Published in final edited form as:

Methods. 2008 April ; 44(4): 338–347. doi:10.1016/j.ymeth.2007.02.007.

A Link Between Neuroscience and Informatics: Large-Scale Modeling of Memory Processes

Barry Horwitz* and Jason F. Smith

Brain Imaging & Modeling Section, National Institute on Deafness and Other Communications Disorders, National Institutes of Health, Bethesda, MD USA

Abstract

Utilizing advances in functional neuroimaging and computational neural modeling, neuroscientists have increasingly sought to investigate how distributed networks, composed of functionally defined subregions, combine to produce cognition. Large-scale, biologically realistic neural models, which integrate data from cellular, regional, whole brain, and behavioral sources, delineate specific hypotheses about how these interacting neural populations might carry out high-level cognitive tasks. In this review, we discuss neuroimaging, neural modeling, and the utility of large-scale biologically realistic models using modeling of short-term memory as an example. We present a sketch of the data regarding the neural basis of short-term memory from non-human electrophysiological, computational and neuroimaging perspectives, highlighting the multiple interacting brain regions believed to be involved. Through a review of several efforts, including our own, to combine neural modeling and neuroimaging data, we argue that large scale neural models provide specific advantages in understanding the distributed networks underlying cognition and behavior.

1. Introduction

Traditionally, investigations into the neural basis of human cognition have primarily focused on the localization of function [1]. The study of deficits in brain-damaged subjects (both human and non-human, e.g., [2,3]) and neurophysiological and neuroanatomical studies in non-human primate and other mammalian preparations (e.g., [4–6]) have provided detailed information regarding potential neural loci of functionally defined modules of cognition. In recent years, two new complimentary techniques have been developed which have revolutionized not only the fundamental nature of the data available to neuroscientists studying the relationship between brain and behavior, but also the nature of the questions they ask, the interpretations of data that they make, and ultimately the theories they construct regarding human cognitive function. The two new techniques are functional neuroimaging and computational neural modeling. Each of these techniques has made its own unique contribution to the field of cognitive neuroscience and have been reviewed elsewhere (c.f., [7] for the contributions of neuroimaging and [8,9] for the impact of computational neural modeling).

Modern functional brain imaging tools such as functional magnetic resonance imaging (fMRI) and electro- and magnetoencephalography (EEG/MEG) can capture vast quantities of neurobiologically relevant data. For many human cognitive functions (e.g., language),

*Address for correspondence: Dr. Barry Horwitz, Bldg. 10, Rm. 8S235, MSC 1407, NIDCD, NIH, Bethesda, MD 20892, USA, Tel. 1-301-594-7755, Fax 1-301-480-5625, Email horwitz@helix.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

neuroimaging data often represent the only neural data one can obtain in normal populations. Uniquely, these imaging techniques acquire data simultaneously from most of the brain over multiple points in time. Thus, rather than having information from a single neuron or a single brain region, about which the obvious questions often reduce to “what does this bit do?”, functional brain imaging permits the investigation of multiple interacting brain elements whose combined and orchestrated neural activity implements various cognitive functions. That is, these methods allow researchers to investigate the neuronal networks of cognition.

Likewise, neural modeling and the study of artificial neural networks have dramatically increased our understanding of how simple, unintelligent computational units can combine to collectively produce intelligent-like behavior. Neural models have begun to lighten the theoretical “black boxes” underlying a wide range of behaviors from classical conditioning [10,11] to visual object recognition [12] to reading [13] and have provided a new language for describing and theorizing about normal [14] and disordered [15,16] cognition.

Here, we argue that the values of neuroimaging and neural modeling are increased when they are combined in the form of large-scale neural models where neuroimaging data (along with data from other sources) are used to constrain biologically realistic neural models that in turn assist in the interpretation of the imaging data. Such models can be used to bridge the different levels of neuroscientific analysis and ultimately form detailed, fully specified theories of cognitive functions [17]. In this article, we will discuss neuroimaging, neural modeling, and the utility of large-scale biologically realistic models in the context of short-term memory. We first present a brief review of short-term memory, highlighting the multiple brain regions involved. This suggests that it is a prime candidate for analysis at a network level. We discuss how short term memory processes have been implemented in neural network studies, and then go on to examine neuroimaging studies of short-term memory with a focus on systems-level network analyses. We argue that large scale neural models provide an additional means of performing network analysis of imaging data with specific advantages. Finally, we review several efforts, including our own, to combine large-scale neural modeling and neuroimaging in the study of short-term memory.

2. Memory Processes –A Brief Overview

Memory, the ability to retain information regarding the past to influence future actions, is a central component of cognition and intelligence. Because of its importance, it has been the subject of study at a number of different levels of analysis, each with its own specific questions, research paradigms, and foci. At the cognitive or behavioral level, theorists have long considered memory as consisting of multiple sub-types. According to one commonly accepted taxonomy of memory based primarily on neuropsychological evidence, primate memory is divided into declarative and non-declarative systems [18]. The non-declarative system encompasses memory processes for which the contents are not directly available for conscious inspection, and includes procedural memory, perceptual priming, classical and operant conditioning, and nonassociative forms of learning. Declarative memory, which will be the focus here, consists of episodic (i.e., memory for specific events or points in time), and semantic (i.e., memory for factual knowledge independent of the encoding event) components.

A useful classification of declarative memory fractionates the system into three constituent (though not necessarily mutually exclusive) processes based upon temporal characteristics: sensory, short-term, and long-term memory. Sensory memory (e.g., Iconic or Echoic memory) is a very short-term store temporally limited to approximately one second or less. Sensory memory traces are unimodal, directly associated with or dependent upon a specific sensory modality and brain regions associated with perception in that modality. Short-term memory (or working memory) is of longer duration, on the order of a few seconds. This is the process

associated with “keeping something in mind” for a short period, such as remembering a phone number for just long enough to dial it correctly. Though the term working memory is often used to indicate additional executive or transformational operations being performed on the contents of short-term memory stores [19,20], here we will use the terms interchangeably. Short-term memory is often considered an interface between sensory/motor processes and long-term memory - passing information to and receiving information from long-term memory stores. The duration of long-term memory is measured using units from minutes to years and contains the accumulated knowledge of facts and past events that are most directly associated with memory in the colloquial sense.

An effort has been made recently to map the different memory components identified at the cognitive/behavioral level, such as short and long-term memory, onto a neural systems level architecture. The neural correlates of short-term memory have been well studied in nonhuman primates using various delayed response paradigms. Early electrophysiological studies of awake, behaving primates performing delayed response tasks clearly demonstrated increased neural activity in prefrontal cortex (PFC) in the period between stimulus presentation and response [21,22]. This delay period activity in PFC is associated with mnemonic maintenance and is predictive of response accuracy in both spatial (e.g., [23]) and object (e.g., [24,25]) matching tasks, and is now believed to be the neural basis for short-term memory [26] (see [27] for a review). Sustained alterations of delay period activity have also been observed in the thalamus [28], parietal cortex [29,30], and the inferior temporal lobes [25,31–34]. The existence of multiple regions exhibiting behaviorally relevant delay period activity has led some researchers to posit that short-term memory is itself dependent upon the interactions among these regions (c.f., [35]). However, task related inferior temporal delay activity can be disrupted by distractor stimuli without reduction in task accuracy [25,34], as can parietal delay activity [30], suggesting that while this activity potentially plays a role in normal short-term memory processes, not all of the activity in these regions is strictly necessary for short-term memory maintenance.

Long-term memory has been most directly associated with the medial temporal lobes, including the hippocampus, and the parahippocampal, perirhinal and entorhinal cortices [36]. This association is most clearly demonstrated by patients such as HM [37] whose temporal lobectomy resulted in a profound anterograde amnesia, as well as a temporally graded retrograde amnesia. Development of an animal model of human amnesia has greatly contributed to our understanding of the medial temporal memory system, resulting in a fairly detailed description of the structure and function of its constituent elements (see [38–40] for reviews). A review of long-term memory is beyond the scope of the current discussion. However, of interest here are the potential interactions between short and long-term memory systems. The delayed paired associates memory task, which involves both short and long-term memory processes, has been well studied in nonhuman primates, yielding important physiological data for understanding the interactions between these two systems (see [41] for a review). In this task, individual neurons in both inferior temporal and prefrontal cortex that have been shown to respond selectively to a single stimulus exhibit a gradually increasing response to a different stimulus when the different stimulus is the pair of the neuron’s preferred stimulus [42,43]. Neurons coding for the cue stimulus alone and for both items of an associated pair are also observed [44–46]. The association related activity in the inferior temporal cortex is dependent upon feedback signals from intact medial temporal structures including perirhinal and entorhinal cortex [45,47] and is resistant to distractor stimuli that disrupt delay activity in this region in working memory tasks [34]. Thus, short-term maintenance of a representation derived from long-term memory (such as in the delayed paired associate task) is likely performed by an augmented or even separate memory circuit than the circuit supporting maintenance of representations derived directly from the environment. That prefrontal cortex, medial temporal cortex and inferior temporal cortex all code the response stimulus during the

delay period and that these three regions are strongly interconnected suggests that they form a functional network for paired associates memory [48].

3. Computational Modeling and Neuroimaging of Memory

Computational modeling has been used extensively to examine memory function. The kinds of models employed have ranged from those that incorporate neurobiological details to those that are cognitively based but whose relationship to an underlying neural substrate is unspecified. Included in the latter case are classical connectionist models for which an existing pattern of activity across a layer or cluster of nodes is considered a short-term memory while the connection weight matrix holds long-term memories of the relationships between items [14,49–51]. While these “biologically inspired” networks have greatly increased our understanding of computational processes in distributed systems of unintelligent processors, the ability to connect these network models to actual computations occurring in the brain has been limited.

Focusing on short-term memory, a number of researchers have attempted to develop computational models of delay activity in prefrontal neurons using a variety of methods (see [52] for a methodological review). Delay period activity can be achieved using single bistable neurons, simple recurrently connected networks, or “synfire” chains of multiple feedforward neurons that form a closed loop [52]. Bistable neurons switch between low and high firing rate stable states, generating persistent activity independent of synaptic connectivity [53]. The utility of synfire chains for short-term memory has only recently begun to be explored; however, their ability to explain observed firing sequences makes them a strong candidate for further study [54]. Recurrent networks are by far the most studied system capable of producing persistent activity. In response to input, the dynamics of the recurrent connections (the outputs of a set of neurons are fed back as additional inputs to the set) in these networks cause the networks to be drawn toward (discrete or continuous) attractor states which can then be maintained without further input [51,55–61]. Recurrent networks have been used to maintain short-term representations in a variety of applications including memory for spatial locations (e.g., [12,62,63]), memory for objects (e.g., [64]) and memory for task demands [65], and can be constructed by connecting model neurons within (e.g., [66]) or between ([35]) brain regions.

There have been a enormous number of human functional neuroimaging studies of short-term and long-term memory (reviews can be found in [67,68] and in several chapters in [69]). In general, there is much agreement between the neuroimaging results and the findings from non-human electrophysiological studies. Focusing on short-term/working memory, functional neuroimaging studies consistently demonstrate activations in prefrontal and posterior parietal cortex (e.g., [67,70–72]). Smith and Jonides [67] have argued that the posterior parietal cortex plays a critical role in the storage of the item in short-term memory, whereas the PFC subserves a maintenance/rehearsal function. Curtis and D’Esposito [73] have proposed that dorsolateral PFC aids in information maintenance by directing attention to internal representations of sensory stimuli and motor plans that are stored in more posterior areas of the brain.

Because functionally defined cognitive processes, such as short-term memory, often involve multiple brain regions [74], efforts have been made to exploit the massive number of simultaneous data points though out the brain available in neuroimaging to analyze the networks supporting these processes. Initially, this effort was centered on the development of imaging data analysis methods related to the evaluation of various measures of interregional interactions (i.e., functional and effective connectivity [75–80]). Beginning in the early-to-mid 90s, investigators began applying various types of computational modeling to brain hemodynamic/metabolic data acquired in humans by means of positron emission tomography (PET) and in nonhuman animals using autoradiographic methods. One such approach that

focused on the interaction between various brain regions employed structural equation modeling (SEM) [79,81,82]. This type of modeling, which we have called a kind of systems-level neural modeling [83], also was applied to fMRI data [84], and is now an important technique used in many fMRI investigations (e.g., [85–88]). Recently, similar approaches using more sophisticated techniques, such as Dynamic Causal Modeling (DCM) [89–91], have further developed the system-level modeling approach. The application of SEM in the analysis of neuroimaging of short-term memory is reviewed by [92]. Numerous interactions have been observed in both normal and patient populations, including altered effective connectivity under increasing memory load [93,94], highlighting the distributed nature of short-term memory (see [95] for similar results using functional connectivity).

However, one problem with systems-level neural modeling is that it is largely based on regional hemodynamic data. This means that its spatial and temporal resolutions are limited to those associated with fMRI (or PET, whose spatiotemporal resolution is even more restricted): a spatial scale on the order of 5–10 millimeters and a temporal scale corresponding to a few seconds. Even the smallest element of an fMRI data set (a voxel) contains multiple and diverse neuronal populations. Moreover, the hemodynamic response function that results in the measured fMRI signal spatially and temporally filters the underlying neural activity, further reducing resolution. This smoothing of the neural activity reduces our ability to determine the neural level computations underlying the regional signals and the consequences of altered inter-regional connectivity on these computations. To overcome the spatiotemporal limitations of functional neuroimaging data, another type of computation modeling – large-scale neural modeling – was developed [66,96,97].

It should be mentioned that the use of EEG/MEG data does yield signals with an excellent temporal resolution (on the order of milliseconds) even though such data provide information based on the activity of neural populations; however, the spatial resolution of such data is limited (see [98]). Nonetheless, there have been efforts in the past to enhance the neuroscientific understanding of such data by using computational neural modeling (e.g., [99–103]) and a resurgence of this approach has taken place in the last few years (e.g., [104–106]).

4. Biologically Realistic Modeling and Functional Neuroimaging

In this section, we discuss the use of neurobiologically realistic models to help account for functional brain imaging data, specifically PET and fMRI data. Another way to think about this is that we discuss the use of functional neuroimaging data to help constrain neural models that purport to provide neurally based mechanisms for sensorimotor and cognitive processes. We first motivate the use of biologically realistic computational models in neuroimaging and then provide specific examples of how these models are being used in the study of short term memory. For reviews of this type of modeling, see [17,83,98,107].

4.1 Large Scale Neurocomputational Models

Four conceptually distinct purposes (which can blur into one another in actual applications) can be served by this type of modeling [108]: (1) formulating and implementing specific hypotheses about how neuronal populations mediate a task; this is called forward modeling, and it is perhaps the most important use for such modeling; (2) determining how well an experimental design paradigm or analysis method works; (3) investigating the meaning in neural terms of a macro-level concept (e.g., functional connectivity); (4) combining different types of data with one another (e.g., fMRI and MEG data).

The central reason for employing large-scale modeling with functional neuroimaging data is that it provides a method to formulate and delineate specific hypotheses about how interacting neural populations carry out high-level cognitive and sensorimotor tasks in humans, and to

generate simulated data based on these hypotheses that can be compared directly to experimental results, especially those acquired using functional neuroimaging and behavioral measures. In essence, this approach allows neuroscientists to bridge the different spatiotemporal scales at which neural data are obtained. The spatial scales range from the brain as a whole to subcellular and molecular dimensions, and the temporal dimensions go from days, months, years to submillisecond intervals. No one dataset transcends all the different levels, and so interpreting spatiotemporally disparate data in terms of a single, unified account relating specific behaviors to their underlying neural mechanisms has been a challenge for the cognitive neuroscience community. We have argued that large-scale neural modeling provides a way to attempt to bridge some of these multiple levels [17]. It is in this sense – simultaneously accounting for neural data across multiple spatiotemporal dimensions – that neural modeling can provide a link between neuroscience and informatics. Hemodynamic-based functional neuroimaging affords information concerning brain location, but does not directly allow us to infer what is being computed by the neuronal populations – that is, functional neuroimaging does not provide direct evidence concerning the neural mechanisms underlying the cognitive function under study. Neural modeling can do this in the sense that in constructing a model, one is forced to make specific hypotheses about the neural mechanisms that will enable the model to perform the task of interest. These hypotheses concerning the neural mechanisms can then be tested by comparing simulated with experimental data.

4.2 Short-Term Memory for Objects

As a first example, we will discuss our laboratory's research investigating the neural substrates of object perception – both visual and auditory. Like visual objects (such as tables, chairs, people), auditory objects can be thought of as perceptual entities subject to figure-ground separation [109]. Besides definable environmental sounds, humans are particularly interested in words and musical patterns, and thus the number of auditory objects we have the ability to distinguish is in the hundreds of thousands.

The starting point for thinking about this problem was to examine the similarities between auditory and visual object processing. Both begin as signals at the receptor surface (retina for vision, cochlea for audition). A large amount of lower brain level processing occurs, resulting in the construction at the level of primary auditory or visual cortex of what we will somewhat loosely call a percept (for vision, some of the lower level processing takes place in the retina itself). Neural processing continues onto higher cortical areas whose activities, at least in primates, result in integrating the percept with other aspects of the world: the percept is transformed into a concept. Understanding the processing of auditory objects at the higher levels is critical for understanding a number of important cognitive functions in humans, especially speech perception.

There are multiple areas in the primate brain that respond to visual stimuli. As first proposed by Ungerleider and Mishkin [110], these areas form primarily two parallel and hierarchically arranged pathways that start in primary visual cortex. One pathway includes regions in ventral occipital, temporal and frontal cortex, and appears to be concerned with processing object features such as form and color. The other major pathway starts in occipital cortex and extends dorsally into parietal cortex and thence onto dorsal frontal cortex. Neurons in these areas seem to be engaged in processing the location of objects in space. Recent investigations in primates, including humans, have given rise to the hypothesis, proposed by Kaas [111], Rauschecker [112] and others [6], that like the visual system, auditory areas in the cerebral cortex contain at least two primary processing pathways – a ventral stream running from primary auditory cortex anteriorly along the superior temporal gyrus that is associated with processing the features of auditory objects, and a dorsal stream that goes into the parietal lobe that is concerned with the spatial location of the auditory input. This notion, however, is more controversial

(especially the part concerning the dorsal pathway being associated with spatial processing), and lacks strong experimental support. It was, nevertheless, the starting point for our work, which focused on the less controversial object processing pathway.

We developed two models – one for visual object processing [66,113] and one for auditory object processing [114]. Both models perform a short term memory task (the delayed match-to-sample (DMS) task), in which a stimulus is presented briefly, there is a delay period during which the stimulus is kept in short-term memory, a second stimulus is presented, and the model decides if the second stimulus is the same as the first. For the visual model, the stimuli consisted of simple geometric shapes (e.g., squares, tees), whereas for the auditory model, the stimuli were simple tonal patterns (e.g., combinations of frequency sweeps and pure tones).

The visual model [66] incorporates four major brain regions representing the ventral object processing stream [110]: (1) primary sensory cortex (V1/V2); (2) secondary sensory cortex (V4); (3) a perceptual integration region (inferior temporal (IT) cortex); and (4) prefrontal cortex (PFC), which plays a central role in short-term working memory (Fig. 1). Every region contains multiple excitatory-inhibitory units (modified Wilson-Cowan [55] units) each of which represents a cortical column. Feedforward and feedback connections link neighboring regions. Based on the experimental observation that the spatial receptive field of a neuron increases as one progresses from primary visual cortex to higher-level areas [115], there are different scales of spatial integration in the first three stages of the model, with the primary sensory region having the smallest spatial receptive field and IT the largest. Model parameters were chosen so that the excitatory elements have simulated neuronal activities resembling those found in electrophysiological recordings from monkeys performing similar tasks (e.g., [33, 116]).

The model for auditory object processing was constructed in a manner analogous to the visual model (consult Husain et al. [114] for details and parameter values). The modules that were included were primary sensory cortex (Ai), secondary sensory cortex (Aii), a perceptual integration region (superior temporal cortex/sulcus, ST), and a prefrontal module (PFC) essentially identical to that used in the visual model (see Fig. 1). As with the visual model, there were feedforward and feedback connections between modules. Because auditory stimuli are perceived over time, rather than space, and can have a wide spectral range, the neurons in the auditory model were posited to have spectrotemporal receptive fields that become larger as one goes from Ai to Aii to ST. Stimuli are presented to an area of the model corresponding to the lateral geniculate nucleus (LGN) for the visual case or the medial geniculate nucleus (MGN) for the auditory case.

There is also a biasing (or attention) signal that is used to tell the models which task to perform: the DMS task, or a sensory control task that requires only sensory processing but no retention in short-term memory. This biasing variable modulates a specific subset of prefrontal units via diffuse synaptic inputs, the functional strength of which controls whether the stimuli are to be retained in short-term memory or not [117]. Activity in each brain area, therefore, is some combination of feedforward activity determined in part by the presence of an input stimulus, feedback activity determined in part by the strength of the modulatory bias signal, and local activity within each region.

Besides simulating neuronal activity, the models allow for the simulation of PET/fMRI data, which is accomplished by temporally and spatially integrating the absolute value of the synaptic activity in each region over an appropriate time course. For simulating fMRI, these values are convolved with a function representing the hemodynamic delay [113]. Details about the parameters used in the two models, and a thorough discussion of all the assumptions employed, are given in Tagamets and Horwitz [66,117] and in Husain et al. [114].

Following presentation of the initial stimulus of a single DMS trial, significant neural activity occurs in all brain regions of the model. Activity in two prefrontal populations is relatively high during the delay interval, when the stimulus must be kept in short-term memory, but low-level activity occurs in all other neural populations. When the second stimulus is presented during the response portion of the task, there is an increase in neural activity in all areas, and a subpopulation in PFC responds only if the second stimulus matches the first. Consequently, both the visual and auditory simulations demonstrate that these neural models can perform the DMS task and that the simulated electrical activities in each region are similar to those observed in nonhuman mammalian electrophysiological studies.

For the visual model the simulated functional neuroimaging data were compared to PET regional cerebral blood flow (rCBF) values for a short-term memory task for faces (the original experimental data came from Haxby et al. [71]). The control tasks consisted of passive viewing of scrambled shapes for the model and a nonsense pattern for the experiment. For the auditory simulation, identical stimuli (tonal contours, each consisting of two frequency sweeps, separated by a constant tone; each tonal contour was 350 msec in duration; the delay period in the DMS task was 1000 msec long) were used for both modeling and experiment [114]. In the auditory case, the control task consisted of performing the DMS task using pure tones. Fig. 2 shows the quantitative comparisons for both the visual and the auditory simulations. The simulated PET values for the visual model of the DMS condition, when compared to the control condition, were similar [66] to those found in experimental PET studies of face working memory [71], as shown in the bar graph on the left of Fig. 2. Similarly, the auditory model was able to generate simulated fMRI data that generally were in close agreement with the experimental data. The bar graph on the right of Fig. 2 shows the percent signal changes (comparing the fMRI activity for tonal contours to that for pure tones) in each brain region for both the simulated and the experimental data. Although our simulated results in primary auditory cortex (Ai) did not match the experimental value (in the simulation, the percent change between tonal contours and tones was near zero), we were able to obtain close quantitative agreement between simulated and experimental data in all the other right hemisphere regions that corresponded to those in the model. A likely reason for the lack of agreement in the primary auditory area is that we included in our model only one type of neuron found in primary auditory cortex (selective for frequency sweeps), even though there are many neuronal types in the brain selective for other features in the auditory input (e.g., loudness, on and off properties of the input) that were not included in the model. Moreover, there was a large amount of scanner noise during the experiment that could have had some effect on the experimental data that was not taken into account in the simulation. Nevertheless, this was the first study in which a biologically realistic neural model generated stimulated fMRI data that generally agreed *quantitatively* with experimental fMRI values in which task design and stimuli were identical to those used in the modeling.

To test the auditory model, intact stimuli were matched with fragmented versions (i.e., with inserted silent gaps) [118]. The ability of the model to match fragmented stimuli declined as the duration of the gaps increased. However, when simulated broadband noise was inserted into these gaps, the matching response was restored, indicating that a continuous stimulus was perceived. This perceptual grouping phenomenon is called the auditory continuity illusion, and has a rich behavioral literature. Husain et al [118]. found that the simulated electrical activities of the neuronal units of the model agreed with published electrophysiological data, and the behavioral activity of the model matched published human behavioral data (see [118] for details). In the model, the predominant mechanism implementing the auditory continuity illusion is the divergence of the feedforward connections along the auditory processing pathway in the temporal cortex.. It is noteworthy that no parameters of the auditory model were changed to simulate the auditory continuity illusion, and thus, the results attest to the robustness of the model to explain phenomena associated with auditory object processing.

4.3. Spatial Working Memory

The previous section focused on the visual and auditory object processing ventral pathways in primate neocortex. Recently, Klingberg, Tegner and colleagues have developed computational models that relate neural to fMRI data for tasks involving visual spatial working memory (vsWM) [119,120]. These tasks required subjects to maintain in mind during a delay period the location(s) on the circumference of a circle of one or more consecutively presented visual cues. Humans can retain the location of at least 4 items at the same time [121]. Unlike the models of Tagamets, Husain, Horwitz and colleagues discussed above [66,114], which used non-spiking, leaky integrator neurons, the neurons employed to model vsWM were spiking (integrate-and-fire) neurons [63], with separate compartments for soma, proximal and distal dendrites. The modeled neurons included both excitatory pyramidal neurons (having both AMPA and NMDA mediated synaptic connections) and inhibitory neurons (fast-spiking with GABA-A synaptic connectivity). These neuronal populations had all-to-all connectivity, both within and between populations. The network model maintained memories during the delay period as states of persistent neural activity that were initiated by cues which induced transient activity in a few adjacent neurons in the network, in agreement with electrophysiological recordings in primate prefrontal cortex.

One problem addressed focused on multi-item vsWM [119]. Although biophysically based computational models have successfully accounted for the persistent neural activity during vsWM for single items (e.g., [63,122], these types of recurrent networks have had difficulty maintaining more than two items simultaneously in working memory. To address this problem, Macoveanu et al. extended the Tegner et al. model [63] by implementing cellular mechanisms known to occur during the childhood development of working memory, such as an increased synaptic strength and improved contrast and specificity of the neural response. Their computational study showed that these mechanisms were sufficient to create a neural network which was able to store information about multiple items through sustained neural activity, in agreement with behavioral experimental data [121]. Importantly, by using fMRI, they found that the information-activity curve predicted by the model corresponded to that in the human posterior parietal cortex during performance of vsWM tasks.

The recurrent network model just discussed represents activity in a single brain region – in some studies it was a single area of the dorsolateral prefrontal cortex (PFC) [63], in the Macoveanu et al. study [119] it was posterior parietal cortex. Human functional brain imaging studies of vsWM studies in humans [67,123,124] and monkeys [125], however, have found sustained delay activity associated with vsWM most consistently in the superior frontal sulcus (SFS) and the intraparietal sulcus (IPS). Extending the single-region Tegner et al model [63] into a two-region model, Edin et al. [120] investigated the consequences of cellular maturational processes, including myelination, synaptic strengthening and synaptic pruning, on vsWM-related brain activity and behavioral performance (vsWM improves during childhood development). They examined which maturational processes could lead to improved working memory. They implemented five structural developmental changes occurring as a result of the cellular maturational processes in their computational network model. The developmental changes in memory activity predicted from the simulations of the model were then compared to brain activity measured with fMRI in children and adults. Edin and colleagues determined that networks with stronger fronto-parietal synaptic connectivity between cells coding for similar stimuli, but not those with faster conduction, stronger connectivity within a region, or increased coding specificity, predicted measured developmental increases in both working memory-related brain activity and in correlations of activity between regions (i.e., in interregional functional connectivity). Stronger fronto-parietal synaptic connectivity between cells coding for similar stimuli was thus the only simulated developmental process that could

account for the observed changes in brain activity associated with development of working memory during childhood.

Although the models used to investigate visuospatial working memory are relatively simple (representing the activity in only one or two brain regions), nonetheless, the results we just presented demonstrate how one can combine modeling and functional neuroimaging to get at the neural mechanisms that may underlie developmental changes in behavior. It seems clear that similar approaches could be utilized to explore the neural underpinnings of neurological and psychiatric disorders

5. Conclusions

In this paper we provided an overview of some recent efforts to use computational neural modeling in conjunction with functional neuroimaging data. In particular, we emphasized research on human memory function. The examples we presented dealt with short-term/working memory for visual and auditory objects, and with visuospatial working memory. We showed that it was possible to account for neuroscientific data at several different levels of investigation, and as a result, the models that were discussed provide a way to bridge these levels. Interestingly, as far as we know, there have been no similar types of study concerned with long-term memory, although this is likely to change in the relatively near future.

Perhaps the most important feature of the models that were discussed was related to the notion of neural mechanism. In each case, the models provided (and implemented) specific hypotheses about the neural mechanisms underlying the cognitive phenomena under study. The models also allowed for the simulation of brain data at different levels that could be compared with experimental data. The significant point here is that there are at least three types of data that each model must successfully account for: (1) electrophysiological; (2) fMRI-derived; and (3) behavioral. Heretofore, most modeling studies dealt with one or two of these. The need to match three distinctly different types of data will impose strong constraints on any neural model, which means that one should be able to reduce dramatically the space of plausible neural mechanisms. Although this will make the modeling harder, it should lead to a stronger understanding of the neural basis of human cognitive function.

Acknowledgments

This work was supported by the NIH-NIDCD Intramural Research Program.

References

1. Gazzaniga, MS., editor. *The New Cognitive Neurosciences*. Cambridge: MIT Press; 2000.
2. Shallice, T. *From Neuropsychology to Mental Structure*. New York: Cambridge Univ. Press; 1988.
3. Murray EA, Bussey TJ. *Trends Cogn Sci* 1999;3:142–151. [PubMed: 10322468]
4. Miller EK, Li L, Desimone R. *Science* 1991;254:1377–1379. [PubMed: 1962197]
5. Funahashi S, Inoue M, Kubota K. *Beh. Brain Res* 1997;84:203–223.
6. Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. *Nat Neurosci* 1999;2:1131–1136. [PubMed: 10570492]
7. Frackowiak, RSJ.; Friston, KJ.; Frith, CD.; Dolan, RJ.; Price, CJ.; Zeki, S.; Ashburner, J.; Penny, W. *Human Brain Function*. San Diego, CA: Elsevier, Academic Press; 2004.
8. O'Reilly, RC.; Munakata, Y. *Computational Explorations in Cognitive Neuroscience*. Cambridge: MIT Press; 2000.
9. Rolls, ET.; Deco, G. *Computational Neuroscience of Vision*. Oxford: Oxford University Press; 2002.
10. Tesauro G. *Biol Cybern* 1986;55:187–200. [PubMed: 3801536]

11. Baxter, DA.; Raymond, JL.; Buonomano, DV.; Cook, DG.; Kuenzi, FM.; Carew, TJ.; Byrne, JH. Quantitative Analyses of Behavior. Volume XII: Neural Network Models of Conditioning and Action. Commons, ML.; Grossberg, S.; Staddon, JER., editors. Hillsdale, NJ: Lawrence Erlbaum and Associates; 1991. p. 13-52.
12. Stringer SM, Rolls ET. Neural Comput 2002;14:2585–2596. [PubMed: 12433291]
13. Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Psychol. Rev 1996;103:56–115. [PubMed: 8650300]
14. Rumelhart, JL. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge: MIT Press; 1986.
15. Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Psychol Rev 1997:801–838. [PubMed: 9337631]
16. Murre JMJ, Graham KS, Hodges JR. Brain 2001;124:647–675. [PubMed: 11287367]
17. Horwitz B. Comptes ren. Biol 2005;328:109–118.
18. Squire LR, Stark CE, Clark RE. Annu Rev Neurosci 2004;27:279–306. [PubMed: 15217334]
19. Baddeley, AD.; Hitch, GJ. Recent Advances in Learning and Motivation. Bower, GA., editor. vol. 8. New York: Academic Press; 1974. p. 47-90.
20. Baddeley A. J Commun Disord 2003;36:189–208. [PubMed: 12742667]
21. Kubota K, Niki H. J Neurophysiol 1971;34:337–347. [PubMed: 4997822]
22. Fuster JM, Alexander GE. Science 1971;173:652–654. [PubMed: 4998337]
23. Funahashi S, Bruce C, Goldman-Rakic PS. J. Neurophysiol 1990;63:814–831. [PubMed: 2341879]
24. Wilson FAW, Scalaidhe SPO, Goldman-Rakic PS. Science 1993;260:1955–1958. [PubMed: 8316836]
25. Miller EK, Li L, Desimone R. J Neurosci 1993;13:1460–1478. [PubMed: 8463829]
26. Sobotka S, Diltz MD, Ringo JL. J Neurophysiol 2005;93:128–136. [PubMed: 15317835]
27. Funahashi S. Neuroscience 2006;139:251–261. [PubMed: 16325345]
28. Watanabe Y, Funahashi S. J Neurophysiol 2004;92:1738–1755. [PubMed: 15140911]
29. Constantinidis C, Steinmetz MA. J Neurophysiol 1996;76:1352–1355. [PubMed: 8871242]
30. Constantinidis C, Wang XJ. Neuroscientist 2004;10:553–565. [PubMed: 15534040]
31. Fuster JM, Jervey JP. Science 1981;212:952–955. [PubMed: 7233192]
32. Miyashita Y, Chang HS. Nature 1988;331:68–70. [PubMed: 3340148]
33. Fuster JM. J. Neurophysiol 1990;64:681–697. [PubMed: 2230917]
34. Takeda M, Naya Y, Fujimichi R, Takeuchi D, Miyashita Y. Neuron 2005;48:839–848. [PubMed: 16337920]
35. Ashby FG, Ell SW, Valentin VV, Casale MB. J Cogn Neurosci 2005;17:1728–1743. [PubMed: 16269109]
36. Squire LR, Zola-Morgan S. Science 1991;253:1380–1386. [PubMed: 1896849]
37. Scoville WB, Milner B. J. Neurol. Neurosurg. Psychiatr 1957;20:11–21. [PubMed: 13406589]
38. Lavenex P, Amaral DG. Hippocampus 2000;10:420–430. [PubMed: 10985281]
39. Murray EA, Richmond BJ. Curr Opin Neurobiol 2001;11:188–193. [PubMed: 11301238]
40. Rolls ET, Kesner RP. Prog Neurobiol 2006;79:1–48. [PubMed: 16781044]
41. Miyashita Y. Science 2004;306:435–440. [PubMed: 15486288]
42. Sakai K, Miyashita Y. Nature 1991;354:152–155. [PubMed: 1944594]
43. Fuster JM, Bodner M, Kroger JK. Nature 2000;405:347–351. [PubMed: 10830963]
44. Naya Y, Sakai K, Miyashita Y. Proc Natl Acad Sci U S A 1996;93:2664–2669. [PubMed: 8610098]
45. Naya Y, Yoshida M, Miyashita Y. Science 2001;291:661–664. [PubMed: 11158679]
46. Naya Y, Yoshida M, Miyashita Y. J Neurosci 2003;23:2861–2871. [PubMed: 12684473]
47. Higuchi S, Miyashita Y. Proc Natl Acad Sci U S A 1996;93:739–743. [PubMed: 8570626]
48. Ranganath C, D'Esposito M. Curr Opin Neurobiol 2005;15:175–182. [PubMed: 15831399]
49. Anderson JA. Math. Biosci 1972;14:197–220.
50. Grossberg S. Psychol Rev 1980;87:1–51. [PubMed: 7375607]

51. Hopfield JJ. *Proc Natl Acad Sci U S A* 1982;79:2554–2558. [PubMed: 6953413]
52. Durstewitz D, Seamans JK, Sejnowski TJ. *Nat Neurosci* 2000;3:1184–1191. [PubMed: 11127836]
53. O'Reilly, RC.; Braver, TS.; Cohen, JD. *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Miyake, A.; Shah, P., editors. New York: Cambridge University Press; 1999. p. 375–411.
54. Abeles, M. *The Handbook of Brain Theory and Neural Networks*. 2nd Edition. Arbib, MA., editor. Cambridge, MA: MIT Press; 2003. p. 1143–1146.
55. Wilson HR, Cowan JD. *Biophys. J* 1972;12:1–24. [PubMed: 4332108]
56. Amari S. *Biol Cybern* 1977;27:77–87. [PubMed: 911931]
57. Hopfield JJ. *Proc Natl Acad Sci U S A* 1984;81:3088–3092. [PubMed: 6587342]
58. Amit, DJ. *Modeling Brain Function*. Cambridge: Cambridge University Press; 1989.
59. Treves A, Rolls ET. *Network* 1991;2:371–397.
60. Treves A, Rolls ET. *Hippocampus* 1992;2:189–199. [PubMed: 1308182]
61. Amit DJ, Brunel N. *Cereb Cortex* 1997;7:237–252. [PubMed: 9143444]
62. Camperi M, Wang XJ. *J. Comput. Neurosci* 1998;5:383–405. [PubMed: 9877021]
63. Tegner J, Compte A, Wang X-J. *Biol Cybern* 2002;87:471–481. [PubMed: 12461636]
64. Brunel N, Wang XJ. *J. Comput. Neurosci* 2001;11:63–85. [PubMed: 11524578]
65. Chadderdon GL, Sporns O. *J. Cogn. Neurosci* 2006;18:242–257. [PubMed: 16494684]
66. Tagamets M-A, Horwitz B. *Cereb. Cortex* 1998;8:310–320. [PubMed: 9651128]
67. Smith EE, Jonides J. *Proc. Natl. Acad. Sci. USA* 1998;95:12061–12068. [PubMed: 9751790]
68. Krause BJ, Hautzel H, Schmidt D, Fluss MO, Poeppel TD, Muller HW, Halsband U, Mottaghy FM. *J Physiol Paris* 2006;99:318–332. [PubMed: 16723213]
69. Cabeza, R.; Kingstone, A., editors. *Handbook of Functional Neuroimaging of Cognition*. Cambridge, MA: MIT Press; 2001.
70. Paulesu E, Frith CD, Frackowiak RSJ. *Nature* 1993;362:342–345. [PubMed: 8455719]
71. Haxby JV, Ungerleider LG, Horwitz B, Rapoport SI, Grady CL. *Human Brain Mapp* 1995;3:68–82.
72. Postle BR, Berger JS, D'Esposito M. *Proc Natl Acad Sci U S A* 1999;96:12959–12964. [PubMed: 10536030]
73. Curtis CE, D'Esposito M. *Trends Cogn Sci* 2003;7:415–423. [PubMed: 12963473]
74. Mesulam M-M. *Brain* 1998;121:1013–1052. [PubMed: 9648540]
75. Horwitz B, Duara R, Rapoport SI. *J. Cereb. Blood Flow Metabol* 1984;4:484–499.
76. Moeller JR, Strother SC, Sidtis JJ, Rottenberg DA. *J. Cereb. Blood Flow Metabol* 1987;7:649–658.
77. Horwitz B. *Human Brain Mapp* 1994;2:112–122.
78. Friston KJ. *Human Brain Mapp* 1994;2:56–78.
79. McIntosh AR, Grady CL, Ungerleider LG, Haxby JV, Rapoport SI, Horwitz B. *J. Neurosci* 1994;14:655–666. [PubMed: 8301356]
80. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. *Magn. Reson. Med* 1995;34:537–541. [PubMed: 8524021]
81. McIntosh AR, Gonzalez-Lima F. *Brain Res* 1991;547:295–302. [PubMed: 1884204]
82. McIntosh AR, Gonzalez-Lima F. *Human Brain Mapp* 1994;2:2–22.
83. Horwitz B, Tagamets M-A, McIntosh AR. *Trends Cogn. Sci* 1999;3:91–98. [PubMed: 10322460]
84. Buechel C, Friston KJ. *Cereb. Cortex* 1997;7:768–778. [PubMed: 9408041]
85. Buechel C, Coull JT, Friston KJ. *Science* 1999;283:1538–1541. [PubMed: 10066177]
86. Bullmore E, Horwitz B, Honey G, Brammer M, Williams S, Sharma T. *NeuroImage* 2000;11:289–301. [PubMed: 10725185]
87. Rowe JB, Stephan KE, Friston K, Frackowiak RS, Passingham RE. *Cereb. Cortex* 2005;15:85–95. [PubMed: 15238443]
88. Kim J, Zhu W, Chang L, Bentler PM, Ernst T. *Hum Brain Mapp*. (in press)
89. Friston KJ, Harrison L, Penny W. *Neuroimage* 2003;19:1273–1302. [PubMed: 12948688]

90. Penny WD, Stephan KE, Mechelli A, Friston KJ. *Neuroimage* 2004;22:1157–1172. [PubMed: 15219588]
91. Mechelli A, Price CJ, Friston KJ, Ishai A. *Cereb. Cortex* 2004;14:1256–1265. [PubMed: 15192010]
92. Schlosser RG, Wagner G, Sauer H. *Neurosci* 2006;139:91–103.
93. McIntosh AR, Grady CL, Haxby JV, Ungerleider LG, Horwitz B. *Cereb. Cortex* 1996;6:571–584. [PubMed: 8670683]
94. Honey GD, Fu CH, Kim J, Brammer MJ, Croudace TJ, Suckling J, Pich EM, Williams SC, Bullmore ET. *Neuroimage* 2002;17:573–582. [PubMed: 12377135]
95. Hampson M, Driesen NR, Skudlarski P, Gore JC, Constable RT. *J Neurosci* 2006;26:13338–13343. [PubMed: 17182784]
96. Horwitz B, Sporns O. *Human Brain Mapp* 1994;1:269–283.
97. Arbib MA, Bischoff A, Fagg AH, Grafton ST. *Human Brain Mapp* 1995;2:225–233.
98. Horwitz B, Friston KJ, Taylor JG. *Neural Networks* 2000;13:829–846. [PubMed: 11156195]
99. Nunez, P. *Electric Fields of the Brain*. New York: Oxford Univ. Press; 1981.
100. Nunez PL. *Brain Topogr* 1989;1:199–215. [PubMed: 2641263]
101. Wood CC, Allison T. *Canad. J. Psychol* 1981;35:113–135. [PubMed: 7248847]
102. Freeman WJ. *Biol. Cybern* 1987;56:139–150. [PubMed: 3593783]
103. Freeman, WJ. *Methods of Analysis of Brain Electrical and Magnetic Signals*. Gevins, AS.; Remond, A., editors. vol. 1. Amsterdam: Elsevier; 1987. p. 583–673.
104. Riera J, Aubert E, Iwata K, Kawashima R, Wan X, Ozaki T. *Phil. Trans. Roy. Soc. Lond. B* 2005;360:1025–1041. [PubMed: 16087446]
105. David O, Friston KJ. *Neuroimage* 2003;20:1743–1755. [PubMed: 14642484]
106. Kiebel SJ, David O, Friston KJ. *Neuroimage* 2006;30:1273–1284. [PubMed: 16490364]
107. Stephan KE. *J. Anat* 2004;205:443–470. [PubMed: 15610393]
108. Horwitz B. *Neuroinformatics* 2004;2:251–266. [PubMed: 15319520]
109. Kubovy M, Van Valkenburg D. *Cognition* 2001;80:97–126. [PubMed: 11245841]
110. Ungerleider, LG.; Mishkin, M. *Analysis of Visual Behavior*. Ingle, DJ.; Goodale, MA.; Mansfield, RJW., editors. Cambridge: MIT Press; 1982. p. 549–586.
111. Kaas JH, Hackett TA, Tramo MJ. *Curr. Opinion Neurobiol* 1999;9:164–170.
112. Rauschecker JP. *Acta Otolaryngol. (Stockh.)* 1997;532:34–48.
113. Horwitz B, Tagamets M-A. *Human Brain Mapp* 1999;8:137–142.
114. Husain FT, Tagamets M-A, Fromm SJ, Braun AR, Horwitz B. *NeuroImage* 2004;21:1701–1720. [PubMed: 15050592]
115. Van Essen, DC. *Cerebral Cortex*. Peters, A.; Jones, EG., editors. vol. 3. New York: Plenum Press; 1985. p. 259–328.
116. Funahashi S, Bruce CJ, Goldman-Rakic PS. *J Neurophysiol* 1989;61:331–349. [PubMed: 2918358]
117. Tagamets M-A, Horwitz B. *Neural Networks* 2000;13:941–952. [PubMed: 11156203]
118. Husain FT, Lozito TP, Ulloa A, Horwitz B. *J. Cogn. Neurosci* 2005;17:1275–1292. [PubMed: 16197683]
119. Macoveanu J, Klingberg T, Tegner J. *Neuroscience* 2006;141:1611–1168. [PubMed: 16777342]
120. Edin F, Macoveanu J, Olesen P, Tegner J, Klingberg T. *J. Cogn. Neurosci.* (in press)
121. Cowan N. *Beh. Brain Sci* 2001;24:87–114.
122. Wang XJ, Tegner J, Constantinidis C, Goldman-Rakic PS. *Proc Natl Acad Sci U S A* 2004;101:1368–1373. [PubMed: 14742867]
123. Rowe JB, Toni I, Josephs O, Frackowiak RS, Passingham RE. *Science* 2000;288:1656–1660. [PubMed: 10834847]
124. Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV. *Science* 2000;279:1347–1351. [PubMed: 9478894]
125. Chafee MV, Goldman-Rakic PS. *J. Neurophysiol* 1998;79:2919–2940. [PubMed: 9636098]

1.

Network diagram of the object processing models [66,114]. The regions of each model form a complex network of feedforward and feedback connections; these interregional connections can be either excitatory (excitatory-to-excitatory elements, shown as solid lines) or inhibitory (excitatory-to-inhibitory elements, shown as dashed lines). Regions specific to the visual model (LGN, V1–V2, V4, IT) are shown in black-bold; regions specific to the auditory model (MGN, Ai, Aii, ST) are shown in gray-italics; the prefrontal module is structured in the same way in both models. The sub-modules for each model are also indicated. In the PFC region, there are four submodules: FS contains stimulus-sensitive units, D1 and D2 contain units active during the delay part of a delayed match-to-sample task, and FR contains units whose activity increases if there is a match between the first and second stimuli of a trial. Abbreviations: LGN – lateral geniculate nucleus; MGN – medial geniculate nucleus; Ai – primary sensory auditory cortex; Aii – secondary sensory auditory cortex; PFC – prefrontal cortex.

Simulation vs. Experiment Comparison

2.

Comparison between model simulation and experimental results. (Left) Bar graphs of the percent signal change in simulations using the visual model (white bars) in each brain region and the corresponding results (black bars) from the comparable PET study of Haxby et al. [71]. The percent signal changes were evaluated as the difference between the DMS task of object shape vs. a passive viewing task of degraded shapes (see Tagamets & Horwitz [66]) for details). (Right) A similar comparison for the auditory model and the corresponding fMRI study. The percent signal changes were computed as the normalized difference of the percent signal changes of both the DMS and control tasks relative to baseline (see Husain et al. [114] for details).