brief; rather, we will focus our efforts on those aspects of the problem inherent in the research question stated in the previous section.

When a mathematician formulates a mathematical structure, he or she does so with some context for that structure already in mind. An example of such a context is "finding a more general concept of 'arithmetic'." Abstract algebra, point-set topology, and lattice theory are all examples of this sort of generalization mathematicians pursue. Within the context of theoretical neuroscience, our interest at present is in finding a proper way to express what mathematicians do in terms of large-scale brain structures and brain functions in a way that remains objectively valid in relationship to neurological and psychological facts. Even more specifically, what we wish to explore is the more rigorous explanation of the still-somewhat-vaguely-worded Piagetian notion of a logic of meanings underpinning the development of human intelligence.

Piaget's finding that the logic of meanings is indissolvably linked to practical sensorimotor actions at the most primitive levels of intelligence was shown by Wells (2006) to be fully congruent with the general theory of the phenomenon of mind. We will therefore take as our starting point for this work the hypothesis that context in ART network adaptation control is to be

sought in practical sensorimotor functions, and we seek for a proper network system architecture for expressing the instantiation of context derived from a neurological foundation for a logic of meanings. Within this objective, the previous research question can be sharpened up a bit by asking: *What manner of attentional control is needed to restrict ART network learning to only those signals within the system that bear upon the sensorimotor logic system for which the ART network constitutes a function subsystem?* In researching this question, we will explore the consequences of the following hypothesis: **(H1) The interaction of neocortex and thalamus is a neurological substrate for the logic of meanings.**

*A. The Thalamus and its Motor Logic Role.* In mammals all sensory information reaching the neocortex, with the single exception of olfactory sensory information, comes to the cortex by way of the thalamus.<sup>1</sup> This has been known for some 70 years and gave rise to a traditional view of the thalamus as a kind of relay network or gateway for sensory data. Something that was never clear, but always a puzzle, about the traditional model is: *Why?* What functional purpose is served by interposing a major subcortical brain structure – the thalamus – in between the sensory nerves and the cortex? Why not route sensory information directly to the early sensory cortices? This is what occurs, more or less, with olfactory signals. Why did evolution favor the development of the thalamus in vertebrates, and why should primates, including *Homo sapiens*, have the largest and most complex thalamus found among mammals?



**Figure 6:** System-level schematic of the sensorimotor network environment of the thalamus.

<sup>1</sup> For a tutorial overview of brain structure see (Wells, 2005).

The early research, from which the traditional picture of the thalamus resulted, was carried out by dissection, post mortems, and through research on anesthetized animals. In recent years, one key new fact has gradually emerged: Most, and maybe all, sensory signals going into the thalamus are also projected to neural motor circuits. Furthermore, it has been found that most, perhaps all, feedback signals returning to the thalamus from the neocortex *also* project to the central nervous system's motor pathways. This is illustrated schematically in figure 6. Recently some leading experts in thalamic research have been led to propose a new paradigm for the role of the thalamus. The new model holds that the thalamus informs the cortex not only about sensory information but also about the *motor state* of the organism. Furthermore, since the feedback pathways from cortex to higher-order (HO) nuclei in the thalamus also appear to contain motor information, the thalamus can be seen as passing on *motor command* information to other regions within the neocortex. This is illustrated in figure 6. Thus, the thalamus is to be seen not merely as a sensory portal but, rather, as a structure having a *sensorimotor* role (Sherman and Guillery, 2006, pp. 357-389).

This new view of the thalamus is, of course, interesting in its own right. But when put in context with the developmental psychology finding that all early meanings the infant and child assigns to entities and events is a practical sensorimotor meaning, the new thalamic paradigm would seem to assume an even larger degree of significance for brain research. Sherman and Guillery comment,

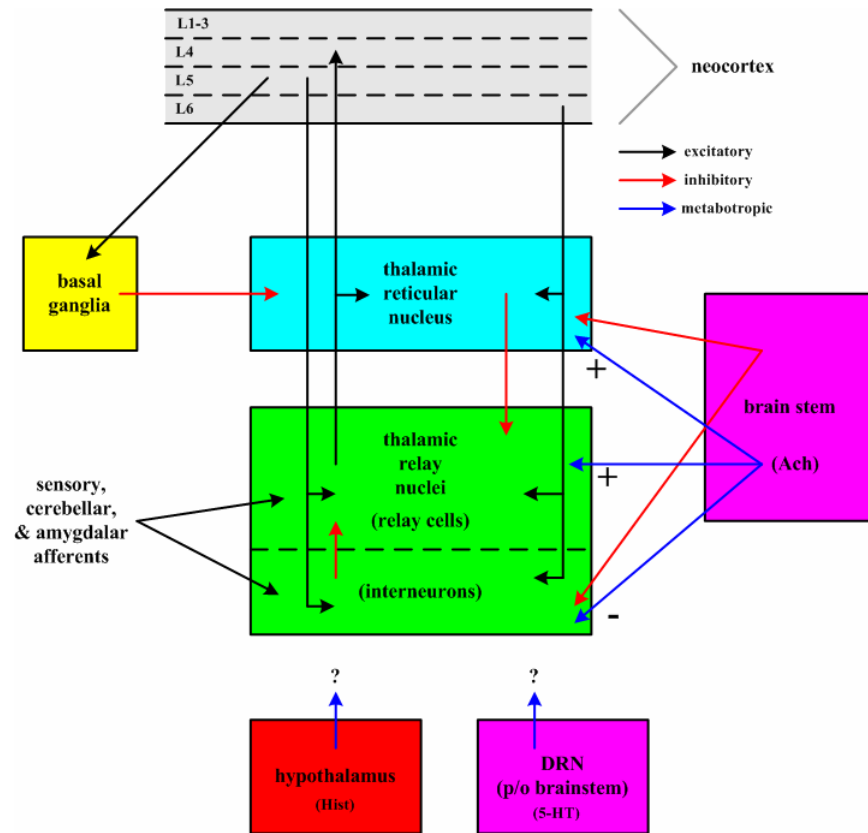
These pathways demonstrate that even primary *sensory* areas such as area 17 (V1)<sup>2</sup> have significant motor outputs that give these areas access to motor controls independent of any *higher* cortical processing.

We have mentioned . . . that several of the axons that bring inputs to first and higher order thalamic nuclei have branches that innervate cell groups with connections to motor centers . . . but so far we have not explored the implications of these connections. In addition . . . most, possibly all, cortical areas have connections with motor or premotor areas. These pathways provide evidence that not only *can* “motor assembly begin before sensory signals reach the highest levels” but that it *must* begin before the sensory signals even reach the thalamus, and that it must accompany corticocortical processing at essentially every stage. The messages that pass along the axons to first order relays, and that also pass along branches of the same axons to motor centers, provide a close and essentially unbreakable link between action and perception at the earliest stages of sensory processing, and those that pass along branching axons to higher order relays provide essentially secure links between cortical outputs to motor centers and perceptual processing through corticocortical connections (Sherman and Guillery, 2006, pp. 361-362).

Thus, the recent neurological evidence supports Piaget's contention that there is an essential link between actions and perceptions. It may be that we now possess a key clue to understanding the neurological substrate for the logic of meanings.

---

<sup>2</sup> V1 is part of the primary visual cortex.



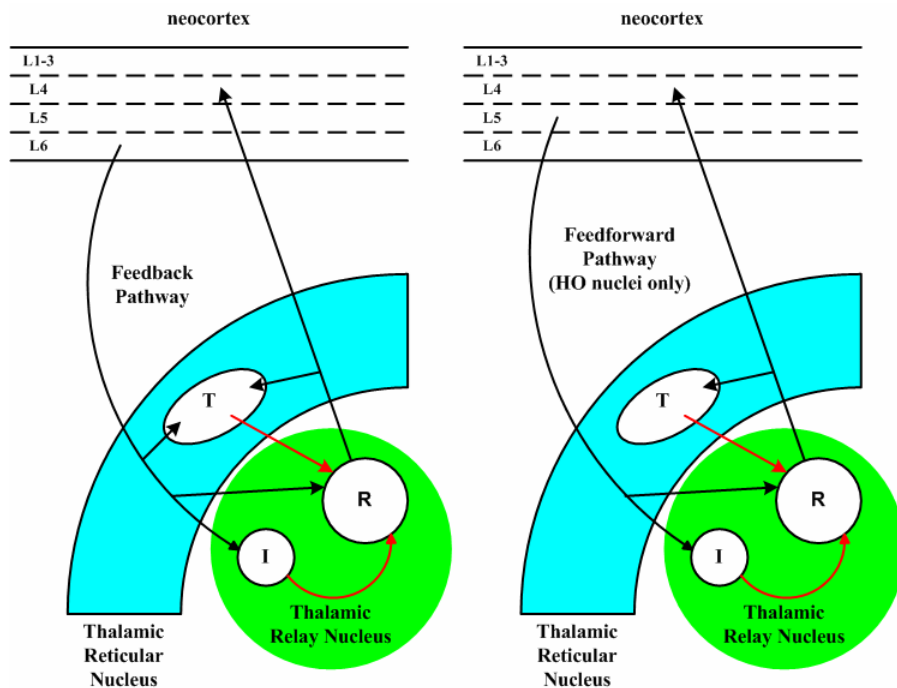
**Figure 7:** The anatomical environment of the thalamus. The thalamus is divided into three distinct regions known as the dorsal thalamus, the ventral thalamus, and the epithalamus. The thalamic reticular nucleus belongs to the ventral thalamus, while the thalamic relay nuclei belong to the dorsal thalamus. The epithalamus is most closely associated with the hypothalamus and is not represented in this figure. In addition to sensory afferents coming up via the spinal cord from the peripheral nervous system, the thalamus also receives signals from the cerebellum (which is involved with motor control), the amygdala (which is involved with affective phenomena such as emotions), brain stem (which is implicated in consciousness and in arousal, waking, and sleeping), the hypothalamus, and the basal ganglia (also known to be involved with the motor functions). The precise connectivity of the hypothalamic inputs and those of the DRN are not yet understood, which is indicated by the question marks in the schematic. DRN = dorsal raphé nucleus.

Hist = histamine pathway. 5-HT = serotonin pathway. Ach = acetylcholine pathway. Hist, 5-HT, and Ach are metabotropic modulator signals, which means they alter how their target neuron cells react to stimuli. + denotes that the metabotropic pathway is excitatory, while - denotes that the metabotropic pathway is inhibitory. L denotes a layer in the neocortex.

B. The Anatomical Environment of the Thalamus. Figure 7 is a schematic illustration of the anatomical environment for thalamic signaling pathways. The dorsal thalamus (see figure caption) makes up the largest part of the thalamus in primates, and projects to topographically organized regions of the neocortex (the brain’s “gray matter”). It is composed of more than 30 distinct relay nuclei (RN) specialized for passing information from particular sensory and other modalities. All the axons carrying these signals must pass through the thalamic reticular nucleus (part of the ventral thalamus). The TRN itself makes no projections to neocortex but rather sends inhibitory feedback signals to the relay nuclei in the dorsal thalamus. Both the RN and TRN are subject to modulations due to signals from other brain structures, including brain stem and the hypothalamus. Ascending afferents converge on what are called *first order* (FO) nuclei, which

also receive feedback signals originating in layer 6 of the neocortex. In addition, the thalamus receives signals originating from layer 5 of the neocortex. These converge on **higher order** (HO) nuclei, so called because the information passed back from the thalamus to neocortex due to these sources contain information “the cortex has already seen” – whereas the FO nuclei convey information “the cortex has not previously seen.”

Figure 8 provides a more detailed schematic representation of the general signaling schema of the thalamus and the differences between FO and HO relay nuclei. One should carefully compare this figure with figure 6 above. The existence of HO relay nuclei is also part of the new paradigm for understanding the thalamus. In the older traditional model of perception processing in the neocortex, it was thought that all such processing took place exclusively by direct corticocortical signaling. Under the new paradigm, direct corticocortical projections are still made (and, indeed, they constitute the majority of signaling pathways within neocortex), but these are augmented by signals from HO relay nuclei which, as stated above, contain motor command information.



**Figure 8:** A more detailed schematic of the signal processing difference between FO and HO nuclei. All RN are thought to conform to the left-hand diagram, i.e. all thalamic relay nuclei receive feedback from layer 6 of the neocortex and it is thought that this feedback signal originates from the same region of the cortex as the ascending thalamic signal projects to. Both the ascending and descending signals project to cells (T) in the TRN that send inhibitory signals back to the associated relay nucleus. FO relay nuclei receive ascending signals from subcortical sources, including spinal cord, that have not yet reached the neocortex. HO nuclei, on the other hand, do not receive these lower ascending signals but do receive a feedforward signal originating in layer 5 of the neocortex as illustrated by the right-hand figure. This descending signal does not project to T cells in the TRN. The HO nuclei in turn project back to the neocortex, but to a different region than that from which the descending layer 5 signal originated. Black lines denote excitatory signals, red lines denote inhibitory signals. R = relay cells. I = interneurons. T = TRN cells. L denotes the layer in the neocortex.

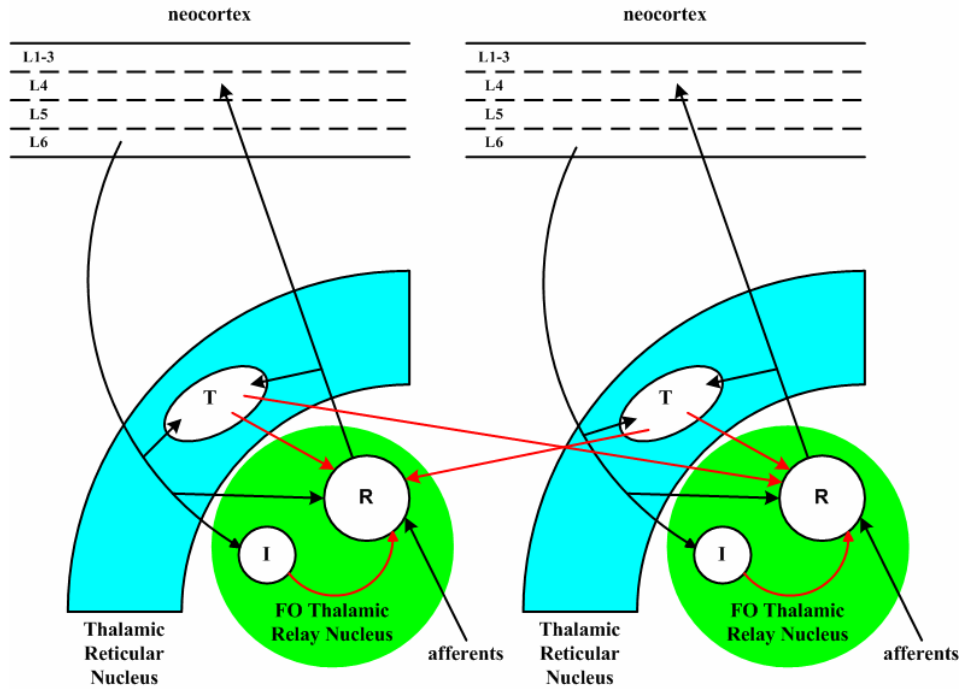


Yet another indirect pathway from neocortex to thalamus to neocortex takes place by way of the basal ganglia. Unlike the pathway through the HO relay nuclei, this pathway is essentially modulatory in its nature. The basal ganglia projections to thalamus are projected to the TRN and its T cells, primarily in the ventrolateral, ventroanterior, and meriodorsal sections of the ventral thalamus. It is not very clear precisely what the function of these projections might be, but because they inhibit T cells the implication is that they remove the self-regulatory negative feedback between the TRN and the RN, which would tend to allow more vigorous signaling activity within the RN. Projections to neocortex from the RN implicated in this pathway are made principally to the pre-motor and supplementary motor cortices and to the prefrontal cortex.

Within the basal ganglia there are a number of feedback loops in this pathway, the function of which is again not yet well understood. However, it is generally thought that some form of signal filtering is effected via this pathway. To make matters even more interesting, it is known that the thalamus' midline nuclei, central medial nucleus, and intralaminar nuclei (all dorsal relay nuclei) make projections into the basal ganglia in addition to the diffuse projections they make to the neocortex. The relay nuclei receive afferent inputs from spinal cord, cerebellum and also from the basal ganglia. Whatever the precise function or functions of these pathways may be, it seems inherently clear that it involves motor signal processing because the basal ganglia and cerebellum are known to be involved in this, as are some of the ascending signals from spinal cord. However, the thalamic projections of this pathway back into neocortex are diffuse, i.e. do not readily seem to conform to a particular topographical mapping the way the major sensory modality afferents do. It is readily conceivable that this pathway is not confined to the corticocortical signaling regions within neocortex depicted in figure 6.

Although different relay nuclei within the dorsal thalamus do not appear to make any direct projections to one another (that is, RN signals do not cross dorsal nuclei boundaries within the thalamus), some recent experimental evidence does hint that there may be *indirect* projections between relay nuclei by way of the TRN. This is rather highly speculative at this time, but if it should turn out to be true then yet another dimension would be added to the new thalamus paradigm. One speculative example of this is illustrated schematically in figure 9 (Sherman and Guillery, 2006). This network presents a scenario for possible control of activity arousal in neocortical functional columns (*ibid.* pp. 239-242).

To understand the idea behind this network, we must first review a few facts concerning neural signaling in the thalamus. Thalamocortical relay neurons exhibit two distinct firing modes called **tonic** firing mode and **bursting** firing mode. Tonic firing by the relay neuron is thought to be more or less a "relay mode" by which information presented to the thalamus is passed along to



**Figure 9:** Schematic illustration of possible cross-coupling between first order relay nuclei. This network connection is merely speculative at the present time, but is one possible guess for understanding how attentional focus might be implemented within the neocortex.

neocortex more or less intact and without distortion. It is thought that tonic mode firing from the thalamus is effective in activating the target column in neocortex to respond to stimuli. Bursting mode, on the other hand, introduces severe nonlinear distortion of the ascending signal. There is an initial burst of activity upon initial application of a *novel* stimulus followed by a long interval in which the relay cell goes silent. It has been suggested that cells operating in burst mode may be more capable of detecting a novel stimulus or of detecting small variations in the general features presented within the ascending stimulus. In burst mode the relay cells are ineffective at detecting low frequency stimuli ( $< 1$  Hz), respond best to intermediate frequencies ( $\sim 4$  Hz), and respond poorly or not at all to higher frequencies ( $> 10$  Hz). In contrast, relay cells operating in tonic mode respond over this entire range of input stimulus frequencies.

Owing to these properties, it has been suggested that burst firing may serve as a “wake up call” from thalamus to neocortex. Because of the severe nonlinear distortion introduced into the signal pathway by burst mode firing, it is thought that burst signals going into the neocortical functional column are difficult for those circuits to “analyze.” Although the notion that a column in the neocortex “analyzes” signals is more poetic than scientific, in the context of an ART network we can liken this notion to the more concrete statement that burst signals probably do not present signaling conditions required for achieving a state of adaptive resonance. In ART no network activity is regarded as “significant” except during a state of adaptive resonance.

Returning now to the model in figure 9, let us assume that the left-hand relay nucleus,  $RN_1$ , is operating in burst mode while the right-hand nucleus,  $RN_2$ , is operating in tonic mode. Then  $RN_2$  is responding to on-going stimuli and passing this information along to its related neocortical column,  $FC_2$ . (Using an ART network model for the latter, we can say  $FC_2$  is thereby placed in a state of adaptive resonance).  $RN_1$ , on the other hand, is not providing its functional column,  $FC_1$ , with signaling conditions required to establish adaptive resonance and so we regard  $FC_1$  as having low responsiveness to the incoming thalamocortical signal. In particular, in the absence of a *novel* stimulus presented to  $RN_1$ , the burst mode state leads to no significant signaling being conveyed to  $FC_1$ .

Now let us suppose  $RN_1$  receives a novel stimulus. It then responds with a burst signal sent to  $FC_1$  that strongly drives the neocortical column into activity. This activity produces a feedback signal from layer 6 back to the T cell associated with  $RN_1$ , causing  $RN_1$  to switch to tonic firing mode. At the same time, this T cell also sends an inhibitory signal into  $RN_2$ , causing it to switch to the burst mode and reducing the inhibitory signal sent by the right-hand T cell to  $RN_1$ . (In computer engineering terminology, this action is called a “flip flop”). The net result is that  $FC_2$  drops out of adaptive resonance and enters a condition of low responsiveness, while  $FC_1$  enters a state of adaptive resonance and becomes highly responsive to its thalamocortical signal. In this way, one can say that neocortical network “attention” has “shifted” from  $FC_2$  to  $FC_1$  as a result of the novel stimulus presented to  $RN_1$ .

*C. Meanings Context.* If this speculative model is true, this would mean the thalamic network is functionally a part of the attentional subsystem depicted in figure 5. Furthermore, it implies that attentional subsystems for different ART networks are linked in agonist-antagonist relationships (which is, in part, what is depicted by the “to/from other networks” signal pathways in figure 5). Such relationships would make very little sense under the traditional model (how could different perception data be mutually antagonistic?), but it makes a great deal more sense in the new paradigm, where perception and motor action are intimately linked. This is because muscles (the engines of motor action) *are* arranged in agonist-antagonist pairs.

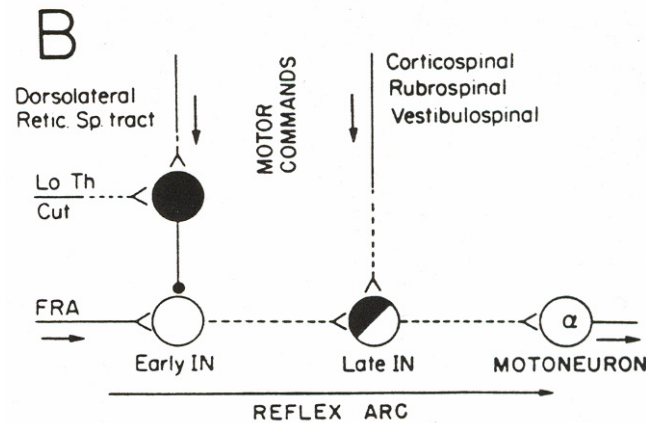
Although the model of figure 9 is, at this time, merely speculative, it does have a number of interesting similarities to the neural organization of the ventral horn of the spinal cord (Wells, 2003c). The feedback loop from R to T and back to R in figure 8 (which is *not* a matter of speculation) recalls to mind the regulating loop from alpha motor neuron to Renshaw cell and back to alpha motor neuron (Wells, 2003c, pt. III, figure 1). Going further, the cross-coupled inhibitory linkages in figure 9 call to mind the cross-coupled inhibition action mediated in the ventral horn via the Ia inhibitory interneuron (Wells, 2003c, pt. III, figures 3 and 4). The model of

figure 9 and the neural circuitry of the ventral horn do, of course, differ in detail but, overall, there are more similarities between the two than there are differences.

Under the new paradigm for the role of thalamus, FO nuclei are regarded as passing along not only sensory but also motor state information, while HO nuclei are regarded as passing along motor command information. Now, to whatever extent the motor state and motor command information constitutes an “image” or a “reflection” of what is going on in the organism’s motor hierarchy, it would not be all that surprising if it turned out that this information is much more closely related to functional aspects of the ventral horn than has been generally thought. This would be especially the case if the putative neural code subsists in *network organization* and *network state* rather than merely in the signals that pass from one neuronal structure to another (which can be regarded as mere vehicles for *effecting* particular network states and a dynamical linking of neuronal subsystems to organize specific functional transformations). However, before we become overly enthusiastic about what may be no more than merely superficial similarities between the model of figure 9 and the aforementioned motor neuron circuits, we must deal with an important conceptual issue.

The ventral horn networks mentioned above are part of the spinal cord network’s *reflex* circuitry. Descending motor commands coming down from the brain, on the other hand, are regarded as the neurophysiological instantiation of *voluntary* motion. How, then, could there be any objective validity to comparing spinal reflex circuits to high-level sensorimotor processes in the brain? A possible answer to this is seen when one considers the larger view of the spinal neuromuscular system for voluntary skeletal muscle control. Here, the current leading hypothesis for voluntary muscle control is the *reflex arc* concept. It is thought that voluntary muscle control by the brain is effected at the spinal cord level by *co-opting* the spinal cord’s reflex circuitry (Wells, 2003c, pt. I).

The general idea is illustrated in simplified form in figure 10. Spinal cord reflexes are stimulated by various sensory nerves found in muscles, tendons, joints, skin, and so forth. One important class of nerve signals in this system is called the *flexor reflex afferents* or FRA. Some authors, e.g. Lundberg et al. (1987), prefer the term *generalized reflex afferents* or GRA to describe the sum total of afferent signals stimulating the reflex pathways. Descending motor commands from the brain converge on interneurons, perhaps located at the output layer of the dorsal horn of the spinal cord and/or at higher-level layers within the ventral horn, which ultimately drive alpha motor neurons. At the same time, other descending signals stimulate inhibitory neurons in the GRA feedback pathways, effectively cutting off immediate nerve stimulus in the reflex path and, in effect, substituting their own mock-sensory signal in its place.



**Figure 10: Schematic illustration of the reflex arc concept.** The circles indicate small subnetworks of neurons at the same synaptic level. IN stands for interneuron. Input signals are actually vectors rather than individual signals. The dashed lines indicate other possible layers of interneurons interposed between the subnetworks shown. The motor commands are descending signals from the reticulospinal and other tracts (see figure 7). Excitatory synaptic connections are indicated by the “<” synapse symbols. Inhibitory synaptic connections are indicated by the “•” symbol on the lower-left subnetwork. FRA is “flexor reflex afferents”, and Lo Th Cut is “low threshold cutaneous inputs”.  $\alpha$ -motoneurons are the motor neurons that drive extrafusal muscle fibers (Wells, 2003d). “Early IN” designates interneurons that directly receive afferent inputs. “Late IN” designates interneurons at deeper layers of the network in the signal pathway. (These would perhaps consist of neurons in the output layer of the dorsal horn or perhaps interneurons in the ventral horn). The solid-black subnetwork at the upper left in the figure represents mid-level interneurons in the dorsal horn. Note that this subnetwork sends inhibitory inputs to the early dorsal horn INs.

According to this hypothesis, voluntary muscle control is “piggy-backed” onto the phylogenetically older reflex circuit arrangement and makes use of the “control circuit functions” already in place for basic spinal cord reflexes. In order for this hypothesis to work, however, it is clear that the brain must be able to supply the proper sort of mock-sensory signals for stimulating the reflex arc. Prior to the new thalamus paradigm, it was very, very unclear how this could be possible, and this was one theoretical argument to oppose the reflex arc concept.

Analogy, of course, is often an untrustworthy guide. However, analogy between the overall organization of the ventral horn (Wells, 2003c, pt. IV) and the new thalamocortical paradigm discussed in this tech brief is probably our best guide given the present-day state of knowledge concerning the function and role of thalamus in relationship to neocortical functional columns. This seems particularly so in light of the psychological findings regarding the logic of actions and the logic of meanings discussed earlier. We can take some additional comfort from the fact that these findings are in complete congruence with a larger general theory of mental phenomena (Wells, 2006) and the analogy does not contradict this larger theory.

Where, then, does this take us to at present? How are we to interpret and understand the phrase “meanings context” in regard to network system architectures and the idea of the neural code? Let us summarize the main points in this tech brief.

**1. *The thalamo-cortical system is a sensorimotor and not merely a sensory system.***

Information being processed and relayed by the thalamus to the neocortex simultaneously involves both data traditionally regarded as “perception data” and data reflecting motor state and motor command information.

**2. *Thalamocortical signaling occurs in two modes.*** One mode, the burst mode, is postulated to be an attention-getting mode for the target neocortical functional columns and inhibits network adaptation of LTM weights (in the ART network model) by preventing the attainment of a state of adaptive resonance in the ART network. The other mode, the tonic mode, provides accurately conveyed sensorimotor information to the neocortex and enables LTM adaptation by supplying the signal conditions necessary for achieving adaptive resonance in the ART network.

**3. *We make the hypothesis that the neural code subsists in the state and dynamically determined structuring of interacting neural network systems and not in the mere signaling that passes from one subsystem to another.*** Furthermore, we make the additional hypothesis that this dynamical neural *structuring* gradually takes the form of basic mathematical structures, i.e. algebraic structuring, topological structuring, and order structuring. This latter hypothesis is motivated by the findings of developmental psychology, which state that the most basic organization of sensorimotor schemes gradually comes to exhibit behaviors that are behavioral isomorphs of these basic mathematical structures.

**4. *We make the hypothesis that the dynamical structuring of the network system architecture constitutes a logic of meanings that rests on a primitive proto-logic of actions.*** Prior to cognizance of objective concepts of entities and events, knowledge structure is *practical* rather than theoretical (Piaget, 1976) and has for its neural substrate the formation and organization of sensori-motor action schemes. These *developed* schemes take their point of origination from innate reflex schemes and capabilities possessed by the human infant at birth.

## **V. The Next Step in the Research Program**

The hypotheses outlined above are quite speculative and they lie well outside the mainstream of current neuroscience thinking. Now, in science all hypotheses are rightfully regarded with the attitude denoted by the motto of the state of Missouri, i.e., “Show me.” We cannot regard them as having a well-established grounding unless we can also *demonstrate* that neural network system “learning” can be achieved that takes on the form of mathematical structures (algebraic, topological, and order structuring). The ability for a network system architecture to *spontaneously* form such structures has yet to be demonstrated.

Prior to recent developments in our understanding of the thalamo-cortical system,

neuroscience had no guide posts or road signs to indicate how such structuring might be expressed *in concreto* by neuronal structures. It is not enough to simply demonstrate a neural network system that functionally implements a groupoid, a semigroup, a monoid, or a group. This is no significant achievement. What *would* constitute a significant achievement would be to produce a network architecture model that achieves this outcome using *only* assumptions and models tied back to neurological and psychological *facts* emerging from neuroscience research.

A criticism often, and often justly, leveled at neural network research is that too many neural network models are *ad hoc* and are not the result of cumulative research (O'Reilly and Munakata, 2000). This is a criticism that cannot be leveled, by an informed person, against ART (which was and is the fruit of many years of cumulative research that maintained good faith and allegiance with the findings of both cognitive and biological neuroscience). But ART is not and does not claim to be a finished theory. ART networks, such as the one illustrated in figure 5, are a beginning, not an end point. The focus of most current research involving ART is aimed at understanding the implications of ART at the scale of the network system architecture. Much of great interest and probable significance has been accomplished in recent years by ART theorists. However, it has been quite some time since deep and *fundamental* considerations have played a well-published role in considering the network architecture problem. This tech brief has tried to help remedy this situation.

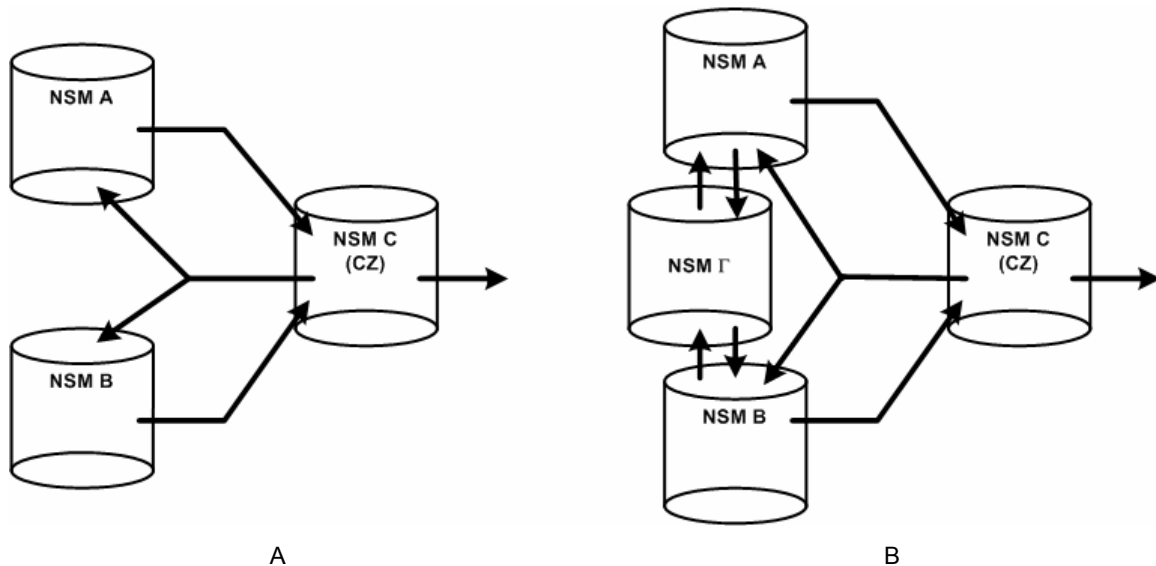
Because no one has yet demonstrated a general method or algorithm by which even the first step of the new paradigm – a spontaneous and unsupervised adaptation process taking ART networks from a simple classification function (a groupoid structure) to a semigroup function – this is the logical place to go next. To accomplish this, one needs to put together a simplified – but not overly simple – “body environment” in which signals have both sensory *and* motor context. Our work here at LCNTR, especially the spinoneuromuscular modeling work being carried out in the Soule Laboratory for Evolutionary Computing, is a good basis for this step.

Next, one must develop models for the thalamocortical subsystem and its functional contribution to the attentional/orienting subsystems of ART networks. Third, this must be integrated into a model such as that illustrated in figure 6. This integration should aim at discovering what constraints are placed upon network systems (figure 5) if these systems are to spontaneously form mathematical structures. These three steps constitute the first *basic research objectives* for this program.

From the viewpoint of theoretical difficulty, the third objective is likely to be the initially most challenging. Here something that deserves serious consideration is the system-level hypothesis made by Damasio in his convergence zone proposal (Damasio 1989a,b). Damasio's model is

based on neurological findings implicating the idea that networks in the early sensory cortices comprise “feature fragments” that are coordinated and linked together by retrograde feedback from immediately downstream networks called *convergence zones*. A very simplified illustration of this concept is given in figure 11. For purposes of initial conceptualizing, it is probably advantageous to regard individual ART networks as constituting *neural network state machines* (NSM). The likely advantage to this is that it places the function of each network in a context in which the well-developed theory of finite automata is cast, thus potentially simplifying the process of relating neural network systems to mathematical structures. It might also turn out to be the case that perhaps individual networks such as figure 5 are better regarded as symbol-representing automata rather than as actual groupoids in their own right. In such a case, the groupoid structure would appear at the level of the first-order CZ networks in figure 11, and higher algebraic structures would develop out of these structures. This representation question is at present open and unresolved.

The model of figure 11 developed out of earlier theoretical considerations made prior to the more recent development of the new paradigm for thalamus. Thus, figure 11 presents no representation of the contribution of the thalamus, although it would not be inconsistent to cast NSM  $\Gamma$  in figure 11(b) in this role. This, too, is a research question for our earliest work in this research program.



**Figure 11:** Simplified version of two forms for Damasio’s convergence zone (CZ) model. The cylinders each represent one ART network. Afferent inputs and thalamic interactions are not illustrated in this figure. Each ART network in this picture is regarded as constituting a neural network state machine (NSM). (A): CZ structure with no lateral interactions between feature fragment networks A and B. (B): CZ structure with lateral interactions between feature fragments A and B mediated through a third network  $\Gamma$ .



## Citations

- Carpenter, G. and S. Grossberg (1987), "ART2: Self-organization of stable category recognition codes for analog input patterns," *Applied Optics*, 26: 4919-4930.
- Carpenter, G., S. Grossberg, and J.H. Reynolds (1991), "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565-588.
- Damasio, Antonio R (1989a), "Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and cognition," *Cognition*, 33: 25-62.
- Damasio, Antonio R (1989b), "The brain binds entities and events by multiregional activation from convergence zones," *Neural Computat.*, vol. 1, pp. 123-132.
- Damasio, Antonio R (1994), *Descartes' Error*, N.Y.: Avon Books.
- Damasio, Antonio R (1999), *The Feeling of What Happens*, N.Y.: Harcourt Brace.
- Damasio, A.R., T.J. Grabowski, A. Bechara, H. Damasio, L.L.B. Ponto, J. Parvizi, and R.D. Hichwa (2000), "Subcortical and cortical brain activity during the feeling of self-generated emotions," *Nature Neurosci.*, 3 (10): 1049-1056.
- Ellias, SA and S Grossberg (1975), "Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks," *Biol. Cybernetics*, 20: 69-98.
- Grossberg, S (1967), "Nonlinear difference-differential equations in prediction and learning theory," *Proc. National Academy of Science*, vol. 58, pp. 1329-1334.
- Grossberg, S (1968), "Some physiological and biochemical consequences of psychological postulates," *Proc. Nat. Acad. Sci.*, 60: 758-765.
- Grossberg, S (1969a), "Some networks that can learn, remember, and reproduce any number of complicated space-time patterns, I," *J. Math. and Mechanics*, 19(1): 53-91.
- Grossberg, S (1969b), "On learning, information, lateral inhibition, and transmitters," *Math. Biosciences* 4: 255-310.
- Grossberg, S (1969c), "Embedding fields: A theory of learning with physiological implications," *J. Math. Psychology*, 6: 209-239.
- Grossberg, S (1970a), "Some networks that can learn, remember, and reproduce any number of complicated space-time patterns, II," *Studies in Applied Math.*, 49(2): 135-166.
- Grossberg, S (1970b), "Neural pattern discrimination," *J. Theor. Biol.*, 27: 291-337.
- Grossberg, S (1972a), "A neural theory of punishment and avoidance, I. Qualitative theory," *Math. Biosci.* 15, 39-67.
- Grossberg, S (1972b), "A neural theory of punishment and avoidance, II. Quantitative theory," *Math. Biosci.* 15, 253-285.
- Grossberg, S (1973), "Contour enhancement, short term memory, and constancies in reverberating neural networks," *Studies in Applied Math.*, 52(1): 213-257.
- Grossberg, S (1975), "A neural model of attention, reinforcement, and discrimination learning," *Int. Rev. Neurobiol.*, 18: 237-263.
- Grossberg, S (1976a), "On the development of feature detectors in the visual cortex with applications to learning and reaction-diffusion systems," *Biol. Cybern.*, vol. 21, pp. 145-159.
- Grossberg, S (1976b), "Adaptive pattern classification and universal recoding I: Parallel development and coding of neural feature detectors," *Biol. Cybern.*, vol. 23, pp. 121-134.
- Grossberg, S (1976c), "Adaptive pattern classification and universal recoding II: Feedback, expectations, olfaction, illusions," *Biol. Cybern.*, vol. 23, pp. 187-202.

- Grossberg, S (1987), "Competitive learning: From interactive activation to adaptive resonance," *Cognitive Science*, 11: 23-63.
- Inhelder, B. and J. Piaget (1964), *The Early Growth of Logic in the Child*, London: Routledge and Kegan Paul Ltd.
- Inhelder, B., H. Sinclair, and M. Bovet (1974), *Learning and the Development of Cognition*, Cambridge, MA: Harvard University Press.
- Levine, DS and S Grossberg (1976), "Visual illusions in neural networks: Line neutralization, tilt aftereffect, and angle expansion," *J. theor. Biol.*, 61: 477-504.
- Lundberg, A, K. Malmgren, and E.D. Schomburg (1987), "Reflex pathways from group II muscle afferents. 3. Secondary spindle afferents and the FRA: a new hypothesis," *Exp. Brain Res.*, 65: 294-306.
- Malsburg, Christoph von der (1994), "The correlation theory of brain function," in *Models of Neural Networks II: Temporal Aspects of Coding and Information Processing in Biological Systems*, E. Domany, J.L. van Hemmen, and K. Schulten (Eds.), NY: Springer-Verlag, pp. 95-119.
- Neumann, John von (2000), *The Computer and the Brain*, 2nd ed., New Haven, CN: The Yale University Press.
- O'Reilly, RC and Y Munakata (2000), *Computational Explorations in Cognitive Neuroscience*, Cambridge, MA: The MIT Press.
- Piaget, J. (1952), *The Origins of Intelligence in Children*, Madison, CN: International Universities Press.
- Piaget, J. (1954), *The Construction of Reality in the Child*, NY: Basic Books.
- Piaget, J. (1970), *Genetic Epistemology*, NY: W.W. Norton & Co.
- Piaget, J. (1975), *The Development of Thought*, A. Rosin (Tr.), N.Y.: Viking Press.
- Piaget, J. (1976), *The Grasp of Consciousness: Action and Concept in the Young Child*, Cambridge, MA: Harvard University Press.
- Piaget, J. and B. Inhelder (1967), *The Child's Conception of Space*, NY: W.W. Norton & Co.
- Piaget, J., J.-B. Grize, A. Szeminska, and V. Bang (1977), *Epistemology and Psychology of Functions*, Boston, MA: D. Reidel Publishing Co.
- Piaget, J. and R. Garcia (1991), *Toward a Logic of Meanings*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sherman, S.M. and R.W. Guillery, *Exploring the Thalamus and Its Role in Cortical Function*, 2nd ed., Cambridge, MA: The MIT Press, 2006.
- Szentágothai, J (1983), "The modular architectonic principle of neural centers," *Rev. Physiol. Biochem. Pharmacol.* **98**: 11-61.
- Wells, RB (2003a), "Synaptic weight modulation and adaptation, part I: Introduction and presynaptic mechanisms," LCNTR Tech Brief, Moscow, ID: The University of Idaho, <http://www.mrc.uidaho.edu/~rwells/techdocs>.
- Wells, RB (2003b), "Synaptic weight modulation and adaptation, part II: Postsynaptic mechanisms," LCNTR Tech Brief, Moscow, ID: The University of Idaho, <http://www.mrc.uidaho.edu/~rwells/techdocs>.
- Wells, RB (2003c), "Spinal sensorimotor system," LCNTR Tech Brief (in four parts), Moscow, ID: The University of Idaho, <http://www.mrc.uidaho.edu/~rwells/techdocs>.
- Wells, RB (2003d), "Muscles," LCNTR Tech Brief, Moscow, ID: The University of Idaho, <http://www.mrc.uidaho.edu/~rwells/techdocs>.
- Wells, RB (2005), "Cortical neurons and circuits: A tutorial introduction," LCNTR Tech Brief, Moscow, ID: The University of Idaho, <http://www.mrc.uidaho.edu/~rwells/techdocs>.
- Wells, RB (2006), *The Critical Philosophy and the Phenomenon of Mind*, an LCNTR E-book, Moscow,

ID: University of Idaho, <http://www.mrc.uidaho.edu/~rwells/Critical%20Philosophy%20and%20Mind>.

Wells, RB (2007), *Introduction to Biological Signal Processing and Computational Neuroscience*, an LCNTR E-book, Moscow, ID: The University of Idaho. Contact the author at [rwells@mrc.uidaho.edu](mailto:rwells@mrc.uidaho.edu).