

Integrating Electrophysiological and Anatomical Experimental Data to Create a Large-scale Model that Simulates a Delayed Match-to-sample Human Brain Imaging Study

M.-A. Tagamets¹ and Barry Horwitz

Laboratory of Neurosciences, National Institute on Aging,
National Institutes of Health, Bethesda, MD, USA

¹Current address: Georgetown Institute for Cognitive and
Computational Sciences, Georgetown University, Washington,
DC, USA

We propose a model that draws together experimental evidence from anatomical, electrophysiological and imaging experiments in order to understand better the neural substrate of human imaging studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). First, we define a simple local circuit that reflects the major role that local connectivity plays in producing PET and fMRI data, which are thought to mainly reflect synaptic activity. Second, in order to account for the role of varying behaviors during the course of a typical imaging experiment, we propose a local circuit that can perform a delayed match-to-sample task. The elements of this circuit behave very much like neurons that have been found in the prefrontal cortex during similar tasks in monkeys. One subpopulation responds selectively only when stimuli are present. Two different populations show the two types of delay-period activity that have been identified, one with high activity both during the cue and the delay period, the other with a rise during the delay period only. Last, a subpopulation shows a brief response only if the second stimulus matches the first, thus mediating the decision about whether the stimuli match. We show that in addition to performing the task, the integrated summed synaptic activities of the model are similar to experimental PET data.

Introduction

Functional imaging methods have been used extensively in recent years to map specific cognitive functions in the human brain. These techniques include positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and electrical and magnetic encephalography (EEG and MEG). Many of the tasks studied also have been investigated in nonhuman primates using single-unit electrode recording methods. Interpreting all these data into a coherent and unified picture has proved difficult for a variety of reasons. These include different spatial and temporal resolutions, different underlying mechanisms relating neural activity to the signals measured, and different experimental paradigms. These and various other issues that are related to interpreting functional imaging data by modeling have been discussed in Horwitz and Sporns (1994). A key suggestion in that paper was that computational modeling methods can be used to understand some of these issues and to interpret human imaging data. In this paper we present such a large-scale computational model for human PET experiments of working memory. Preliminary aspects of this work have appeared in Tagamets *et al.* (1995) and Tagamets and Horwitz (1997a).

At the time of writing there has been only one published modeling study of human functional imaging data (Arbib *et al.*, 1995). This model, referred to as *synthetic pet* by the authors, computed PET data during a simulated saccade task by summing the absolute values of afferent synaptic activities within the areas of the model. Part of the purpose of their study was to test the hypothesis that inhibitory afferent synaptic activities can produce an increase in PET activity, an effect that has

been suggested by experimental studies (Ackermann *et al.*, 1984; Horwitz and Sporns, 1994; Jueptner and Weiller, 1995). That this turned out to be the case in their model emphasizes the potentially counterintuitive nature of interpreting imaging data. In a more recent modeling study, we have shown how the balance of local excitation and inhibition can influence the net effect of afferent inhibition in imaging data (Tagamets and Horwitz, 1997b).

Here we present a description of a model whose design takes into account some of the issues that are specific to modeling quantitative data such as PET and fMRI. We describe a systematic approach that allows us to relate other sources of experimental data, such as single-cell recordings and anatomical studies, to human brain imaging results. We then demonstrate the use of such a model by applying it to a study of a human delayed match-to-sample task (Haxby *et al.*, 1995). The model is made up of multiple interconnected brain areas that are thought to play a role in a visual delayed match-to-sample shape discrimination task. The specific questions that we address in this work are the following:

1. What are the relative roles of local and between-area connections in generating both electrical neuronal activities and quantitative functional imaging data?
2. What circuits, both local and global, might underlie a task such as short-term memory?
3. Where might the modulation of this memory take place?

We first introduce a simple canonical local circuit, based on anatomical and experimental data, (Douglas *et al.*, 1995), which takes into account the relatively important effects of local circuitry on total integrated synaptic activity within an area. All regions in our model are made up of populations of these canonical units and differ only in how they are connected to one another. Parameters are selected so that these units have response properties similar to those seen experimentally in single-electrode recordings in nonhuman primates. We use a bottom-up approach in the design of the network and in selection of parameters, choosing the exact structure of the basic unit and its dynamics first, then connecting groups into local regions, and finally connecting the regions to one another to form the full network.

The full model simulates multiple brain regions along the ventral object vision pathway (Ungerleider and Mishkin, 1982) extending into the frontal lobe. The tasks we simulate have been studied in both human brain imaging experiments (e.g. Corbetta *et al.*, 1991; Sergent *et al.*, 1992; Haxby *et al.*, 1995; Courtney *et al.*, 1996) and in single-unit recordings in monkeys (e.g. Fuster *et al.*, 1982; Haenny *et al.*, 1988; Fuster, 1990; Wilson *et al.*, 1993; Miller *et al.*, 1996).

We introduce a new model of working memory in the

prefrontal cortex and we relate this to the question of relative timing of different aspects of a cognitive task such as that which is typically performed during PET experiments. To simulate a delayed match-to-sample task, the model network is presented with multiple pairs of stimuli and the dynamic activity in each brain region is followed over the course of the simulated experiment. Two subpopulations of elements mediate the matching task by holding a representation of the last seen visual stimulus in memory during a delay period, after which there is a response from another subpopulation of units only if the second stimulus matches the first. A second set of similar trials serves as a control task, in which there is no requirement for matching, and the stimuli are not held in memory. The memory is modulated by a weak, diffuse signal to a specific subpopulation of the memory units. During simulation of each task, a function of the integrated synaptic activity is taken to represent regional cerebral blood flow (rCBF), in a manner similar to that used by Arbib *et al.* (1995). The resulting network shows the dynamic electrical activity in each brain area that is similar to that found in single-unit monkey experiments, and the integrated activity in each brain region resembles the focal activations observed during functional neuroimaging studies. Although the models we construct are equally applicable to PET and fMRI neuroimaging data, we will restrict our focus in this paper to PET/rCBF data.

Methods: A Computational Framework for Modeling Functional Imaging Data

Local Connectivity Patterns and the Basic Unit for the Model

The local response and total synaptic activity within a cortical area depends on the interaction of the afferent connections, originating from other areas, and on local connectivity, which shapes the resulting response. This response, in turn, depends on the proportions of excitatory and inhibitory synapses in the local circuits. It is believed that most synapses in the cortex are excitatory and that a vast majority of synapses onto a neuron come from other neurons within the same area as the cell (Douglas and Martin, 1991; Douglas *et al.*, 1995). This has prompted the idea of ‘amplification’ of neuronal responses in local circuits in response to a relatively small amount of afferent excitation. For quantification of rCBF from total synaptic activity, this amplification effect, as well as the effect of afferent inhibition, needs to be taken into account. From the data in Douglas *et al.* (1995) we can derive the proportions of local and afferent connections as shown in Table 1. These local circuit connectivities are computed as follows. As outlined in Douglas *et al.* (1995), we have that (1) 85% of the synapses in cortex are excitatory, and (2) of those, 85% are to other excitatory neurons. This implies that 15% of total synapses are from inhibitory neurons [from (1)] and that $85\% \times 85\% = 72\%$ of synapses are excitatory-to-excitatory [from (1) and (2) above], with the remaining 13% (i.e. $85\% - 72\%$) of excitatory connections being made to inhibitory neurons. The total afferents coming into an area from other regions are thought to make up a relatively small proportion of local connectivity. For example, various estimates have been made for the total afferents from the LGN to V1, with the 10% cited in the Douglas *et al.* (1995) paper on the low side. If we assume a conservative 10% of total connections arising from other regions, and, as a first approximation we assume that inter-regional connections are excitatory-to-excitatory, then we end up with a total local excitatory-to-excitatory connectivity of 72% (total) – 10% (inter-regional) = 62% .

A Canonical Local Circuit

The resolution of PET data used by most groups is usually at least 1 cm^2 . A scale of about an order of magnitude below this, $\sim 0.5\text{--}1 \text{ mm}^2$, approximately corresponds to the area of a hypercolumn in the visual cortex, which has been estimated to be $\sim 750 \mu\text{m}$ in diameter in V1 (Ts'o *et al.*, 1986) and $\sim 450 \mu\text{m}$ in diameter in the inferotemporal cortex (Tanaka, 1993). We choose the basic unit of the model to be at this scale, representing a local assembly whose elements tend to have similar

Table 1

Proportions of different types of synapses in the basic unit (computed from data in Douglas *et al.*, 1995)

Type	Effect	Approximate percent of total
Local	excitatory → excitatory	60
Local	excitatory → inhibitory	15
Local	inhibitory → excitatory	15
Afferent	any	10–20

The derivation of these proportions is described in the text.

responses (e.g. Hubel and Wiesel, 1977; Tanaka, 1993; Wilson *et al.*, 1994). One element is used to represent the local subpopulation of excitatory pyramidal cells and another represents the inhibitory cell population within the column. Such a configuration forms an example of the well known Wilson-Cowan unit (Wilson and Cowan, 1972). In our model these elements are connected by following the rules for proportions of local connectivity outlined in Table 1. The resulting basic circuit is shown in Figure 1A. The circle labeled ‘E’ represents the excitatory population and the circle labeled ‘I’ indicates the inhibitory population. We fix local synaptic efficacies to 0.6, 0.15 and -0.15 for all excitatory-to-excitatory, excitatory-to-inhibitory and inhibitory-to-excitatory within-unit connections, yielding local proportions approximately as shown in Table 1. These values are used throughout the model. Each region is modeled by one or more groups of 81 basic units arranged as shown in Figure 1B. These units are arranged spatially in a 9×9 configuration, which represents a patch of cortex $\sim 0.5\text{--}1 \text{ cm}$ in diameter.

Computing Activity in the Model

In order to distinguish the two types of activities, electrical (spiking) and blood flow, we will refer to responses of individual units in the model as the *electrical* activity and the modeled imaging data as *rCBF* activity. The electrical activity of each E and I element is determined by a sigmoidal function of the summed synaptic inputs that arrive at the unit. This corresponds to average spiking rates from single-cell recordings. At the same time, the absolute values of all synaptic activities are summed and local blood flow is computed by integrating these values over space and time. This provides the quantitative data for matching imaging studies. The details of these functions are described in the following.

Electrical Activity: The Activation Rule and Parameter Values

The dynamics of the network are determined by the sigmoid activation rule:

$$\frac{dE_i(t)}{dt} = \Delta \left(\frac{1}{1 + e^{-K_E [w_{EE}E_i(t) + w_{EI}I_i(t) + in_{iE}(t) - \tau_E + N(t)]}} \right) - \delta E_i(t) \quad (1)$$

$$\frac{dI_i(t)}{dt} = \Delta \left(\frac{1}{1 + e^{-K_I [w_{EI}E_i(t) + in_{iI}(t) - \tau_I + N(t)]}} \right) - \delta I_i(t) \quad (2)$$

$E_i(t)$ and $I_i(t)$ represent the electrical activations of the i th excitatory and inhibitory elements at time t respectively. K_E and K_I are the gains, or steepness, of the sigmoid functions for excitatory and inhibitory units respectively, τ_E and τ_I are the input thresholds of excitatory and inhibitory units, Δ is the rate of change, δ is the decay rate, and $N(t)$ is an added noise term. w_{EE} , w_{EI} and w_{EI} are the weights within a unit: excitatory-to-excitatory (value = 0.6), inhibitory-to-excitatory (value = -0.15), and excitatory-to-inhibitory (value = 0.15) respectively. $in_{iE}(t)$ and $in_{iI}(t)$ are the total inputs coming from other areas into the excitatory and inhibitory units at time t :

$$\begin{aligned} in_{iE}(t) &= \sum_j w_{ji}^E E_j(t) + \sum_j w_{ji}^I I_j(t) \\ in_{iI}(t) &= \sum_k w_{ki}^E E_k(t) + \sum_k w_{ki}^I I_k(t) \end{aligned} \quad (3)$$

w_{ji}^E and w_{ji}^I are weights coming from excitatory/inhibitory unit j in another area into the i th excitatory and inhibitory units respectively.

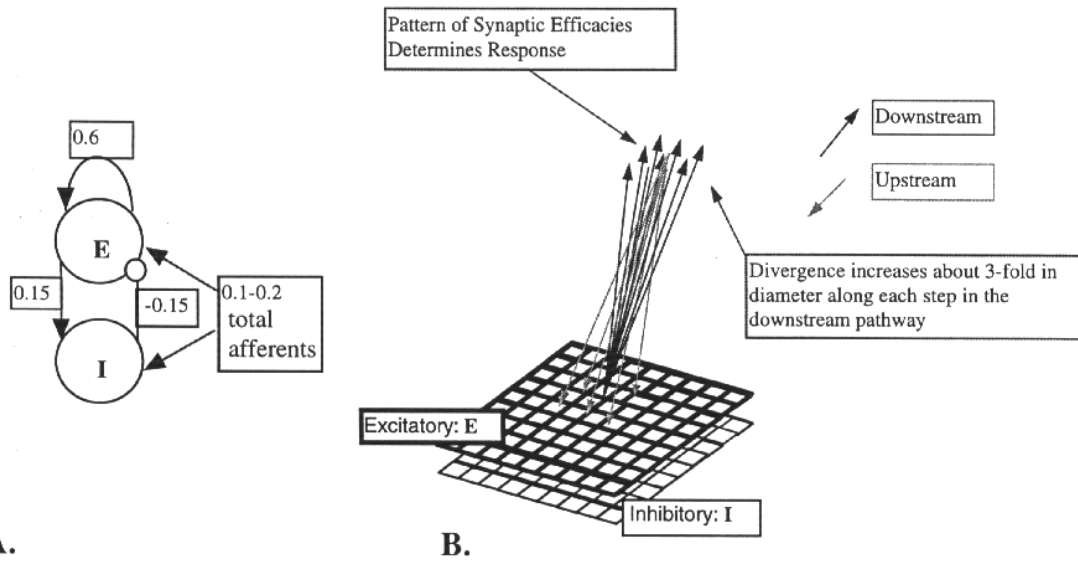


Figure 1. (A) The basic unit that results from applying the data in Table 1. E represents the excitatory population and I the inhibitory population in a local assembly such as a cortical column. Local synaptic activity is dominated by the local excitation and inhibition, while afferents account for the smallest proportion, as indicated by the synaptic weights shown. (B) A cortical area is modeled by one or more 9×9 sets of basic units. The excitatory population is shown in bold lines above the inhibitory group, shown in lighter lines. Individual units in the excitatory and inhibitory populations within a group are connected as shown in (A).

Electrical activations in the model range between 0 and 1, and can be interpreted as reflecting the percentage of active units within a local population.

Choosing Parameter Values

The values for the parameters K_E , K_I , τ_E , τ_I , Δ and δ are chosen by both analytical and empirical means in order to achieve the desired dynamics within a single basic circuit. Specifically, thresholds τ_E and τ_I are selected so that if both excitatory and inhibitory elements are receiving half their possible inputs, the unit will have an equilibrium state in which each element is equal to 0.5, i.e. half the possible maximum value of 1. These values are easily computed if one notes that equations (1) and (2) are at equilibrium when the argument of the exponential (in parentheses) in the denominator is equal to 0. K_E and K_I determine the slopes of the sigmoid functions, and can roughly be viewed as determining the responsiveness of the E and I elements in the circuit. These are chosen so that the above-mentioned equilibrium point is stable, but near the bifurcation point. The bifurcation point, or the values for the K s at which the stability of this fixed point changes from being an attractor to unstable is when $K_I \approx 2K_E$. An expression for this relation between K_E and K_I in a single E-I unit can be derived by the use of elementary theory of differential equations, solving for the zero-crossings of the characteristic equations of the system. The parameters Δ and δ determine the time resolution during simulations. Table 2 summarizes the resulting values for all the parameters and their effects on computed rCBF within a single basic unit. These parameter values are the same throughout all areas of the model network. The values for local weights within a basic unit have already been discussed, and are also fixed at the same values throughout the network. Thus, both the distinct electrical behaviors and the computed blood flow in the different regions are determined by the connectivity patterns among units within an area and between different areas, as will be discussed later.

Computing PET Activations

rCBF in the model is computed within an area as:

$$\text{rCBF} = \sum_{t,i} IN_i(t) \quad (4)$$

$IN_i(t)$ is the total sum of absolute values of the inputs to both the E and I elements of unit i , both from within the unit and from all sources, at time t , and is computed as:

Table 2

Values of parameters that are used in the activation rule of the model, and effects on simulated rCBF that come from changing these values

	Value	Effect	Value	Effect
	Excitatory elements	Raising value causes rCBF to:	Inhibitory elements	Raising value causes rCBF to:
K (gain)	9.0	↓	20.0	↑
Threshold	0.3	↓	0.1	↑
Noise	± 0.1	↑	± 0.1	↑
Rate of change (Δ)	0.5	NC	0.5	NC
Decay (δ)	0.5	↓	0.5	↑

Numbers in columns 1 and 3 (Value) give the parameter values that were used in all areas of the model, in excitatory elements and inhibitory elements respectively. The Effect columns show the changes in computed rCBF that come from changing these values. These changes are computed from modeled rCBF in a single area that receives constant input to the excitatory element.

Key: ↑ means that raising the parameter value raises computed rCBF in the area. ↓ means that raising the parameter value reduces computed rCBF in the area. NC indicates that the value of the parameter has no effect on computed rCBF.

$$IN_i(t) = w_{EE}E_i(t) + w_{EI}E_i(t) + w_{IE}I_i(t) + \sum_{k,i} w_{ki}E_k(t) + \sum_{m,i} |w_{mi}I_m(t)| \quad (5)$$

The first three terms on the right-hand side of equation (5) above are the synaptic weights from within the unit i itself, while the last two are the summed activities coming from all the other units which are connected to unit i . This is similar to the method used in Arbib *et al.* (1995) for computing simulated blood flow. For comparison with experimental PET data, both the simulated data and the PET data are normalized to a common reference area, e.g. area V1, for one of the tasks. This allows us to examine both the relative amount of blood flow in different areas in various tasks in addition to computing the percentage changes between tasks.

A Large-scale Model that Performs a Delayed Match-to-sample Task

In this section we propose a model that mediates a delayed match-to-sample task such as those used in both single-unit studies and in rCBF/PET studies. In particular, we model a working memory task that performs

shape matching. This involves constructing a network of interconnected model areas that represents a cortical system. Each area is composed of one or more 9×9 populations of E-I units like those shown in Figure 1B in the previous section. Separate subpopulations mediate different aspects of the task and will be described in more detail below.

The approach that we use in constructing the model can be divided into four main components:

1. Choosing the areas to be included in the model. This is based on properties of brain areas that have been identified by various methods, including lesion studies, imaging studies, and single-cell recording data.
2. Determining the desired response properties for units in each area. These are mainly based on data from single-cell recordings in monkeys.
3. Determining appropriate connectivity patterns between areas. The connectivity patterns influence both the electrical and the rCBF activations. The major connections are based on monkey data and the specific topography of the patterns of connectivity are based on both monkey data and computational studies.
4. Defining the way in which a PET study is to be simulated.

The goal is to account simultaneously for three different phenomena. First, the simulated electrical responses of neuronal units in each area of the model must approximate the behaviors of single-unit recording data in analogous areas in primates. Second, the model needs to account for how such behaviors could mediate the tasks that are being modeled. Third, during a simulation of a PET study with the model, the absolute values of the integrated synaptic activities in each area should be similar to the quantitative values that are seen in human blood flow data during these tasks. This must hold for all the interconnected regions of the model simultaneously for at least two different tasks, representing the task of interest and the control.

Areas Included in the Model

Figure 2 depicts some of the major areas that are known to be involved in visual processing. The task that we model, visual shape matching, involves mainly the occipitotemporal visual pathway. Single-cell recordings in primates have provided data about specific visual response properties in these areas (Hubel and Wiesel, 1977; Desimone and Schein, 1987; Tanaka, 1993), as have imaging studies in humans (Corbetta *et al.*, 1991; Haxby *et al.*, 1991; Sergent *et al.*, 1992; McIntosh *et al.*, 1994b; Courtney *et al.*, 1996). This pathway includes areas V1, V2, V4, TEO, the inferotemporal cortex (IT) and lateral prefrontal cortex (Felleman and Van Essen, 1991). The areas included in our model are shown in bold in Figure 2. Although it is not known exactly what the homology is between humans and nonhuman primates in these areas, anatomical and imaging studies suggest that at least the earlier parts of this pathway are similar between the two species (Burkhalter and Bernado, 1989; Sereno *et al.*, 1995; DeYoe *et al.*, 1996). The percentages shown are from anatomical data and reflect approximate proportions of afferent connections that have been estimated in anatomical monkey studies (Doty, 1983; Tanaka *et al.*, 1990). In the following we outline some of the salient properties of these areas and their corresponding implementation in the model. The details of the specific connectivity patterns that achieve these behaviors are given in Appendix A.

V1/V2

The early visual areas V1 and V2 are combined in the model. In single-unit recordings, both areas have neurons that respond to elemental visual features such as oriented lines, edges and colors (Hubel and Wiesel, 1977; Peterhans and von der Heydt, 1993; Roe and Ts'o, 1995), and neither area has been shown to exhibit any substantial changes in single-unit activity that can be attributed to influences such as attention. Additionally, these two areas appear to have similar patterns of afferent connections (Rockland and Van Hoesen, 1994).

Most neurons in areas V1 and V2 are at least to some degree tuned to oriented lines (Hubel and Wiesel, 1977). We simplify orientation selectivity to units that respond preferentially to either horizontal or vertical line segments in the visual field, thus forming two populations of

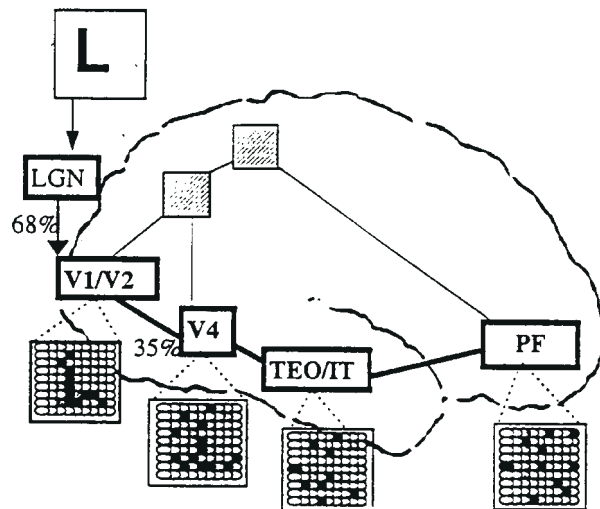


Figure 2. The model is made up of areas along the occipitotemporal pathway, which is thought to be the main pathway that is involved with object vision. Different populations of units process incoming data in successively abstracted stages, indicated by the representations shown in response to an input of a letter L.

units that represent the V1/V2 area in the model. This generalizes the 'vocabulary' of the model to objects that are composed of such line segments, while keeping the complexity of the network to a minimum. These coding properties are achieved in the model by oriented converging connections from the LGN.

V4

The shape perception properties of neurons in area V4 are similar to the earlier areas, but appear to encode somewhat more complex properties of shape. Most neurons in V4 have been found to respond to form (Desimone and Schein, 1987). This includes orientation, as in V1, and simple shape features such as length and width of lines (Desimone and Schein, 1987) and concentric circles and X patterns (Gallant *et al.*, 1993). This suggests that V4 is a continuation of the shape pathway and probably encodes shape from combinations of the features encoded in the earlier areas. In the model, V4 is made up of three populations of units that encode longer line segments, both horizontal and vertical, and also corners formed by adjacent pairs of horizontal and vertical lines. These properties are attained by a similar principle as in V1, with a cooperative combination of afferent convergent connections.

In single-unit recordings, V4 has been shown to exhibit changes that are associated with a cognitive state such as attention (Haenny *et al.*, 1988; Spitzer *et al.*, 1988; Chelazzi *et al.*, 1995). This is usually characterized by stronger responses of units in this area when the feature is behaviorally relevant. This is implemented in the V4 area of the model by feedback connections from both the TEO/IT and the prefrontal areas.

TEO/IT

Single-unit recordings in the inferior temporal area of monkeys have shown that neurons in this region respond to a variety of shape characteristics. Some cells are selective for bars of a particular length or width, similar to area V4, though most neurons in area IT appear to be selective to more complex features of shape (Desimone *et al.*, 1984; Tanaka, 1993). Populations of neurons appear to be organized similarly to earlier areas, with column-like groupings that tend to share responses to specific features of shapes (Tanaka, 1993). The TEO/IT area is modeled as a single region whose elements respond best to shapes that are made up of connected lines, such as squares and 'L' shapes. The units in this area of the model obtain their selectivity in a two-stage procedure: (i) generation of an initial rough map, created by sparse, random connections from V4; and (ii) sharpening the tuning by a competitive Hebbian learning rule (Tagamets, 1994). This learning rule and its consequences are described more fully in Appendix B.

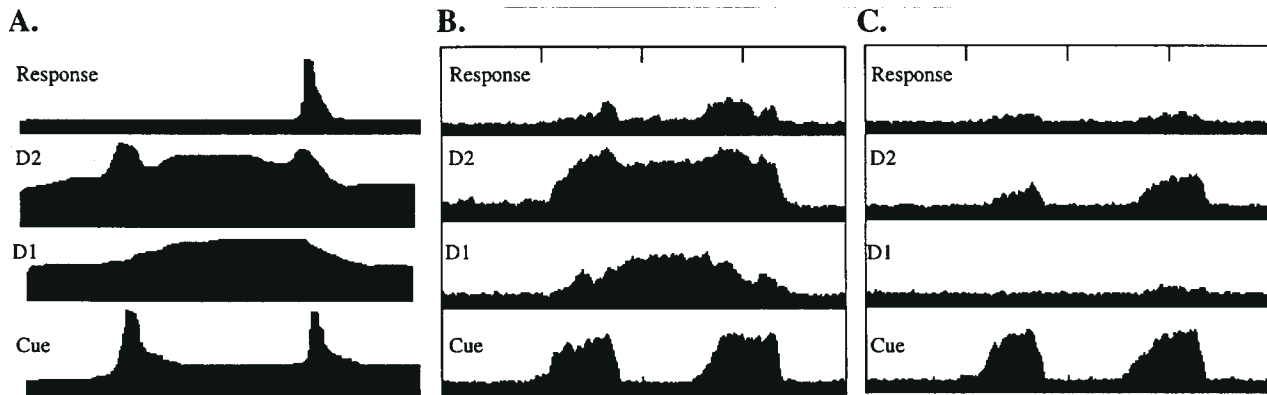


Figure 3. (A) The leftmost panel shows idealized patterns of behavior of the different types of neurons that have been identified in prefrontal cortex of the macaque in spatial delayed-response experiments (Funahashi *et al.*, 1990; Goldman-Rakic, 1995). Two types of delay units seem to be part of the working memory circuit. D1 units become active only when the stimulus has disappeared from view, while the D2 units are active both during the stimulus and during the delay period. In addition there are cue units that respond only when a stimulus is in view, and response units that show a brief activation after presentation of the test stimulus. (B,C) Behavior of averaged activity over five trials of the similar units in the prefrontal area of the model, with attention (B), and without (C). In the model, the response units indicate whether a match has occurred. As discussed in the text, it is likely that the response units of the model subserves a different purpose than the response units shown in (A).

Prefrontal

The selectivity of single-cell responses to stimuli in visually responsive areas of the prefrontal cortex are similar to those in area IT (e.g. Fuster, 1990; Wilson *et al.*, 1993). The prefrontal area has also been found to have a substantial population of neurons that maintain a high level of activity during the delay period of a delayed match-to-sample object-matching task (Fuster, 1973; Fuster *et al.*, 1982; Wilson *et al.*, 1993; Miller *et al.*, 1996) as well as in spatial delayed-response tasks (Funahashi *et al.*, 1990; Funahashi and Kubota, 1994). Some studies suggest that such delay period activity depends to a large extent on the behavioral relevance of the stimulus to the animal (Fuster, 1973; Rosenkilde *et al.*, 1981). For example, data from the study by Fuster (1973) in the prefrontal area indicated that there was little delay period activity if the animal was not expecting a reward, while ~50% of the recorded neurons had delay activity during actual performance of the task, when a reward was expected. This suggests modulation of the working memory by some form of 'attention'.

In the spatial delayed-response task, delay-active neurons in the prefrontal cortex have further been classified by Funahashi *et al.* (1990) as being of two types: those that are active only during the delay, and others that display activity during both the initial stimulus and the delay periods. Typical electrical behaviors of these neurons, as seen in single-cell recordings from monkeys, are illustrated in Figure 3A. Experimental results suggest that the mechanisms underlying memory for shape matching are similar to those that involve spatial stimuli (Miller *et al.*, 1996).

A Local Memory Circuit

These data suggest a memory circuit that preserves the representation of the cue shape during the delay period. We have implemented a short-term memory circuit as shown in Figure 4, made up of four subpopulations of units in the model prefrontal area, with electrical responses as follows:

1. Cue units (C) respond only when there is a stimulus present in the LGN area.
2. Delay only units (D1) become active only during the delay period after presentation of the first stimulus.
3. Delay + cue units (D2) are active during presentation of the stimulus and during the delay.
4. Response units (R) display a brief activation if the second (test) stimulus matches the first and the first stimulus has been remembered.

The prefrontal area of the model is made up of a 9×9 array of such circuits, each circuit roughly representing a local assembly such as a

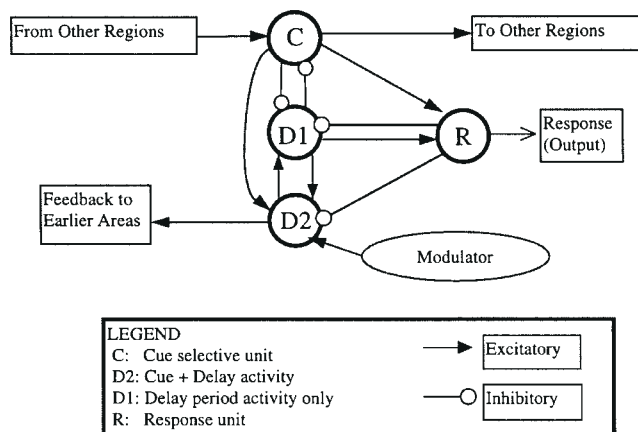


Figure 4. A working memory is added to the prefrontal areas of the model as a local circuit composed of different types of units, as identified in studies such as shown in Figure 3. Each element of the circuit shown is a basic unit shown in Figure 1A. Inhibitory connections are effected by excitatory connections onto inhibitory units. An attentional modulation is modeled as a diffuse, weak activity directed into the D2 units. These D2 units also are the source of feedback into earlier areas. Details of the connection weights are given in Appendix C.

column. Representations of objects in the prefrontal area of the model are formed by distributed activities of these circuits.

Although the response units shown in Figure 3A are from a spatial delayed matching task and have been shown to respond without a match, we posit that there must also exist a subpopulation of units that mediate the decision of whether a test stimulus matches a cue. The justification for this is twofold: (i) a decision at the cognitive level requires the existence of some neuronal elements that respond differentially to matching and nonmatching stimuli; and (ii) data from Miller *et al.* (1996) suggest the existence of such units in the prefrontal cortex. The specific connection weight values among the units shown in Figure 4 are given in Appendix C.

The effect of attention in the model is implemented by a low-level, diffuse incoming activity to the D2 units, as shown in Figure 4. While we do not model the source of this modulation, our model does make it explicit that the D2 units are the recipients. When the attentional level is low, there is no delay period activity in the D1 and D2 units. In this case, the response units of the model also do not become activated, even if the second stimulus matches the first. With a higher level of attention, a given

stimulus activates the delay units. After the stimulus disappears, both the D1 and D2 units remain active until presentation of the second stimulus. If the second stimulus matches the first, the response units show a brief activity. It is this circuit that implements the working memory and matching tasks in the model.

Between-area Connectivity

We have used three main criteria for choosing the inter-area connection patterns: (i) overall percentage of all connections between areas; (ii) the fanout, or divergence, of connections; and (iii) the specific topography – oriented longitudinally, clustered, diffuse, etc. While all of these criteria affect both qualitative and quantitative results, the first has the most influence on quantification of synaptic activities and the third has the most effect on specific electrical response properties of individual units.

Incoming connection strengths between areas of the model are shown in Appendix A. These values were obtained by starting with approximate values from Figure 2 and adjusting to achieve the desired behaviors. In our model, all afferents constitute ~10–20% of the total connections in each area. The topology of the connections is based both on known features of the topography of receptive fields and on modeling studies. Connections between areas in the early visual pathway tend to be topographic, with neighboring units sending connections to neighboring units in the next area. It has been estimated that the size of the receptive field grows by ~3-fold in each successive stage in the feed-forward direction, so that by the inferotemporal area the receptive fields are very large (e.g. Desimone *et al.*, 1984) and there is little or no retinotopy. The widening of the receptive fields in the model is implemented by increasing both the divergence and sparseness of connections in the feed-forward direction. The sparseness and specific weight values are randomized, while maintaining the total percentage of afferents per unit on average. The topography of afferents in the earlier parts of the pathway are based on theoretical and modeling studies (e.g. Linsker, 1988; Sporns *et al.*, 1991; Tononi *et al.*, 1992). For example, orientation selectivity in V1 is achieved by oriented converging inputs from the LGN area. Connections from V4 to the TEO/IT area are tuned with a learning rule as described earlier. A more detailed description of how these connections are generated is given in Appendices A and B.

Simulating PET and the Match-to-sample Task

A typical PET study involves two or more different tasks, such as matching objects in a visual display in task 1 and passive viewing of a similar display in task 2. Our model needs to account for both electrical activity and blood flow in the various model cortical regions in both tasks. Each task can be divided into several qualitatively different components that correspond to different phases during a single trial of a PET study. A simple delayed match-to-sample task can be divided into at least five components that differ in the time during which they occur and in the type of cognitive activity that is required during the course of a PET study:

1. The cue period, when the initial stimulus is presented. This is usually on the order of one to a few seconds during which there is visual input which the subject must attend to and encode for later recall.
2. A delay period, during which there is no visual input and the subject presumably keeps the cue item in memory. This typically lasts from one to tens of seconds.
3. A test period, when a second stimulus is presented. At this time there is visual input which the subject must compare to the previously seen cue.
4. A decision event, when the subject makes a choice.
5. The intertrial period, before the next trial begins.

During a simulated PET study, the model must respond appropriately at each of these phases by mediating the task with electrical activity that is similar to that seen in electrophysiological studies. At the end of a series of repetitions of the task, i.e. after a simulated PET session, the net summed blood flow activity must be similar to that seen in human PET studies of the match-to-sample task.

The shape stimuli used in this study are simple figures composed of horizontal and vertical line segments (see the LGN area in Fig. 6). For each run of a simulation, stimuli are presented by setting a patterned group of units in the LGN area to a high activity level, and the other areas update

their activities according to the activation rules given in equations (1) and (2). The simulations are performed using a high-level specification language written in C++. Each iteration during the simulation corresponds to ~5 ms in real time. Noise at a level of ~10–50% of the total synaptic activity is added to the afferent activities at each iteration in order to model the nonspecific effects that might arise from sources that are outside the scope of this model. During delay periods all units in the LGN are set to a uniformly low activity level. This creates a sequence of stimuli that is similar to those generally used for both PET and single-cell studies. We explicitly simulate part of the human working memory experiment of Haxby *et al.* (1995). In this experiment, stimuli were presented for 4 s and the shortest delay period was 1 s.

A typical PET subtraction study is designed to measure differences in brain activity in two different tasks in order to identify changes that are due to the cognitive attribute of interest (Posner *et al.*, 1988). The tasks are usually designed so that they differ only in the attribute of interest. In our case, the goal is to examine areas of the cortex that are specifically involved in short-term memory processing in the visual pathway. The number of activated units in the LGN, serving as sensory input to the system, are the same in both tasks. In the match-to-sample task, the effect of interest is the mechanism of matching ability and the modulation of the delay period activity in order to activate the memory circuits described earlier. This is implemented by a low-level incoming activity into the delay-plus-cue (D2) units, as described above. A control task is simulated by presenting a degraded version of the same stimuli and a lower level of modulatory activity to the D2 units in the prefrontal region. ‘High’ attention in the model is ~1.5% of the total synaptic activity to these units, while ‘low’ attention is ~1%. The use of degraded stimuli is common in the type of PET study that we are modeling (cf. Haxby *et al.*, 1994, 1995; Courtney *et al.*, 1996). This is presumably because the salience or familiarity of stimuli such as faces can produce an involuntary attentional effect even though there is no explicit requirement for attention. This, in turn, implies that results from human matching studies such as those cited above can arise from an interaction of bottom-up and top-down effects.

Results

Electrical Activity

The electrical activity of the model is shown in Figures 3B,C, 5 and 6. For Figures 3 and 5 we used stimuli and delays of 1 s duration in order to compare the behavior over time with the experimental data shown in Figure 3A. In Figure 3 we compare the electrical activities of units in the prefrontal area of the model (B,C) with experimental data (A). Figure 3B,C shows the activities of the elements of a single frontal unit in the model averaged over five trials. Figure 3B shows the high attention case, while Figure 3C shows the same units with a low attentional level. Note the similarity of the time course of the delay units to the experimental data shown in Figure 3A.

In order to demonstrate the variability of electrical responses in a single unit, Figure 5 shows the activity of sample units from areas V4, TEO/IT and prefrontal of the model during five trials of the task with low attention (Fig. 5A) and high attention (Fig. 5B). In the time bar along the bottom of the figure, stimuli are shown in black along the x-axis. In this example, we used the same stimulus for all five trials in order to demonstrate the variability that results from noise that is added during simulations with the model. As can be seen from the figure, a given delay (D1) unit may not be active during every delay period, and, as a consequence, the corresponding response units do not activate during these trials. However, because of the distributed nature of the representation in the prefrontal area of the model, a sufficiently high attention always results in a subgroup of units that are active during the delay in the entire population of 81 units, as shown in Figure 6. Thus, during a high attention state there is always a subgroup of units that respond when the test

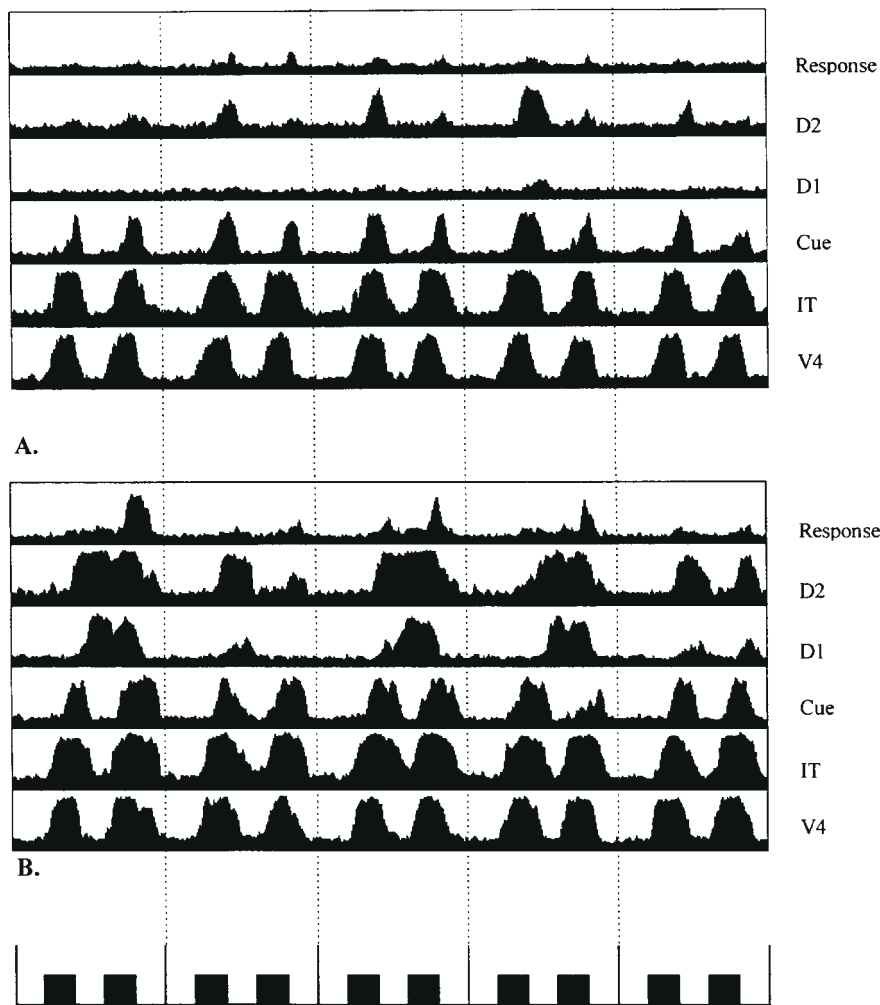


Figure 5. Time course of electrical activities of sample units from different areas of the model during five trials of a task. Stimuli are indicated by the filled blocks in the timeline at the bottom of the figure. Each block corresponds to 1 s of simulated time. Trials are separated by dashed lines. (A) Low attention level. Electrical activity of a sample unit from each area, including each type of frontal unit in the memory circuit. (B) High attention level. Electrical activity of the same sample units as in (A). Note the difference in activity patterns between the two types of delay units. Delay activity only occurs when the attention level is high, and the corresponding response unit shows a brief burst of activity only when the memory has been retained during the delay. During simulations, random noise is added at each iteration, yielding the variability in electrical responses seen in this figure.

shape matches the cue, while with low attention they almost never do. If the test stimulus does not match the cue, there is little or no activity in the response units, and the delay units decay to baseline level over time.

Figure 6 shows ‘snapshots’ of the electrical activity level of all units in each of the areas during the cue period (top), delay period (center) and test stimulus period (bottom), with low attention (Fig. 6A) and with high attention (Fig. 6B). In order to demonstrate the shift-invariance that the learning rule achieves (see Appendix B), the two identical stimuli are placed at different locations on the array during the cue period and the test period. Although there is little overlap in active units in areas V1/V2 and V4, the patterns in TEO/IT and the prefrontal area are very similar for the two stimuli.

rCBF Activations

The blood flows in the four areas of the model were computed by summing total synaptic activity among all subpopulations within each area. The results are shown in Table 3 along with experimental PET data for the 1 s delay task in Haxby *et al.* (1995). The table shows values for simulated rCBF activations in each of the modeled areas for both tasks, along with the

percentage differences between tasks in each area. The rCBF values in Table 3 are normalized to the activity in V1 during high attention in each case (modeled and experimental) so that direct comparisons can be made between the simulated and experimental data.

Although no direct connections from the prefrontal cortex to area V4 have been demonstrated in the monkey, we found that it was necessary to add such a connection in order to replicate the common finding in PET studies that in humans, an area which appears to be homologous to monkey V4 shows a substantial degree of enhancement during shape processing when there is an attentional component to the cognitive task (Corbetta *et al.*, 1991; Sergent *et al.*, 1992; Haxby *et al.*, 1994; Courtney *et al.*, 1996). We found that including only the prefrontal-to-TEO/IT connections, as suggested by the anatomical data from Felleman and Van Essen (1991), for example, results in much higher changes in blood flow in these more anterior areas than in V4 when a sufficient attentional effect was present to allow the delay activities to occur. However, in the study by Haxby *et al.* (1995), for example, V4 showed an ~8% increase in blood flow activity, while the TEO/IT and prefrontal areas both showed ~4% change. The other studies cited do not specify the exact amounts

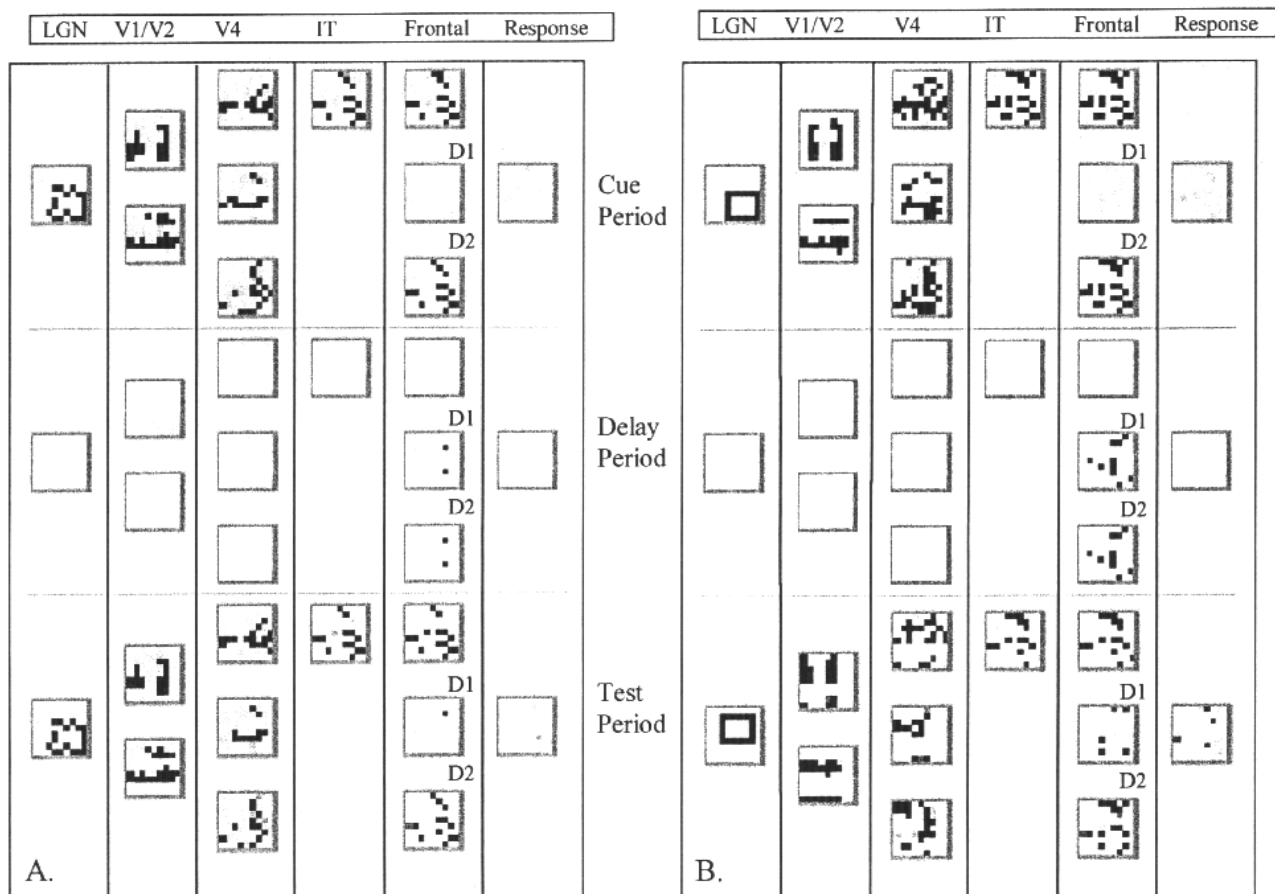


Figure 6. Snapshot frames of the electrical activities of all units in all areas in the model during a single time step at three different phases of the task. The LGN is modeled as a single set of units whose electrical activities are clamped into different patterns, such as the square shown here, and serves as input to the system. Darker units are more activated. Reading down in the prefrontal column, shown are cue units, D1 units and D2 units. Response units, acting as the decision/output of the matching task, are shown in separate columns on the right. (A) Low attention. Very little delay activity occurs. (B) High attention level. Note the delay activity in a subpopulation of the D1 and D2 units, and the response to the matching stimulus.

Table 3

Comparison of simulated blood flow and experimental PET data

	V1	V4	TEO/IT	Prefrontal
Simulated				
High attention	1.00	0.8958	0.8303	0.6964
Low attention	0.9696	0.8519	0.8103	0.6732
Percent change	3.1	5.2	2.5	3.5
Experimental				
High attention	1.00	0.8924	0.8050	0.7057
Low attention	0.9738	0.8258	0.7724	0.6776
Percent change	2.7	8.1	4.2	4.1

Data are normalized to the value of the high-attention V1 activations in each of the two cases (simulated and experimental). The experimental data are from that used in the paper by Haxby *et al.* (1995).

by which blood flow increased in the reported areas, but the maps of significant activation in a passive viewing task vs. shape matching generally show a greater spread of activity in posterior areas along this pathway than in anterior ones. McIntosh *et al.* (1994), in their path analysis, also needed to add a connection from the prefrontal area to V4 in order to explain the results of a similar visual task. The prefrontal-to-V4 link may reflect an indirect path via subcortical areas, which has been implicated in

attentional processing (e.g. Alexander *et al.*, 1986). In our simulations, we have shown that a diffuse projection via this pathway may enhance blood flow in V4 without specifically encoding particular stimuli.

Discussion

The goal of this work has been to identify issues that are relevant to modeling quantitative brain imaging data with large-scale networks of neuronal-like elements. Matching such data imposes the constraint that the parameters, responses and connectivity of the network be both quantitatively and qualitatively similar to experimental data. The major factors that we have focused on are:

1. Connectivity:

- The importance of the balance of local connectivity patterns, with the excitatory-to-excitatory connections being the most significant source of local synaptic activity and afferents making a relatively small contribution.
- Constraining the total connectivity strengths between different areas to approximate the published data for the known connections.
- Using a systematic rule-based method for determining connection topography between areas. This achieves the

same type of receptive field properties that have been shown to exist in the ventral visual pathway and makes it possible to identify the effects of rules, such as increasing divergence, on both electrical and rCBF responses.

2. Temporal dynamics: the network must behave appropriately during a sequence of different phases such as those that are involved in each of the tasks. We used the different stages of a delayed match-to-sample task and demonstrated a model of working memory that could mediate such a task and show appropriate behaviors during each stage.
3. The need for having a network that can perform several different tasks of the kind that are commonly used in imaging studies. The model described here shows how matching can be mediated by a single parameter that corresponds to attention, and, in particular, suggests that the attentional component affects a particular type of neuron that responds to both a stimulus and during the delay, the D2 unit, which has been identified in delayed-response tasks in the monkey.

Highly interconnected dynamical systems with recurrent feedback, such as the cortex, can be quite sensitive to minor perturbations in even a single area. If a model exists for the tasks that are involved, these effects can be examined by simulation and/or analysis of the model network and by relating the observed changes to potential cognitive differences. Modeling can help to identify the salient components that can affect experimental brain imaging data, since it affords the ability to test effects of different components of imaging studies, such as the assumed connectivity and temporal dynamics.

For example, because of the noise in imaging data, statistical methods that are commonly applied to the subtraction paradigm may not identify all areas involved in a task, even though they may actually be activated. This is particularly acute, since in many cases the differences in blood flow that are seen in PET studies of cognition are small, <5%, and even minor variance can cloud the results. Methods that make use of the variability in rCBF data have been applied to data without subtraction. This has taken the form of correlation analysis (Horwitz *et al.*, 1992a,b), structural equation modeling (McIntosh and Gonzalez-Lima, 1994; Horwitz *et al.*, 1995) and principal component or eigenimage analysis (Friston, 1994). However, there are still a number of unanswered questions about the interpretation of such results. For example, it is likely that under some conditions, the same area mediates two different tasks, but in different ways. Area V2 is thought to be involved in processing both shape and color, but by separate populations of neurons that appear to be interleaved with each other at a resolution below that of PET data (Roe and Ts'o, 1995). This issue is also relevant in interpreting the role of the prefrontal area. If there are different populations that mediate different aspects of cognitive activity at different phases of a task, as in matching, then if such populations are located close to each other, imaging may not reflect changes due to one population vs. another. We have shown that modeling can help to elucidate the distinctive roles of such populations.

The circuit that implements the working memory in the model (Fig. 4) is local to the prefrontal cortex. The fact that the behaviors of the different subpopulations of units in the model are similar to spiking activities in the analogous neuron types in prefrontal cortex (see Fig. 3) suggests that a similar circuit may mediate the task in the prefrontal cortex. Delay-active units evidently occur mostly in the deep layers (V and VI) (Fuster, 1973) and do not appear to be localized to specific areas spatially (Fuster *et al.*, 1982). This suggests that the circuit depicted in

Figure 4 might form a local assembly, such as a cortical column, in the prefrontal cortex. Such circuits in the prefrontal cortex might also imply that local connectivity is more dense in the lower layers in this area than in areas further back along this pathway.

In summary, we have defined a basic local unit that represents a cortical column, and extended it to build a large-scale network with a local memory circuit that models neuronal behaviors in the prefrontal cortex during different phases of a cognitive task. For connecting the different areas, we used some basic principles for creating connection patterns that give the types of tuning commonly seen in the different areas, such as orientation selectivity in V1/V2 and the widening of receptive fields in the V1 to IT pathway. The goal was not to specifically model such behaviors, but rather to identify a modeling framework for interpreting human blood flow data. Given the increasing number of functional neuroimaging studies being performed, such a method is useful for integration of the extensive literature on animal studies to help identify key factors that will allow human imaging data to be better understood.

Notes

We would like to thank Dr Stanley Rapoport for support during development and testing of the model and Ankoor Shah for his dedicated help in programming. We are grateful to Drs Susan Courtney and Edward Bullmore for their careful reading of the manuscript and for their helpful comments.

Address correspondence to M.-A. Tagamets, GICCS, Room EP04, New Research Building, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007-2197, USA. Email: malle@giccs.georgetown.edu.

References

- Ackermann RF, Finch DM, Babb TL, Engel J Jr (1984) Increased glucose metabolism during long-duration recurrent inhibition of hippocampal pyramidal cells. *J Neurosci* 4:251-264.
- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357-381.
- Arbib MA, Bischoff A, Fagg AH, Grafton ST (1995) Synthetic PET: analyzing large-scale properties of neural networks. *Hum Brain Map* 2:225-233.
- Burkhalter A, Bernardo KL (1989) Organization of corticocortical connections in human visual cortex. *Proc Natl Acad Sci USA* 86:1071-1075.
- Chelazzi L (1995) Neural mechanisms for stimulus selection in cortical areas of the macaque subserving object vision. *Behav Brain Res* 71:125-134.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1991) Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J Neurosci* 11:2383-2402.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1996) Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb Cortex* 6:39-49.
- Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57:835-868.
- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051-2062.
- DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J (1996) Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci USA* 93:2382-2386.
- Doty RW (1983) Nongeniculate afferents to striate cortex in macaques. *J Comp Neurol* 218:159-173.
- Douglas RJ, Martin KAC (1991) A functional microcircuit for cat visual cortex. *J Physiol* 440:735-69.
- Douglas RJ, Koch C, Mahowald M, Martin KAC, Suarez HH (1995) Recurrent excitation in neocortical circuits. *Science* 269:981-984.

- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1-47.
- Friston KJ (1994) Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Map* 2:56-78.
- Funahashi S, Kubota K (1994) Working memory and prefrontal cortex. *Neurosci Res* 21:1-11.
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1990) Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *J Neurophysiol* 63:814-831.
- Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J Neurophysiol* 36:61-78.
- Fuster JM (1990) Inferotemporal units in selective visual attention and short-term memory. *J Neurophysiol* 64:681-697.
- Fuster JM, Bauer RH, Jervey JP (1982) Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. *Exp Neurol* 77:679-694.
- Gallant JL, Braun J, Van Essen DC (1993) Selectivity for polar, hyperbolic and Cartesian gratings in macaque visual cortex. *Science* 259:100-103.
- Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14:477-485.
- Haenny PE, Maunsell JH, Schiller PH (1988) State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Exp Brain Res* 69:245-259.
- Haxby JV, Grady CL, Horwitz B, Ungerleider LG, Mishkin M, Carson RE, Herscovitch P, Schapiro MB, Rapoport SI (1991) Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc Natl Acad Sci USA* 88:1621-1625.
- Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady CL (1994) The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *J Neurosci* 14:6336-6353.
- Haxby JV, Ungerleider LG, Horwitz B, Rapoport SI, and Grady CL (1995) Hemispheric differences in neural systems for face working memory: a PET-rCBF study. *Hum Brain Map* 3:68-82.
- Horwitz B, Sporns O (1994) Neural modeling and functional neuroimaging. *Hum Brain Map* 1:269-283.
- Horwitz B, Grady CL, Haxby JV, Schapiro MB, Rapoport SI, Ungerleider LG, Mishkin M (1992a) Functional associations among human posterior extrastriate brain regions during object and spatial vision. *J Cognit Neurosci* 4:311-322.
- Horwitz B, Soncrant TT, Haxby JV (1992b) Covariance analysis of functional interactions in the brain using metabolic and blood flow data. In: *Advances in metabolic mapping techniques for brain imaging of behavioral and learning functions* (Gonzalez-Lima F, Finkenstadt T, Scheich H, eds), pp. 189-217. Dordrecht: Kluwer.
- Horwitz B, McIntosh AR, Haxby JV, Furey M, Salerno JA, Schapiro MB, Rapoport SI, Grady CL (1995) Network analysis of PET-mapped visual pathways in Alzheimer type dementia. *NeuroReport* 6:2287-2292.
- Hubel DH, Wiesel TN (1977) Functional architecture of macaque visual cortex. *Proc R Soc Lond B* 198:1-59.
- Jueptner M, Weiller C (1995) Review: does measurement of regional cerebral blood flow reflect synaptic activity? – Implications for PET and fMRI. *NeuroImage* 2:148-156.
- Linsker R (1988) Towards an organizing principle for a layered perceptual network. *Neural Informat Process Syst* 485-494.
- McIntosh AR, Gonzalez-Lima, F (1994) Structural equation modeling and its application to network analysis in functional brain imaging. *Hum Brain Map* 2:2-22.
- McIntosh AR, Grady CL, Ungerleider LG, Haxby JV, Rapoport SI, Horwitz B (1994) Network analysis of cortical visual pathways mapped with PET. *J Neurosci* 14:655-666.
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16:5154-5167.
- Peterhans E, von der Heydt R (1993) Functional organization of area V2 in the alert macaque. *Eur J Neurosci* 5:509-24.
- Posner MI, Petersen SE, Fox PT, Raichle ME (1988) Localization of cognitive operations in the human brain. *Science* 240:1627-1631.
- Rockland KS, Van Hoesen GW (1994) Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cereb Cortex* 4:300-313.
- Roe AW, Ts'o DY (1995) Visual topography in primate V2: multiple representation across functional stripes. *J Neurosci* 15:3689-3715.
- Rosenkilde CE, Bauer RH, Fuster JM (1981) Single-cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Res* 209:375-394.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889-893.
- Sergent J, Ohta S, Macdonald B (1992) Functional neuroanatomy of face and object processing: a positron emission tomography study. *Brain* 115:15-36.
- Spitzer H, Desimone R, Moran J (1988) Increased attention enhances both behavioral and neuronal performance. *Science* 240:338-340.
- Sporns O, Tononi G, Edelman GE (1991) Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. *Proc Natl Acad Sci USA* 88:129-133.
- Tagamets M-A (1994) Self-organization of spatio-temporal behaviors in an oscillatory neural network. CS-TR-3237, University of Maryland Computer Science Department, College Park, Maryland.
- Tagamets M-A, Horwitz B (1997a) Modeling brain imaging data with neuronal assembly dynamics. In: *Computational neuroscience: trends in research 1997* (Bower JM, ed.), pp. 949-953. New York: Plenum Press.
- Tagamets M-A and Horwitz B (1997b) A computationally based account of GABA_A agonist in human brain imaging studies. *NeuroImage* 5:5386.
- Tagamets M-A, Horwitz B, Reggia JA (1995) A large-scale neural model linking local neuronal dynamics to positron emission tomography (PET)/regional cerebral blood flow (rCBF) data. *Soc Neurosci Abstr* 21:1988.
- Tanaka K (1993) Neuronal mechanisms of object recognition. *Science* 262:685-688.
- Tanaka M, Lindsley E, Lausmann S, Creutzfeldt, OD (1990) Afferent connections of the prelunate visual association cortex (areas V4 and DP). *Anat Embryol* 181:19-30.
- Tononi G, Sporns O, Edelman GE (1992) Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. *Cereb Cortex* 2:310-335.
- Ts'o D, Gilbert CD, Wiesel TN (1986) Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J Neurosci* 6:1160-1170.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Ingle J, Goodale MA, Mansfield RJW, eds), pp. 549-586. Cambridge, MA: MIT Press.
- Wilson FAW, Ó Scalaidhe SP, Goldman-Rakic PS (1993) Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955-1958.
- Wilson FAW, Ó Scalaidhe SP, Goldman-Rakic PS (1994) Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proc Natl Acad Sci USA* 91:4009-4013.
- Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12:1-24.

Appendix A

Between-region Connection Weight Patterns

Connections between areas of the model are generated by an automated procedure which allows one to specify the fanout, the average weight value, the range of variability around this average, and the density, i.e. percentage of the connections to actually construct. Thus, for example, specifying a fanout of 4×4 with values 0.02 ± 0.01 and 50% density will result, on average, in an array of eight connections from unit i to unit j (i.e. 50% of 16) that have a mean value of 0.02 and that vary from 0.01 to 0.03. Thus the total value, on average, of a connection from the source to the destination will be $8 \times 0.02 = 0.16$.

Table A1 gives the fanout, mean value, variability, density (shown as percentage of connections to make), and the mean total weight afferent to the destination area (computed as described above) of all the between-area connections in the model.

Appendix B: The Learning Rule

In order to achieve electrical response characteristics that are similar to those observed in monkey TEO/IT, i.e. shift invariance and tuning for specific shapes, we use an unsupervised competitive Hebbian learning rule to train the feed-forward connections from area V4 to area TEO/IT.

Table A1

Connectivity among regions of the model (unless otherwise noted, connections are E-to-E)

From	To	Fanout	Mean value and variability	Percent to create	Net total afferents	Comments
LGN	V1	7 × 7	37 @ 0.003 ± 0.003 2 × 8 @ 0.006 ± 0.003 1 × 4 @ 0.02 ± 0.002	100	0.239	highest values oriented either vertically or horizontally
V1	V4	1 × 5	0.04 ± 0.01	50	0.100	
V4	TEO/IT	5 × 5	0.0028 ± 0.0028	50	0.105	learned
TEO/IT	PF	1 × 1	0.2 ± 0.02	100	0.2	
PF (D2)	V4	5 × 5	0.0014 ± 0.0007	100	0.035	
PF (D1)	TEO/IT	1 × 1	0.01 ± 0.002	100	0.01	
PF (D2)	TEO/IT	1 × 1	0.03 ± 0.001	100	0.03	to inhibitory
TEO/IT	V4	4 × 4	0.00125 ± 0.0006	100	0.02	

The competition is achieved by normalizing all of the efferent weights from each source unit and rescaling to a specified total connectivity strength (Tagamets, 1994). It is this connectivity strength that determines the proportion of weights between the two areas. Specifically, the learning rule is given by the equations:

$$\tilde{w}_{ji}(t+1) = w_{ji}(t) + d \cdot \left[s[a_j(t) - T_{\text{pre}}] s[a_i(t) - T_{\text{post}}] \right] \quad (\text{A1})$$

which establishes the amount by which the weight w_{ji} is to change, and

$$w_{ji}(t+1) = C \frac{\tilde{w}_{ji}(t+1)}{\sum_k \tilde{w}_{ki}(t+1)} \quad (\text{A2})$$

which performs the normalization and scaling of the weights.

s is the threshold function:

$$s(a - X) = \begin{cases} a & \text{if } a - X > 0 \\ 0 & \text{otherwise} \end{cases}$$

$a_i(t)$ and $a_j(t)$ are the electrical activities of units i and j respectively, and C is the scaling factor. d is the learning rate and T_{pre} and T_{post} are the presynaptic and postsynaptic learning thresholds respectively. The values that we used for these parameters are as follows: $C = 0.035$, $d = 0.005$, $T_{\text{pre}} = 0.7$ and $T_{\text{post}} = 0.15$. Since each of the three V4 populations connects to the single TEO/IT region, the total weight into TEO/IT is $3 \times C = 0.105$, as shown in Table A1.

The competition comes from the normalization step, equation (A2). A particular weight which was not changed in equation (A1) because either the pre- or the postsynaptic unit failed to reach the learning threshold will decrease during normalization if at least one other weight increased. This learning rule, then, in effect implements a redistribution of weights among efferents from a unit.

For the simulations performed here, the network is trained with two distinct shapes, a square and a T shape. Each shape is presented to the

Table A2

Connectivity of the working memory circuit

Source	Destination	Destination element	Weight
Cue	D2	excitatory	0.07
Cue	response	excitatory	0.05
D1	response	excitatory	0.06
D1	D2	excitatory	0.105
D2	D1	excitatory	0.10
D1	cue	inhibitory	0.02
Cue	D1	inhibitory	0.05
Response	D1	inhibitory	0.03
Response	D2	inhibitory	0.065

LGN area for 25 iterations, during which the shape is shifted in location by two pixels on the LGN array every five iterations. The shifting stimulus during learning allows the TEO/IT area to learn shift-invariant representations of the different stimuli. The two stimuli are alternated, with blank periods of 25 iterations between presentations, for a total of eight times each. Learning is continuous during the training period, taking place at each iteration during the simulation.

After training, the IT area has a well-tuned distributed response to each of the training stimuli and weaker, less tuned responses to other, unlearned stimuli. The trained network is then used for the match-to-sample PET simulation described in the main text of the paper.

Appendix C: The Working Memory Circuit: Connection Strengths

Each element of the working memory circuit shown in Figure 4 is a basic unit, such as that shown in Figure 1A, and is made up of one excitatory element and one inhibitory element. All inhibitory connections within the working memory circuit, shown as arcs with circles at the destination units, are excitatory synapses onto the inhibitory elements of the destination units. All connections are single (i.e. with a fanout of 1×1), with the weight values shown in Table A2.