Pergamon

PII: S0893-6080(97)00060-9

1997 SPECIAL ISSUE

A Neural Global Workspace Model for Conscious Attention

JAMES NEWMAN,[1] BERNARD J. BAARS[2] AND SUNG-BAE CHO[3]

[1]Colorado Neurological Institute, Denver, [2]The Wright Institute, Berkeley and [3]Department of Computer Science, Yonsei University

Abstract—*Considerable progress is being made in interdisciplinary efforts to develop a general theory of the neural correlates of consciousness. Developments of Baars' Global Workspace theory over the past decade are examples of this progress. Integrating experimental data and models from cognitive psychology, AI and neuroscience, we present a neurocognitive model in which consciousness is defined as a global integration and dissemination system — nested in a large-scale, distributed array of specialized bioprocessors — which controls the allocation of the processing resources of the central nervous system. It is posited that this global control is effected via cortical 'gating' of a strategic thalamic nucleus. The basic circuitry of this neural system is reasonably well understood, and can be modeled, to a first approximation, employing neural network principles.* © 1997 Elsevier Science Ltd.

Keywords—Attention, Binding, Consciousness, Global Workspace theory, Nucleus reticularis thalami, Programmable blackboard, Thalamocortical circuit, Thalamus.

1. INTRODUCTION

Consciousness has been widely portrayed as an intractable or irrelevant problem for cognitive science (e.g. Harnad, 1994; Penrose, 1994; O'Nuallain et al., 1997). Certainly its serious consideration is a fairly recent development (Jackendorf, 1987; Baars, 1988; Johnson-Laird, 1988; Edelman, 1989; Crick and Koch, 1990a), although a brief enthusiasm for the subject surfaced, and submerged, three decades ago (Eccles, 1966; Penfield, 1975). While it is not widely realized, the experimental neuroscience which served as the basis for that earlier enthusiasm is proving increasingly relevant to the present recrudescence of interest in conscious processes (see e.g. Stryker, 1989; Newman, 1995a, 1997). Beginning with historical developments in both AI and neuroscience, this paper reviews a growing body of evidence that some of the basic mechanisms underlying consciousness can be modeled, to a first approximation, employing variations upon current neural network architectures (see also Taylor, 1992; Taylor, 1996; Baars et al., in press; Newman et al., 1997).

Baars (1983, 1988, 1992, 1994) has developed a set of 'Global Workspace Models', based upon a unifying pattern, and addressing a substantial domain of evidence

explicitly related to conscious experience. These models explicate an architecture in which many parallel, non-conscious experts interact via a serial, conscious and internally consistent Global Workspace (GW), or its functional equivalent. GW, or blackboard, architectures were first developed by cognitive scientists in the 1970s and this framework is closely related to the Unified Theories of Cognition of Simon, Newell and Anderson (see Newell, 1992).

The HEARSAY model of speech understanding (Reddy et al., 1973) was one of the earliest attempts to simulate a massively parallel/interactive computing architecture. The notion of a global workspace was initially inspired by this architecture, consisting of a large number of knowledge modules, or 'local experts', all connected to a single 'blackboard', or problem-solving space. Activated experts could compete to post 'messages' (or hypotheses) on the blackboard for all the other experts to read. Incompatible messages would tend to inhibit each other, while the output of cooperating experts would gain increasing access to the blackboard until a global solution emerged. Blackboard architectures are relatively slow, cumbersome and error-prone, but are capable of producing solutions to problems too novel or complex to be solved by any extant modular knowledge source. Once such 'global solutions' are attained, however, the original problems can be allocated to modular processors for 'non-conscious' solution.

Corresponding Author, Present address: 740 Clarkson Street, Denver, CO 80218, USA; Tel.: (303) 832-9062; e-mail: newmanjb@aol.com.

McClelland (1986) attested to the significance of this set of models to subsequent developments in cognitive science when he described HEARSAY, not only as "a precursor of the interactive activation model", but "of the approach that underlies the whole field of parallel distributed processing" (p. 122). We consider McClelland's own 'Programmable Blackboard Model of Reading' as a connectionist example of a global workspace architecture, and discuss its applicability to modeling conscious processes in a concluding section.

Another class of models that may turn out to be compatible with GW theory comes from 'distributed artificial intelligence' (DAI), which Durfee (1993) characterizes as the study of "how intelligent agents coordinate their activities to collectively solve problems that are beyond their individual capabilities" (p. 84). He cites examples of DAI applications, such as generic conflict resolution, unified negotiation protocols, and search-based models of coordination/cooperation. DAI applications appear to more closely approximate human interpersonal behaviour than purely logic-driven AI. They require that agents learn to be "knowledgeable and skilled in interacting with others" (p. 86). DAI models would appear to reflect an intelligent balance between competitive self-interest and cooperative problem-solving that is essential to optimizing overall outcomes in complex 'social' organizations. This, like GW theory, is consistent with other well-known metaphors in cognitive science, such as Minsky's 'Society Theory' (Minsky, 1979) and Gazzaniga's 'Social Brain' (Gazzanigga, 1985).

A similar, globally-integrative balancing of priorities appears to characterize the optimal processing of conscious information. Conscious percepts are characterized by unified gestalts of shape, texture, color, location and movement, despite the fact that these contributions to perception are initially processed in parallel areas of the cortex, in both hemispheres. Moreover, conscious intentions are generally single-minded and goal-directed. Of course, conflicts can and do arise, but a central purpose of consciousness seems to be resolving such conflicts (employing both integrative and inhibitory algorithms).

While such global states can be highly adaptive — indeed, are essential to explicit learning — GW theory maintains that the vast majority of cognitive tasks performed by the human brain are automatic, and largely non-conscious (Baars, 1988; Newman and Baars, 1993; Baars, 1997; Newman, 1997). Consciousness generally comes in play when stimuli are assessed to be novel, threatening, or momentarily relevant to active schemas or intentions.

The defining properties of stimuli which engage conscious attention (i.e. the global allocation of processing resources) are that they: 1) vary in some significant degree from current expectations; or 2) are congruent with the current, predominant intent/goal of the organism. In contrast, the processing of stimuli which are predictable, routine or over-learned is automatically allocated to non-conscious, highly modularized cognitive systems (Newman, 1995b, p. 691).

Generally, we are conscious of what has the highest relevance to us at that moment. This may be a momentary threat, a sudden insight, a pleasant sensation, etc. (in relaxed moments, there may be no particular focus or intent, simply a stream of associations). Yet, while the range of our awareness is immense (limited only by our most developed cognitive capacities), we contend that the basic mechanism for the allocation of these capacities remains constant under virtually all contingencies; and the basic neural circuitry of that resource-allocation mechanism is reasonably well understood. Indeed, in subsequent sections, we suggest how it might be modeled based upon already existing neural network simulations (McClelland, 1985; Hampshire and Waibel, 1992; Taylor and Alavi, 1993; Llinas et al., 1994).

The relevance of Global Workspace theory extends beyond NN modeling, however. Indeed, it bears upon central philosophical problems in consciousness studies, such as the *homunculus* and Cartesian theater. The two are, of course, related. The image of a 'little man in our head' observing and manipulating the play of conscious images is beguiling, but absurd. For who is this strange being lodged in our brains? And who is watching him?

In Global Workspace theory the single homunculus is replaced by a large 'audience of experts'. The 'theater of consciousness' then becomes a workspace, with stage (Baars, 1997). Almost everyone in an audience has potential access to center stage (although most prefer to simply observe, or exert indirect influences). The focus of conscious activity, at any moment, corresponds to the 'work' produced by the most active coalition of experts, or modular processors: whoever has managed to win the competition for 'the spotlight'. There is no fixed, superordinate observer. Individual modules can pay as much or as little attention as suits them, based upon their particular expertise. At any one moment, some may be dozing in their seats, others busy on stage. In this sense, the global workspace resembles more a deliberative body than a theater audience. Each expert has a certain degree of 'influence', and by forming coalitions with other experts can contribute to deciding which issues receive immediate attention and which are 'sent back to committee'. Most of the work of this deliberative body is done 'off stage' (i.e. non-consciously). Only matters of greatest relevance in-the-moment gain access to consciousness.

While the GW is a teaming multiplicity, what is explicitly represented in consciousness is largely coherent and adaptive. The overall workspace serves as a 'global integration and dissemination system', in which all experts can participate, but only select coalitions dominate, momentarily, producing an orderly succession

of global representations. The stream of consciousness arises out of the operations of the GW system — and, over time, our sense of being a coherent 'I' (the memory and intention systems vital to this aspect are beyond the scope of this paper; see Baars et al., in press; Newman, 1997). It is this unitary awareness, not any agent or homunculus, that is globally superordinate. Of course, such a system is prone to inefficiencies and pathological perturbations, but this is consistent with the scientific literature concerning human consciousness (see Baars, 1988).

If we are to proceed beyond pleasing metaphors, however, it is necessary to operationalize the GW model in explicit neurocognitive terms. This process begins in the next section. To introduce it, we offer the following working definition:

> consciousness reflects the operations of a global integration and dissemination system, nested in a large-scale, distributed array of specialized bioprocessors; among the various functions of this system are the allocation of processing resources based, first, upon biological contingencies of novelty, need or potential threat and, secondly, cognitive schemas, purposes and plans.

2. MODELING GLOBAL, COMPETITIVE ATTENTION

We have introduced the theoretical background for the model. Newman and Baars (1993) and Newman (1997) present detailed accounts of its neural architecture. We would stress, however, that consciousness is a dynamic process, not a static structure. Also, it is not localized to some 'brain center', but arises out of the coordinated activities of widely distributed networks of neurons. Resource allocation is integral to these activities. The neural bases of resource allocation, or attention, have been extensively explored (see, e.g. Heilman et al., 1985; Mesulam, 1985; Posner and Rothbart, 1991; Posner, 1994; LaBerge, 1990, 1995). But, of course, not all forms of attention are conscious. As an example from AI, McClelland (1986) notes that in simulations of reading, activated modules must be 'sticky', that is "interactive activation processes continue in older parts of the programmable blackboard while they are being set up in newer parts as the eye moves along..." (pp. 150–151). This 'stickiness' would seem to entail a type of attention. It normally proceeds quite automatically, however, both in a reading machine and in a literate human being. Only when the process is disrupted by, say, a mis-spelled or unknown word, does that word becomes the focus of our conscious awareness. Normally, we are only conscious of the overall sense of the passage of text, and the images and thoughts it evokes, not particular semantic or syntactical operations. These linguistic processes became second nature to us

long ago. Such 'particular operations' are hardly trivial aspects of language acquisition, but as Kihlstrom (1987) noted, in humans they tend to be "automatized through experience and thus rendered unconscious" (p. 285).

Conscious awareness clearly involves a higher order of resource allocation, which Newman and Baars (1993) call 'global attention'. The term "refers to a level of cognitive processing at which a single, coherent stream of information emerges out of the diverse activities of the CNS" (p. 258). The focus of that stream could (under atypical circumstances) be an individual word; but the conscious mind seldom confines itself to the processing of such rudimentary representations. Rather it seems to be decisively biased towards multifaceted, yet unified images. Thus, we are able to perceive a Necker Cube as projecting out of a two-dimensional page, alternately to the left, then to the right; but we are curiously incapable of perceiving these two perspectives simultaneously.

The processing load of global attention (like working memory), is both highly chunked and highly restricted (Baars, 1988). The non-conscious allocation of processing resources operates under no such constraints. For example, neuroscience has shown that specialized areas in the visual cortex process, in parallel, the contour, movement, color, spatial location, etc. of a stimulus (LaBerge, 1995). Yet our awareness is of a single, coherent object (and often includes tactile, auditory and associative aspects). Thus, neuroscience is faced with the 'binding problem' of how these multifarious representations, generated by widely separated areas, are integrated into real-time 'objects' of perception (see Crick and Koch, 1990a; Newman and Baars, 1993).

One would expect the neural mechanism for global attention to be complex, and widely distributed, which it is. But the basic circuitry can be described, to a first approximation, in terms of repeating, parallel loops of thalamo–cortico–thalamic axons, passing through a thin sheet of neurons known as the *nucleus reticularis thalami* (nRt). The loops are formed by long-axoned, excitatory neurons. The neurons of nRt are largely GABA-ergic, inhibitory neurons. Most, if not all, of the looping axons give off collaterals as they pass through nRt, while nRt neurons themselves project mainly to cells of the particular thalamic nucleus lying directly beneath them. There is an orderly topography to this array of axon collaterals and underlying thalamic nuclei (Scheibel and Scheibel, 1966; Mitrofanis and Guillery, 1993). It essentially mirrors, in miniature, the modular architecture of the cortex (see Newman and Baars, 1993; LaBerge, 1995; Newman, 1997, for reviews).

Evidence for the central role of this 'thalamocortical circuit' (LaBerge, 1995) in attention and consciousness has been accumulating for decades (Jasper, 1960; Scheibel, 1980; Jones, 1985; Steriade and Llinas, 1988; Llinas and Pare, 1991). Skinner and Yingling (1977) first proposed a neural model for its role in selective attention. Our 'wagon wheel' model (next section) represents a
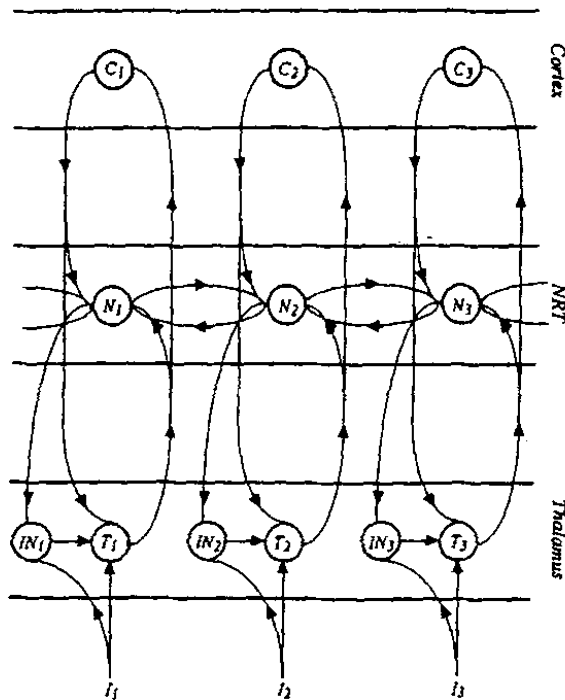
FIGURE 1. The wiring diagram of the main model of the thalamus–NRT–cortex complex. Input I_j is sent both to the thalamic relay cell T_j and the inhibitory interneuron IN_j, which latter cell also feeds to T_j. Output from T_j goes up to the corresponding cortical cell C_j which returns its output to T_j. Both the axons T_jC_j and C_jT_j send axon collaterals to the corresponding NRT cell N_j. There is axonal output from N_j to IN_j, as well as collaterals to neighbouring NRT cells. There are also dendro–dendritic synapses between the NRT cells (from Taylor and Alavi, 1993).

synthesis of both the accumulated evidence, and related models (Skinner and Yingling, 1977; Scheibel, 1980; Crick, 1984; Taylor and Alavi, 1993; Llinas et al., 1994; LaBerge, 1995). These related models vary in their details, as the precise connectivities and physiology of the thalamocortical circuit are not fully worked out.

Most attentional models are based upon conventional simulations of mechanisms such as center-surround inhibition, or winner-take-all (WTA) competitions, among local circuits. Various researchers have described the network of nRt neurons as a mosaic, or array, of neural 'gatelets' acting to selectively filter the flow of sensory inputs to the cortex (Skinner and Yingling, 1977; Scheibel, 1980; Crick, 1984). The WTA dynamic may seem analogous to the 'competition' posited by GW theory. The problem with such conventional networks is that they are poorly suited to global forms of competition, because prohibitively long-range and geometrically increasing numbers of connections would be required. Moreover, most long-range, reciprocal connections in the CNS are excitatory. Inhibitory effects tend to be local.

Taylor and Alavi (1993), however, have modeled a

competitive network for global attention based upon a highly simplified version of the 'thalamus–NRT–cortex complex'. Their model is unique, in that it takes into account the effects of dendro–dendritic interactions throughout nRt. The dendrites of nRt cells project out tangentially within the reticular sheet, bidirectionally. The physiology of information processing in dendritic trees is highly complex, and not well understood (Mel, 1994); but Koch and Poggio (1992) review evidence for the dendritic trees playing a role in several types of second-order, multiplicative computations. We will have more to say about this subsequently.

Figure 1 [taken from Taylor and Alavi (1993)] illustrates three thalamocortical circuits, as well as the nonlinear, dendro–dendritic connections between N1, N2, N3,... within NRT. We would refer the reader to the original paper for a detailed description of the simulations carried out, employing a network of 100 thalamocortical loops. To briefly summarize the results, the addition of dendro–dendritic connections to the looping circuits provided "the basis for a simple version of the global gating model... that instantiates a form of competition in the spatial wavelength parameters of incoming inputs..." (p. 352). In this version of the model, the entire nRt network oscillates with a wavelength,

with the net strength given by the component of the input with the same wavelength.

The way in which global control arises now becomes clear. Only those inputs which have special spatial wavelength oscillations are allowed through to the cortex, or are allowed to persist in those regions of the cortex strongly connected to the NRT: the thalamus–NRT system acts as a spatial Fourier filter (p. 353).

Simulation runs demonstrated the global, wave-like properties of the competitive model. The overall pattern of activation in cortical units was shown to be exclusively dependent upon the wave pattern spanning across all of the NRT units (Figure 2). As LaBerge (1995) notes, the actual physiology of nRt gating in alert states remains unclear, but it is firmly established that nRt is the source of global oscillatory activity (at 8–13 Hz) initiating the descent into sleep.

The RN cells are known to inhibit each other, and when inhibition hyperpolarizes an RN cell sufficiently, it produces a rebound burst. In this way a network of connected RN inhibitory cells can spread activity to every cell within the network, apparently without decrement in the intensity of the activity (p. 184).

Here then, is a plausible circuitry for a global, winner-take-all competition among the large array of specialized cortical processing areas.

Llinas et al. (1994) offer an interesting variation upon this circuitry, in which thalamocortical loops of the 'non-specific' intralaminar nuclei operate in parallel with the specific (input) loops described above. The synchronous
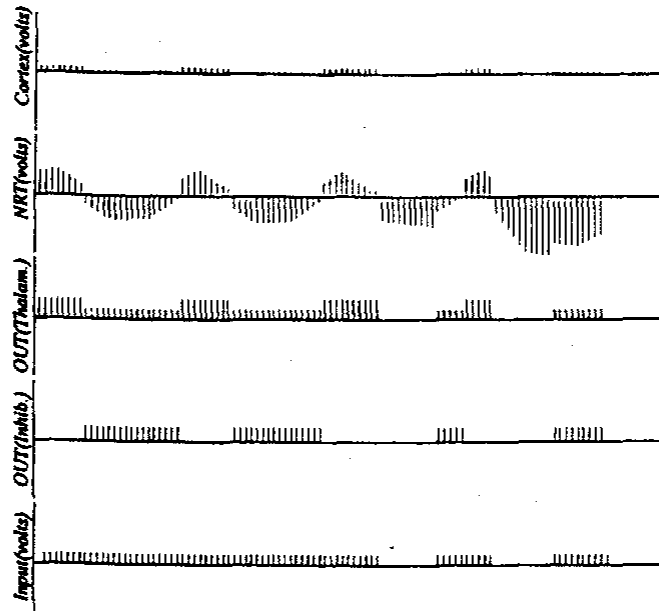
FIGURE 2. One of 15 simulation runs for the thalamus–NRT–cortex model showing full global control with semi-constant spatial input. Note that cortex activity is influenced by the NRT alone (from Taylor and Alavi, 1993).

activation of specific and non-specific loops is postulated to provide a basis for "perceptual unity... by which different sensory components are gathered into one global image" (p. 251). Their modeling is concerned with high-frequency EEG oscillations (and omits dendro–dendritic connections), yet appears to parallel much of what we discuss above.

> When the interconnectivity of these nuclei is combined with the intrinsic properties of the individual neurons, a network for resonant neuronal oscillations emerges in which specific corticothalamic circuits would tend to resonate at 40 Hz. According to this hypothesis, neurons at the different levels, and particularly those in the reticular nucleus, would be responsible for the synchronization of 40-Hz oscillations in distant thalamic and cortical sites...these oscillations may be organized globally over the CNS, especially as it has been shown that neighboring reticular cells are linked by dendrito–dendritic and intranuclear axon collaterals (Deschenes et al., 1985; Yen et al., 1985, pp. 253–254).

3. A NEURAL MODEL FOR GLOBAL RESOURCE ALLOCATION

We have introduced a set of convergent models for the basic circuitry of a Global Workspace system involved in the integration and dissemination of the processing resources of the nervous system. This 'bare bones' version accounts for how a global, winner-take-all competition might be mediated between various external inputs and cortical modules, to produce "a single, coherent stream of information out of the diverse activities of

the CNS" (Newman and Baars, 1993). There remains to be explained how the thalamocortical circuit fits in with the second half of our working definition for the conscious system: the allocation of processing resources
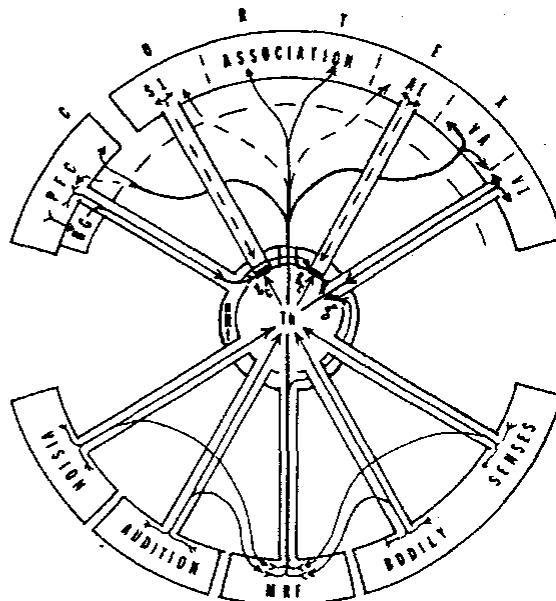

FIGURE 3. 'Wagon wheel' model of CNS systems contributing to global attention and conscious perception. A1, primary auditory area; BG, basal ganglia; g_c, 'closed' nRt gate; g_o, 'open' nRt gate; MRF, midbrain reticular formation; nRt, nucleus reticularis thalami; PFC, prefrontal cortex; S1, primary somatosensory area; Th, ventral thalamus; V1, primary visual cortex (from Newman et al., 1997).

based, first, upon biological contingencies of novelty, need or potential threat and, secondly, cognitive schemas, purposes and plans. In keeping with our definition, we will first add a sub-cortical component for orienting to "novelty, need, or potential threat", and then discuss the much more complex aspects of cortically-mediated effects upon the system.

This extended version of the model is schematically illustrated in Figure 3 as a 'wagon wheel', with the thalamus (Th) as its 'hub'. The reticular nucleus (nRt) corresponds to the metal sleeve fitted around the hub. The upper rim of the wheel represents the cerebral cortex (PFC/S1...V1), and closely associated basal ganglia (BG). The lower half shows the major sensory systems and subcortical nuclei whose projections converge upon the thalamus. The outer 'spokes' represent the sensory pathways for vision, audition and the bodily senses. These project, in an orderly topography, to modality-specific nuclei in the thalamic 'hub'. As they ascend towards the thalamus, these pathways give off collaterals to the midbrain reticular formation (MRF) (see also Newman et al., 1997). Scheibel (1980) reviewed three decades of experimental evidence indicating that these midbrain collaterals serve as the basis for an initial 'spatial envelope', or global map, of the environment surrounding the animal.

> Most reticular [MRF] neurons...appear multimodal, responding to particular visual, somatic and auditory stimuli, with combinations of the last two stimuli most numerous. The common receptive fields of typical bimodal cells in this array show a significant degree of congruence. For instance a unit responding to stimulation of the hind limb will usually prove maximally sensitive to auditory stimuli originating to the rear of the organism. These twin somatic and auditory maps retain approximate register and overlap the visuotopic map laid down in the...superior colliculus...These data might be interpreted to mean that each locus maps a point in the three-dimensional spatial envelope surrounding the organism. Further studies suggest the presence of a deep motor map closely matching and in apparent register with the sensory map. (p. 63)

More recent research has supported Scheibel's portrayal of the superior colliculus as the visual component of what Crick and Koch (1990b) termed a 'saliency map' for eye movements, involved in orienting the animal to biologically relevant stimuli. Subsequent findings have both confirmed Scheibel's analysis, and revealed a number of 'top-down' projections that modulate activities in MRF. LaBerge (1995) writes:

> the superficial area [of the superior colliculus] receives strong cortical inputs from V1, V2 and V3 [primary and secondary visual cortex], the deep layers in the monkey SC receive their main cortical inputs from the posterior parietal area (Lynch et al., 1985), from the prefrontal areas (Goldman and Nauta, 1976), and the frontal eye fields (Leichnetz et al., 1981). The deep layers contain a map of visual space that is stacked adjacent to maps for auditory

and somatosensory spaces in a manner that cells corresponding to points in space lie along the same vertical axis (Merideth and Stein, 1990). Stimulation of these cells by microelectrodes produces movements of eyes and head... (LaBerge, 1995, p. 145)

LaBerge goes on to describe basal ganglia inputs that "are of particular importance because they tonically inhibit activity in the SC cells". It has long been known that the frontal eye fields, and posterior parietal area "exert strong influences on eye movements and must be considered together with the superior colliculus in accounting for...orienting of attention" (p. 142). These facts emphasize two key aspects of the 'conscious system' we are modeling: 1) it is polymodal, integrating not just visual, auditory and somatosensory inputs, but motor and 'higher-order' cortical effects; and 2) it is extended, with input/output relations reaching from the brain stem core to association cortices. Indeed, the general term we have used to describe it elsewhere is the 'extended reticular-thalamic activation system', or 'ERTAS' (Baars, 1988; Newman and Baars, 1993; Newman, 1995a, 1995b, 1997).

The third key aspect of the system (as exemplified by the 'wagon wheel' model) is that it converges on the thalamus. We have already discussed this in terms of the thalamocortical circuit, which connects to "virtually every area of the cerebral cortex" (LaBerge, 1995, p. 221). Scheibel (1980) described the MRF portion of the system as:

> sweep[ing] forward on a broad front, investing the [intralaminar complex of the] thalamus and nucleus reticularis thalami. The investiture is precise in the sense that the sites representing specific zones of the spatial envelope (receptive field) project to portions of the nucleus reticularis concerned with similar peripheral fields via projections from both sensory thalamus and sensory association cortices (p. 62).

The fact that Scheibel's (1980) 'spatial envelope' projects with some topographic precision upon nRt, would appear to enable it to disinhibit particular arrays of nRt gatelets, selectively enhancing the flow of sensory information to the cortex. The 'intralaminar complex' (Newman and Baars, 1993) is also integral to the ERTAS system, as the non-specific portion of the thalamocortical circuit. It is intralaminar projections which relay MRF activation to the cortex (illustrated by the vertical MRF-Th projection, above which it branches out to all areas of CORTEX). As noted above, Llinas et al. (1994) hypothesize the perceptual unity of consciousness (binding) to be brought about by the global synchronization of specific and non-specific circuits via nRt. Scheibel (1980) earlier concluded as much concerning the role of this extended activation system in 'selective awareness':

> From these data, the concept emerges of a reticularis

complex [nRt] selectively gating interaction between specific thalamic nuclei and the cerebral cortex under the opposed but complementary control of the brain stem reticular core [MRF] and the frontal granular cortex [PFC]. In addition, the gate is highly selective; thus, depending on the nature of the alerting stimulus or central excitation, only that portion of the nucleus reticularis will open which controls the appropriate subjacent thalamic sensory field. The reticularis gate [thus] becomes a mosaic of gatelets, each tied to some specific receptive zone or species of input. Each is under the delicate yet opposed control of: (a) the specifically signatured sensory input and its integrated feedback from cortex [S1...V1]; (b) the reticular core [MRF] with its concern more for novelty (danger?) than for specific details of experience; and (c) the frontal granular cortex-medial thalamic system [PFC/BG] more attuned to upper level strategies of the organism, whether based on drive mechanisms (food, sex) or on still more complex derivative phenomenon (curiosity, altruism). Perhaps here resides the structuro- functional substrate for selective awareness, and in the delicacy and complexity of its connections, our source of knowing, and of knowing that we know (p. 63).

Here, as well, is a summary description of a neural substrate for the global allocation of the processing resources of the CNS. All that it lacks is the mechanisms for a global competition (Taylor and Alavi, 1993) and binding (Llinas et al., 1994) introduced in the previous section. But we must tie the operations of this thalamus-centered system more closely to those of the cortex and basal ganglia, or most of the functions routinely studied by cognitive science have no place in the model. This introduces an exponentially higher level of complexity (one of the hazards of dealing with global systems).

One of the values of GW theory, however, is that it provides a framework for understanding this complexity. First, it holds that the vast majority of cognitive functions are carried out, non-consciously, via changing arrays of specialized, modular processors. This is reflected, anatomically, in the immense number of cortico-cortical connections in the human brain, outnumbering those with subcortical nuclei by nearly ten to one. Thalamocortical projections are comparatively sparse, but serve at least two essential functions: 1) transmitting sensory inputs to the primary cortical areas (S1, A1, V1, Figure 3); and 2) providing a means to selectively amplify/synchronize cortex-wide activation (Section 2).

GW theory also reminds us that conscious functions operate upon an information load about the size of working memory. Thus, we are talking of a highly coarse-grained level of processing. In this context, global attention is (at least) a second-order operation, acting upon a highly selective stream of information. All this is to say that a relatively low density of widely distributed, yet highly convergent, circuits could be all that are required to create a conscious system; and these are the very characteristics of the neural model we have described.

However, most neural network modelers take a

cortically-centered view of cognition, from which the brain stem functions so far described probably seem rather primitive or trivial (i.e. orienting, controlling eye movements) when compared to cortically-mediated processes such as language acquisition, pattern recognition, motor planning, etc. What evidence is there that cortical (and other forebrain systems) depend upon projections to the thalamus for effecting high-level cognitive processes?

Early support for such effects, mediated by prefrontal projections, was provided by animal experiments undertaken by Skinner and Yingling (1977). They found that selective activation of one portion of a fronto-thalamic tract could shut down sensory processing in visual, but not auditory, cortex. Activation of another 'spoke' of the prefrontal-thalamic tract shut down auditory processing, but allowed visual inputs to reach posterior cortex. Skinner and Yingling wrote "This result implies that selective attention emerges via selective inhibition in certain sensory channels that the animal must **know in advance** are irrelevant to its situation" (p. 54). To inhibit orienting based upon advanced knowledge is clearly a sophisticated use of cognition. Several lines of research have converged in recent years to support this concept. Summarizing the current state of knowledge of prefrontal regulation of subcortical systems, Newman (1997) wrote:

> It is now generally accepted that the prefrontal lobes (with the cingulate cortex) constitute an 'executive' over the limbic system mediating such functions as working memory, inhibition of conditioned responses, and goal-directed attention (see Fuster, 1980; Goldman-Rakic, 1988b; Damasio, 1994; Posner, 1994). More recent research on the basal ganglia (see reviews by Groenewegen and Berendse, 1994; Parent and Hazrati, 1995) have suggested that they constitute a 'motor programming extension' of the frontal lobes as well — routed through the thalamus (p. 112–113).

Newman (1997) goes on to cite evidence (Parent and Hazrati, 1995) that the BG 'extension' (like the thalamo-cortical loops) sends rich, collateral projections to nRt that effect not only its 'gating' of motor programs, but hippocampal-mediated episodic memory functions (see also Newman, 1995b).

Finally, we would note that cortico-thalamic projections to nRt and associated specific nuclei are both more topographically precise (Mitrofanis and Guillery, 1993) and more pervasive than had once been thought (Jones, 1985). Llinas and Pare (1991) estimate that, for every axon the thalamus sends to the cortex, the cortical area it projects to reciprocates with ten. Given the modular architecture of the neocortex, one might reasonably predict that these cortico-thalamic projections exert highly differentiated influences upon the flow of information through the thalamus. Efforts by experimental neuroscience throughout the 1980s to elucidate the precise effects of cortico-thalamic projections were

frustratingly inconclusive. But a recent review by Buser and Rougeul-Buser (1995) notes:

> The situation has however recently swung back, due to some new and perhaps consistent findings, indicating that the visual cortex appears to have a major action down onto the lateral geniculate nucleus, which may generate thalamic oscillations (Funke and Eyse, 1992; McCormick and Krosigk, 1992; Krosigk et al., 1993; Sillito et al., 1994) (p. 252).

While additional research is clearly needed, these recent findings suggest that Scheibel's (1980) early model of the converging influences of projections upon a thalamic hub — with the addition of basal ganglia inputs to nRt and the intralaminar complex — remains a viable model for "global attention", including the influences of cortically generated "schemas, purposes and plans". Newman (1997) discusses the contributions of the 'cortico-basal ganglia–thalamo–cortical loop' to memory and volitional processes in greater detail. The complexities of this system are beyond the scope of the models presented here, although Monchi and Taylor (1995) and Taylor and Michalis (1995), among others, have developed neural models simulating functions of the BG and hippocampal systems.

What we propose to do instead is present a much simpler, but highly relevant, connectionist model that simulates the sorts of second-order operations one would predict in a GW system employing a gating network to selectively filter and integrate inputs as a function of central knowledge stores. The basic heuristic for this type of model is described in Newman and Baars (1993). It posits that

> prefrontal cortex acts as an executive attentional system by actively influencing information processing in the posterior cortex through its effects upon the nucleus reticularis. In this manner, the highly parallel [processing] functions of the posterior cortex are brought into accord with increasingly complex and intentional cognitive schemes generated within the prefrontal regions of the brain (p. 281).

A defining property of an executive system is that it acts upon other sub-systems, modifying their inputs for its particular purposes. Posterior cortical areas act more like arrays of quasi-autonomous processing modules (or local experts) — the bread and butter of NN simulations. Note that an executive system is not an *essential* requirement for consciousness. That this is the case is illustrated by the literature on extensive damage to the frontal lobes of the brain. PFC damage results in significant deficits in such purposeful activities as: the inhibition of inappropriate responding; switching of response set, planning and monitoring of actions, etc.; but produces little or no alteration in basic mental status. Indeed, many patients with frontal lobe pathology perform at pre-morbid levels on intelligence tests (Walshe, 1978; Damasio, 1994). In terms of the GW model we have

presented, it is not executive attentional processes, but the selective binding of coalitions of active cortical modules via a thalamocortical competition which is the *sine qua non* for the generation of a coherent stream of conscious representations. Examples of these aspects of the GW model have already been offered.

4. SECOND-ORDER MODELS FOR GLOBAL GATING

Let us return to the 'wagon wheel' model illustrated in Figure 3, and transform its components into a connectionist GW, with an executive system. To simplify things, the network will have only two sensory modules, one for processing auditory (A1) inputs, and one for visual (V1). In order to provide second-order control over processing in both modules, we will add a gating module (nRt) with the same number of units as connections in each sensory module. Each gating unit sends its output to a corresponding connection in A1 and V1. The connections between the gating units and sensory units are multiplicative. As Rummelhart et al. (1986) write about such connections:

> if one unit of a multiplicative pair is zero, the other member of the pair can have no effect, no matter how strong its output. On the other hand, if one unit of a pair has value 1, the output of the other passe[s] unchanged to the receiving unit....In addition to their use as gates [such] units can be used to convert the output level of a unit into a signal that acts like a **weight** connecting two units (p. 73).

In this manner, a fully connected gating module can actually program the connection strengths of one or more input modules to process a particular type of input, for example phonemes, or letters, into words. For maximum flexibility, it is preferable that the gating module not have fixed connections either, but simply relay gating (connection strength) information from a central module to which its units are connected. The central module contains (in this case) word-level knowledge needed to program the sensory modules to process words. Another central module might be specialized for knowledge for processing visual scenes or tactile shapes. To complete the system, each programmable input unit sends a corresponding connection to a central module unit.

The highly simplified network just described is really a variation on a 'Programmable Blackboard Model for Reading' developed by McClelland (1985, 1986). Its four modules correspond to those labeled in Figure 4: a Central Module (PFC); Connection Activation System (PFC-nRt); and two Programmable Modules (A1, V1). The connections described above are shown in Figure 5 (note: McClelland's modules are identical, and used only for reading (not hearing) words, but theoretically they could be programmed to process *any* type of input).
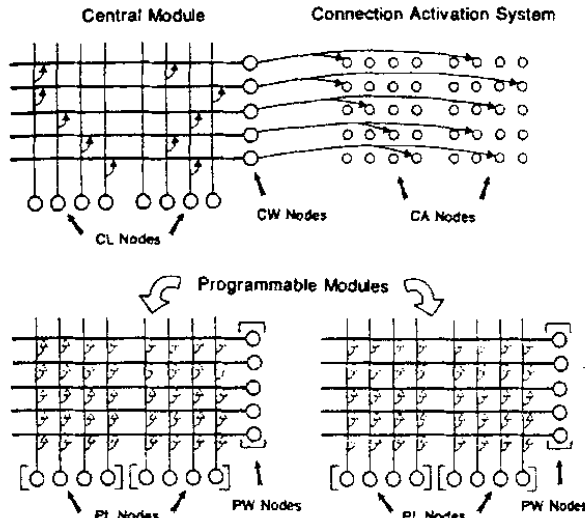
FIGURE 4. A simplified example of a Connection Information Distributor (CID), sufficient for simultaneous bottom-up processing of two two-letter words. The programmable modules consist of the programmable letter (PL) and programmable word (PW) nodes, and programmable connections between them (open triangles). The central module consists of a set of central letter (CL) nodes and a set of central word (CW) nodes, and hard-wired connections between them (filled triangles). The connection activation system includes the central word nodes, a set of connection activator (CA) nodes, and hard-wired connections between them. Connections between the central knowledge system (central module plus connection activation system) and the programmable blackboard are shown in Figure 6 (from McClelland, 1985).
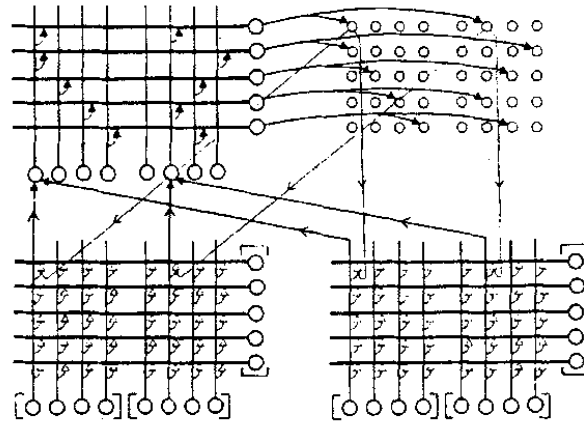

FIGURE 5. Each CA node projects to the corresponding connection in both programmable modules, and each central letter node receives projections from the corresponding programmable letter node in both programmable modules. The inputs to two central letter nodes, and the outputs from two CA nodes are shown (from McClelland, 1985).

In the brain, of course, the primary areas (A1, V1, S1) send no direct projections to PFC; but they do send convergent projections (as in Figure 5) to secondary association areas, which send projections directly to PFC (as well as posterior association areas). Although these feed-forward projections to PFC are less topographically precise (e.g. the receptive fields of visual neurons in the secondary areas are much larger), they maintain a fair degree of parallel distribution, indicating that much of the prefrontal cortex is as modular in its organization as the posterior 'association' cortex. Moreover, PFC 'modules' reciprocate these parallel, feedforward projections, although in a more divergent pattern (Goldman-Rakic, 1988a; LaBerge, 1995). Interestingly, this convergence/divergence pattern is paralleled by the connections in Figure 5 for the central module.

In the actual prefrontal cortex there are hundreds (if not thousands) of 'central modules'. Feed-forward inputs allow them to use and store highly processed information from the posterior (sensory) cortex. Of course, feedback (or re-entrant) connections enable PFC to influence processing in the posterior areas as well. But such divergent and indirect feedback pathways are poorly suited to exercising momentary, direct effects upon processing at the input level. Nor could such centrally-stored knowledge be employed to guide, or anticipate, how

inputs are processed (re the 'knowing in advance' Skinner and Yingling (1977) attributed to PFC-Th circuits). This is where direct projections to the primary processing areas (actually the thalamocortical circuit) could prove quite valuable. Instead of the sensory input units (A1-Th; V1-Th) responding based upon fixed connection strengths, a central module could program input modules to process (i.e. pay attention to) particular categories of inputs. McClelland (1986) calls this form of activation 'connection information distribution' (CID) and compares its benefits to those of:

> the invention of the stored program....The use of centrally stored connection information to program local processing structures is analogous. It allows the very same processing structures to be programmed to perform a very wide range of different tasks.... [CID] also carries out a form of what is known in production systems as 'resolution', binding the right tokens in the blackboard together into higher-order structural patterns (p. 165).

Finally, he notes analogous aspects in the CID's operations to 'working memory', a process which has been tied by neuroscientists to a prefrontal/thalamic/hippocampal system (e.g. Fuster, 1980; Goldman-Rakic, 1988b). These comparisons between the Wagon Wheel and Programmable Blackboard models, of course, have purely heuristic value (although McClelland's (1986) PABLO simulation of his model contained a sufficient programmable blackboard to read lines of text up to 20 characters long). But the use of gating networks to generate useful 'higher-order structural patterns' is fairly widespread.

For engineering problems such as object recognition and robot motion control, the concept of combining modular networks using gating connections has been actively exploited to develop highly reliable systems (Jacobs et al., 1991; Hampshire and Waibel, 1992;
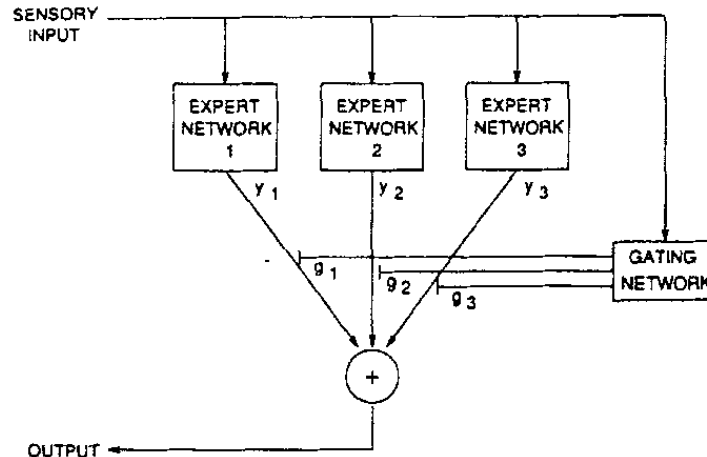
FIGURE 6. Schematic diagram of modular neural networks with three expert networks and a gating network. The output of the entire architecture, denoted by Y, is $Y = g_1y_1 + g_2y_2 + g_3y_3$, where y_i denotes the output of the ith expert network.

Jacobs and Jordan, 1993; Cho and Kim, 1995). The key issue in this approach is how to combine the results of the individual networks to give the best estimate of the optimal overall result. Architectures used in this approach consist of two types of networks: an expert and a gating network. Basically, the expert networks compete to learn the training instances, and the gating network facilitates cooperation by the overall mediation of this competition. The expert networks may be trained separately using their own preassigned sub-tasks and differing modalities (e.g. vision and touch), or the same modality at different times (e.g. the consecutive 2-D views of a rotating 3-D object). The gating network need only have as many output units as there are expert networks.

To train such a gating network, Hampshire and Waibel (1992) developed a new form of multiplicative connection, which they call the 'Meta-Pi' connection. Its function is closely aligned with predecessors described in McClelland (1986). The final output of the overall system is a linear combination of the outputs of the expert networks, with the gating network determining the proportion of each local output in the linear combination. Figure 6 illustrates this architecture with three expert networks.

The final output of the overall system is a linear combination of the outputs of the expert networks, with the gating network determining the proportion of each local output in the linear combination. The Meta-Pi gating network allocates appropriate combinations of the expert networks when stimuli are assessed to be novel, while an automatic ('non-conscious') decision process operates in instances where a single expert can execute the task. This coupling of modular, expert networks and gating controls produces new levels of cooperative behavior. The expert networks are local in the sense that the weights in one network are decoupled from the weights in other expert networks. However,

there is still some indirect coupling because if some other network changes its weights, it may cause the gating network to alter the responsibilities that get assigned to the expert network.

These examples from engineering applications of multiplicative, gating networks are not based upon the Wagon Wheel model or, for that matter, any specific neural circuitry. Yet Koch (1997) notes that

Multiplication is one of the most common operations carried out in the nervous system (for example, for estimating motion or the time-to-contact with an approaching stimulus) (p. 208).

We are not aware of any studies of either the axon collateral or dendro–dendritic projections in nRt demonstrating multiplicative properties, but Mel (1994) has modeled such connections in the NMDA-rich dendritic trees of cortical pyramidal cells. He postulates that they perform nonlinear pattern discrimination and correlative operations. Given the role of the bidirectional dendritic trees of nRt cells in globally synchronizing the thalamo-cortical circuit (Taylor and Alavi, 1993; LaBerge, 1995), it seems likely that they will eventually be found to have important computational functions as well.

Even if it transpires that synchronous oscillations, not multiplicative connections, are the basis for the 'gating' functions of nRt upon the thalamocortical circuit, NN models based upon Meta-Pi connections may still be useful for simulating global workspace systems. The use of Meta-Pi connections has already been extended to synchronous oscillators in modular cortical neural networks. Indeed, computational simulations of phase-locked oscillations characteristic of neurons involved in the 'binding' of visual (Grossberg and Somers, 1991; Sompolinsky et al., 1991) and auditory (Vibert et al., 1994) features of an attended object have already been extended to synchronous oscillators using Meta-Pi connections. Such oscillatory circuits have also been

employed in modeling storage and retrieval in pattern recognition tasks (Yao and Freeman, 1990).

In this paper, we have introduced a collection of neuroscience and NN models for attention and binding, resource allocation, and second-order gating, which share important features and parallels with a Neural Global Workspace System for conscious attention (Newman and Baars, 1993). While the NN models we have presented only implement partial aspects of the GW system, and even our Wagon Wheel model largely neglects the influences of memory and affective systems upon the stream of consciousness, the outlines of a general framework for understanding conscious processes should be discernable (see Newman, 1997 for a fuller account). This is certainly great progress, given the virtual *terra incognita* consciousness has been for most of the history of science.

REFERENCES

Baars, B. J. (1983). How does a serial, integrated and very limited stream of consciousness emerge out of a nervous system that is mostly unconscious, distributed, and of enormous capacity? In G. R. Brock & J. Marsh (Eds), *CIBA Symposium on Experimental and Theoretical Studies of Consciousness* (pp. 282–290). London: John Wiley and Sons.

Baars, B. J. (1988). *A cognitive theory of consciousness.* Cambridge: Cambridge University Press.

Baars, B. J. (1992). *Experimental slips and human error: Exploring the architecture of volition.* New York: Plenum Press.

Baars, B. J. (1994). A global workspace theory of conscious experience. Baars, B. J. and Newman, J. (1994). A neuro-biological interpretation of a Global Workspace theory of consciousness. In A. Revonsuo and M. Kamppinen (Eds), *Consciousness in philosophy and cognitive neuroscience.* Hillsdale, NJ: Erlbaum.

Baars, B. J. (1997). *In the theatre of consciousness: The workspace of the mind.* Oxford: Oxford University Press.

Baars, B. J., Newman, J. & Taylor, J. G. (in press). Neuronal mechanisms of consciousness: A Relational Global Workspace framework. In S. Hameroff, A. Kaszniak, J. Laukes, & A. Scott (Eds), *Towards a science of consciousness: The second Tucson discussion and debates.* Cambridge, MA: MIT Press.

Buser, P., & Rougeul-Buser, A. (1995). Do cortical and thalamic bioelectric oscillations have a functional role? A brief survey. *Journal of Physiology (Paris), 89,* 249–254.

Cho, S.-B., & Kim, J. H. (1995). Multiple network fusion using fuzzy logic. *IEEE Trans. Neural Networks, 6,* 497–501.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences, USA, 81,* 4586–4590.

Crick, F., & Koch, C. (1990a). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences, 2,* 263–275.

Crick, F., & Koch, C. (1990b). Some reflections on visual awareness. *Cold Spring Harbor Symposium on Quantitative Biology, 15,* 953–962.

Damasio, A. R. (1994). *Descartes' Error.* New York: G.P. Putnam Sons.

Deschenes, M., Madariage-Domich, A., & Steriade, M. (1985). Dendrodendritic synapses in the cat reticularis thalami nucleus: A structural basis for thalamic spindle synchronization. *Brain Research, 334,* 165–168.

Durfee, E.H. (1993). Cooperative distributed problem solving between (and within) intelligent agents. In P. Rudomin *et al.* (Eds), *Neuroscience: From neural networks to artificial intelligence* (pp. 84–98). Heidelberg: Springer-Verlag.

Eccles, J. C. (1966). *Brain and conscious experience.* Heidelberg: Springer-Verlag.

Edelman, G. M. (1989). *The remembered present, a biological theory of consciousness.* New York: Basic Books.

Funke, K., & Eyse, U. T. (1992). EEG-dependent modulation of response dynamics of cat dLGN relay cells and the contribution of corticogeniculate feedback. *Brain Research, 573,* 217–227.

Fuster, J. M. (1980). *The prefrontal cortex.* New York: Raven Press.

Gazzanigga, M. S. (1985). *The social brain, discovering the networks of the mind.* New York: Basic Books.

Goldman-Rakic, P. S. (1988a). Changing concepts of cortical connectivity: parallel distributed cortical networks. In P. Rakic and W. Singer (Eds), *Neurobiology of the cortex* (pp. 177–202). Berlin: John Wiley and Sons Ltd.

Goldman-Rakic, P. S. (1988b). The prefrontal contribution to working memory and conscious experience. In O. Creutzfeld and J. Eccles (Eds), *The brain and conscious experience.* Rome: Pontifical Academy.

Groenewegen, H. J., & Berendse, H. W. (1994). The specificity of the 'nonspecific' midline and intralaminar thalamic nuclei. *Trends in Neuroscience, 4*(2), 52–58.

Grossberg, S., & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks, 4,* 452–466.

Hampshire II, J.B., & Waibel, A. (1992). The Meta-Pi network: Building distributed knowledge representations for robust multisource pattern recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence, 14,* 751–769.

Harnad, S. (1994). Guest editorial — Why and how we are not zombies. *Journal of Consciousness Studies. 1*(2), 164–168.

Heilman, K. M., Watson, R. T. & Valenstein, E. V. (1985). Neglect and related disorders. In K.M. Heilman & E.V. Valenstein (Eds), *Clinical neuropsychology.* New York: Oxford University Press.

Jackendorf, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation, 3,* 79–87.

Jacobs, R. A., & Jordan, M. I. (1993). Learning piecewise control strategies in a modular neural network architecture. *IEEE Trans. Systems, Man, and Cybernetics, 23,* 337–345.

Jasper, H. H. (1960). Unspecific thalamocortical relations. In J. Field, H. W. Magoun & V. E. Hall (Eds), *Handbook of neurophysiology, Vol. 1* (pp. 1307–1322). Washington, DC: American Physiological Society.

Johnson-Laird, P. N. (1988). *The computer and the mind.* Cambridge, MA: Harvard University Press.

Jones, E. G. (1985). *The thalamus.* New York: Plenum Press.

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science, 237,* 285–292.

Koch, C. (1997). Computation and the single neuron. *Nature, 385,* 207–210.

Koch, C., & Poggio, T. (1992). Multiplying with synapses and neurons. In T. McKenna, J. Davis & S. F. Zornetzer (Eds), *Single neuron computation* (pp. 315–345). Boston, MA: Academic Press.

Krosigk, von M., Bal, T., & McCormack, D. (1993). Cellular mechanisms of a synchronized oscillation in the thalamus. *Science, 261,* 361–364.

LaBerge, D. L. (1990). William James symposium: Attention. *Psychological Science, 1*(3), 156–162.

LaBerge, D. L. (1995). *Attentional processing: The brain's art of mindfulness.* Cambridge, MA: Harvard University Press.

Llinas, R. R., & Pare, D. (1991). Commentary: of dreaming and wakefulness. *Neuroscience, 44*(3), 521–535.

Llinas, R., Ribary, U. Johot, M. & Wang, X-J. (1994). Content and context in temporal thalamocortical binding. In G. Busaki *et al.* (Eds), *Temporal coding in the brain* (pp. 251–272). Heidelberg: Springer-Verlag.

McClelland, J. L. (1985). Putting knowledge in its place: A scheme for

programming parallel processing structures on the fly. *Cognitive Science, 9,* 113–146.

McClelland, J. L. (1986). The programmable blackboard model of reading. In J. L. McClelland & D. E. Rummelhart (Eds), *Parallel distributed processing,* Vol. 2 (pp. 122–169). Cambridge, MA: MIT Press.

McCormick, D. A., & Krosigk, M. (1992). Corticothalamic activation modulates thalamic firing through glutamate metabotropic receptors. *Proceedings of the National Academy of Science USA, 89,* 2774–2778.

Mel, B. W. (1994). Information processing in dendritic trees. *Neural Computation, 6,* 1031–1085.

Mesulam, M. (1985). *Principles of behavioral neurology.* Philadelphia: F.A. Davis.

Minsky, M. (1979). The society theory. In P. H. Winston & R. H. Brown (Eds), *Artificial intelligence, an MIT perspective, Vol. 1* (pp. 423–450). Cambridge, MA: MIT Press.

Mitrofanis, J., & Guillery, R. W. (1993). New views of the thalamic reticular nucleus in the adult and developing brain. *Trends in Neuroscience, 16,* 240–245.

Monchi, O. & Taylor, J. G. (1995). A model of the prefrontal loop that includes the basal ganglia in solving a recency task, *Proceedings of the International Neural Network Society Annual Meeting,* July 1995. Washington, DC: International Neural Network Society Press.

Newell, A. (1992). SOAR as a unified theory of cognition: Issues and explanations. *Behavioral and Brain Sciences, 15*(3), 464–492.

Newman, J. (1995a). Review: Thalamic contributions to attention and consciousness. *Consciousness and Cognition, 4*(2), 172–193.

Newman, J. (1995b). Reticular-thalamic activation of the cortex generates conscious contents. *Behavioral and Brain Sciences, 18*(4), 691–692.

Newman, J. (1997). Putting the puzzle together: Towards a general theory of the neural correlates of consciousness. *Journal of Consciousness Studies, 4* (1 and 2), 47–66 and 99–120.

Newman, J., & Baars, B. J. (1993). A neural attentional model for access to consciousness: A Global Workspace perspective. *Concepts in Neuroscience, 4*(2), 255–290.

Newman, J., Baars, B. J. & Cho, S-B. (1997). A neurocognitive model for attention and consciousness. In S. O'Nuallain, P. McKevitt & E. MacAogdin (Eds), *Two sciences of mind: Readings in cognitive science and consciousness.* Philadelphia, PA: John Benjamins of North America.

O'Nuallain, S., McKevit, P. & MacAogdin, E. (1997). *Two sciences of mind: Readings in cognitive science and consciousness.* Philadelphia, PA: John Benjamins of North America.

Parent, A., & Hazrati, L.-N. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia–thalamo–cortical loop. *Brain Research Reviews, 20,* 91–127.

Penfield, W. (1975). *The mystery of the mind: A critical study of consciousness and the human brain.* Princeton, NJ: Princeton University Press.

Penrose, R. (1994). *Shadows of the mind—In search of the missing science of consciousness.* Oxford: Oxford University Press.

Posner, M. I. & Rothbart, M. K. (1991). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds). *The neuropsychology of consciousness* (pp. 11–34). London: Academic Press.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Science USA, 91,* 7398–7403.

Reddy, D. R., Erman, L. D., Fennell, R. D., & Neely, R. B. (1973). The Hearsay speech understanding system: An example of the recognition process. *Proceedings of the International Conference on Artificial Intelligence,* 185–194.

Rummelhart, D. E., Hinton, G. E. & McClelland, J. L. (1986). A general framework for parallel distributed processing. In J. L. McClelland & D. E. Rummelhart (Eds). *Parallel distributed processing,* Vol. 1 (pp. 43–76). Cambridge, MA: MIT Press.

Scheibel, M. E., & Scheibel, A. B. (1966). The organization of the nucleus reticularis: A Golgi study. *Brain Research, 1,* 43–62.

Scheibel, A. B. (1980). Anatomical and physiological substrates of arousal: A view from the bridge. In J. A. Hobson & M. A. B. Brazier (Eds), *The reticular formation revisited* (pp. 55–66). New York: Raven Press.

Sillito, A., Jones, H., Gerstein, G., & West, D. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature, 369,* 479–482.

Skinner, J. E. & Yingling, C. D. (1977). Central gating mechanisms that regulate event-related potentials and behavior. In J. E. Desmedt (Ed.), *Progress in clinical neurophysiology: Attention, voluntary contraction and event-related cerebral potentials, Vol. 1* (pp. 30–69). Basel: Karger.

Sompolinsky, H., Golomb, D., & Kleinfeld, D. (1991). Cooperative dynamics in visual processing. *Physical Review A, 43*(12), 6990–7011.

Steriade, M., & Llinas, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. *Physiological Reviews, 68*(3), 649–742.

Stryker, M. P. (1989). Is grandmother an oscillation? *Nature, 338,* 297–337.

Taylor, J. G. (1992). Towards a neural network model of the mind. *Neural Network World, 2,* 797–812.

Taylor, J. G. (1996). A competition for consciousness? *Neurocomputing, 11,* 271–296.

Taylor, J. G. & Alavi, F. N. (1993). Mathematical analysis of a competitive network for attention. In J. G. Taylor (Ed.), *Mathematical approaches to neural networks* (pp. 341–382). Amsterdam: Elsevier Science Publishers B.V.

Taylor, J. G. & Michalis, L. (1995). The functional role of the hippocampus in the organization of memory. *Proceedings of the International Neural Network Society Annual Meeting,* July 1995. Washington, DC: International Neural Network Society Press.

Vibert, J., Pakdaman, K., & Azmy, N. (1994). Interneural delay modification synchronizes biologically plausible neural networks. *Neural Networks, 7,* 589–607.

Walshe, K. W. (1978). *Neuropsychology, A clinical approach.* Edinburgh: Churchill Livingstone.

Yao, Y., & Freeman, W.J. (1990). Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks, 3,* 153–170.

Yen, C. T., Conely, M., Hendry, S. H. C., & Jones, E. G. (1985). The morphology of physiologically indentified GABAergic neurons in the somatic sensory part of the thalamic reticular nucleus in the cat. *Journal of Neuroscience, 5,* 2254–2268