1989

A Simple Model of Prefrontal Cortex Function in Delayed-Response Tasks

Stanislas Dehaene

Laboratoire de Sciences Cognitives et Psycholinguistique, France

Jean-Pierre Changeux

Département des biotechnologies
Institut Pasteur, France

Abstract

Both psychologists and neurobiologists have used delayed response (DR), $A\overline{B}$, and delayed matching-to-sample (DMS) tasks as tools to study the functions of prefrontal cortex in primates and humans. We describe a simulation model that relates behavioral and electrophysiological data relevant to these tasks into a minimal neural network.

The inputs to the network are two visual objects and a positive or negative reinforcement signal. As the output, the network orients toward one of the two objects. We subdivide the architecture of the network into two levels, both of which embody constraints from neuroanatomy in a simplified form. Level 1 consists of a sensory-motor loop with modifiable synaptic weights and provides a capacity for grasping. Level 2 contains memory and rule-coding units and modulates the lower level 1. When level 1 only is simulated, the network fails to learn the tasks. The errors made by the network resemble those of young monkeys, infants, or adults with prefrontal lesions. In particular, the systematic $A\overline{B}$ error can be reproduced. With level 2 on top of level 1, the network acquires systematic rules of behavior by mere reinforcement and rapidly adapts to changes in the reinforcement schedule. Learning takes place by selection among a repertoire of possible rules. The properties of the model are discussed in terms of actual behavioral and physiological data, and several critical experimental predictions are presented. In particular, we address the issues of prefrontal functions, "systematicity" in neural networks, and "mental Darwinism."

Introduction

The brain can be viewed as a non-equilibrium system in evolution which constantly undergoes changes in its internal organization as a consequence of interactions with its environment. The general "Darwinian" scheme of variation-selection has often been used to model such epigenetic modifications, though on the basis of diverse elementary principles (Taine 1870; Jerne 1967; Changeux et al. 1973; J.Z. Young 1973; Changeux and Danchin 1976; Edelman 1978, 1987; Piatelli-Palmarini in press). To adequately relate a given function to the relevant neural structure, any model must deal with the hierarchy of levels that may be delimited within the nervous system, from molecules and cells up to the cognitive levels (Changeux and Dehaene in press). Several of the selectionist models concerned lower-level circuits such as the epigenesis of the neuromuscular junction (Gouzé et al. 1983), the development of afferent innervation of sympathetic ganglia (Purves and Lichtman 1980), or the evolution of the climbing fiber contacts on Purkinje cells (Mariani and Changeux 1981a, 1981b). Others dealt with higher brain functions on both theoretical (Edelman 1978, 1987) and empirical (Merzenich et al. 1988) grounds. Recently, at the level of cell assemblies, the learning of stable configurations of neuronal activities (Toulouse et al. 1986)

and of temporal sequences (Dehaene et al. 1987) have been modeled in terms of pruning an initial repertoire of "prerepresentations" through interactions with the outside world.

In the present article, we examine the usefulness of the notion of learning by selection at a higher cognitive level, namely the acquisition of elementary rules of behavior in the prefrontal cortex. We examine first the relevant psychological and neurobiological data and attempt to account for them by a minimal set of theoretical hypotheses, cast in the form of a simple neural network. Two aspects of this modeling are worth emphasizing. First, higher cognitive functions cannot be accounted for irrespective of their relations with the lower levels; accordingly, we will not model an isolated cortical area, but a highly simplified "organism" containing a multilevel neuronal network and interacting with its environment. Second, according to Fodor and Pylyshyn (1988) and Pinker and Prince (1988), high-level "rational" processes are not adequately described by rules of association; rather, higher cognitive functions like language seem to be modeled better by rules of symbol manipulation. Accordingly, and although we will deal with a much more primitive level than that of language acquisition, our model has been designed to address such general questions as the representation of rules and variables in neural networks. We will see that a simple selectionist model can acquire systematic rules by mere reinforcement.

Biological Premises

We shall focus our discussion on a limited set of data concerning the so-called *delayed-response tasks*,[1] for which a number of psychological and neurobiological observations are available in the literature.

Description of Delayed-Response Tasks

Delayed-response tasks have been used with human adults or babies and intact or lesioned monkeys (for review see Diamond 1988; Fuster 1980, 1984; Milner and Petrides 1984). The basic scheme is always identical.[2] The subject is first presented with a cue object, followed by a delay period of variable duration. Then, two objects are presented simultaneously at two separate locations and the subject must choose one of them. The rules defining the correct choice may vary, and specific tasks may be defined. In the $A\overline{B}$ and specific delayed response (DR) tasks,[1] the rule is to choose the object that stands at the position of the cue object before the delay. In delayed matching-to-sample (DMS), one must choose the object identical to the cue, irrespective of its position. Finally, a third task, delayed alternation (DA), might be considered to fall in the same framework (Fuster 1984); here, one must alternate responses to one location and the other; thus, the subject's previous response can be considered as the "cue" guiding behavior at the current trial.

Formally, many abilities are required in order to succeed in delayed-response tasks. First, all tasks are *sensorimotor* in that they require the subject "to perform motor acts in accord with events in the recent past" (Fuster 1984, p. 408). Second, they tax *short-term memory*, since characteristic features of the cue must be memorized throughout the delay period. Third, the tasks require the subject to focus on one particular feature of the cue (e.g. color) and to neglect the others (e.g. location); thus, some *selective attention* is necessary, at least in DMS and DA. In $A\overline{B}$ and DR, selective attention seems less critical, since the choice items vary only in location, not in the irrelevant dimension of color. Fourth, during the test phase, a decision must be made on the basis of a *comparison* of the test objects with the stored representation of the cue. Finally, learning a task is a process of *induction* through time and space: the subject must discover which features are relevant and what rule governs the application of reinforcement. Hence, despite their simplicity, delayed-response tasks appear to involve several high-level cognitive functions.

Behavioral Observations in Intact and Lesioned Subjects

Piaget's $A\overline{B}$ task (Piaget 1954) has been extensively used with human infants between 6 and 12 months. Likewise, Jacobsen's DR task (Jacobsen 1935) has become a classical test of prefrontal function in primates. The two tasks differ only minimally: the within-trial procedures are identical, both requiring the subject to reach, after the delay, at the location where the cue was first presented. However, in DR, the location of the cue is changed randomly from one trial to the next, whereas in $A\overline{B}$ the location is changed only after a criterion of success at that location has been reached.

Diamond has studied and compared these tasks extensively with human infants, infant monkeys, and intact or lesioned adult monkeys (Diamond 1985; Diamond and Doar 1989; Diamond and Goldman-Rakic 1989; Diamond et al. 1989; for review, see Diamond 1988). In short, infants older than 12 months and adult monkeys with preserved prefrontal cortex can learn both tasks. In contrast, human infants from 7 1/2–9 months, infant monkeys from 1 1/2–4 months, and frontally lesioned monkeys fail in a similar manner. Although they succeed in the absence of a delay, their performance deteriorates with delays as brief as 1–2 seconds. In $A\overline{B}$, the pattern of errors is peculiar. At long delays, performance is essentially random; but at an intermediate duration, a systematic error occurs: the subject is able to reach the success criterion at the initial location (A), but he continues to reach to A even after the location has been switched to B. The duration of delay for which this systematic $A\overline{B}$ error occurs increases with age in human infants (Diamond 1985). Such perseveration is similar to the performance of human frontal patients, as revealed by the Wisconsin Card Sorting test (Milner 1963). This test requires subjects to sort cards ac-

cording to a given criterion (e.g., color, shape, number). Frontal patients learn to sort the cards, but persevere in using the same criterion even after the sorting criterion was changed without notice by the experimenter.

DMS and DA have not been the focus of such an extensive behavioral study. However, solid electrophysiological evidence, examined in the next section, confirms their similarity with $A\overline{B}$ and DR and their links to prefrontal cortex.

Electrophysiological Recordings

Single-unit recordings in the behaving monkey have revealed several types of prefrontal units that are active during delayed-response tasks (review in Fuster 1980, 1984). Many units are driven by *sensory* aspects of the cue (Niki and Watanabe 1976a). However, they appear not to respond passively to sensory features. Rather, stimuli cause prefrontal firing only if they draw attention because they are relevant to behavior; cells have even been reported to respond to the presence (or the absence) of an expected reinforcement (Niki and Watanabe 1979). A second type of units codes for *intended movements*, responding differentially to a cue depending on the direction of the movement it elicits (Niki 1974; Niki and Watanabe 1976a, 1976b, 1979).

Both sensory and movement types of cells may or may not keep a sustained firing during the delay period, when no stimulus is physically present. Some sensory units show a steady or slowly decreasing firing rate during the delay, suggesting they are involved in *short-term* memory for the characteristics of the cue (Niki 1974). Other cells discriminate the direction of the planned movement and show a progressive increase in activity during the delay (Niki and Watanabe 1976a; Fuster et al. 1982); they appear to be involved in *response anticipation*, and their activity predicts occasional errors that the animal makes (Niki and Watanabe 1976a; Watanabe 1986).

Limited information is available about the anatomical organization of these units in prefrontal cortex. The cells described above are found essentially in the principal sulcus (during spatial delayed-response tasks) and in the inferior convexity and the orbital prefrontal cortex (during delayed identity-matching tasks) (Goldman-Rakic 1988). These regions show a segregation of afferents and efferents into separate modules (Goldman-Rakic and Schwartz 1982; Goldman-Rakic 1984) and some hints of a columnar organization (Fuster et al. 1982; Goldman-Rakic 1984). However, their subdivision into smaller functional territories has not yet been achieved.

In short, a significant amount of behavioral and neurobiological data is available about the delayed-response tasks; this justifies an attempt to embody the two sets of data into a minimal model which will now be presented in formal terms. The goals of the model are twofold: first, to provide a simple, coherent and biologically plausible (even if simplified) account of the available data; and second, to lead to critical predictions that may be tested experimentally and may open new avenues for future research.

Description of the Model

This section is divided into four parts. First, we describe the environment of the "formal organism" and the tasks that it must learn. We then describe the basic components of the neuronal network, its architecture, and the learning rule used.

Environment of the Formal Organism

I. Objects: The environment consists of objects defined by their features along a number of dimensions, such as position, color, shape, and size.

Limitations:

1. In this minimal model, only two dimensions are considered, with two features values along dimension 1 and three along dimension 2. Dimension 1 may be likened to position, with left and right as possible features, and dimension 2 to color, with three possible hues.
2. Since only two positions are used, the model does not take into account experiments where the relative position of objects is varied (e.g. Niki 1974) or $A\overline{B}$ experiments with multiple hiding wells (e.g. Cummings and Bjork 1983).
3. Two objects at most can be presented at a given time.

II. Reinforcement Signal: Positive reinforcement or punishment is provided to the system under conditions defined in paragraph V.

Formalization:

A reinforcement parameter r belongs to the interval $[-1, +1]$. 0 is neutral (no reinforcement). +1 and −1 correspond respectively to maximal positive reinforcement and maximal punishment.

III. Tasks: Objects and reinforcement are presented in tasks composed of several trials. A trial in a task involves four steps: (1) presentation of one object (cue period); (2) delay period without any object presented; (3) presentation of two objects; if the organism chooses one of them, the appropriate reinforcement or punishment is delivered (choice period); (4) inter-trial period.

IV. Rules Defining the Correct Choice: During the choice period, the organism is given an opportunity to choose one of the two objects presented. Which object should be chosen depends on the type of trial. In a type 1 trial, the correct object is the one whose feature along dimension 1 matches the object presented during the cue period. In a type 2 trial, the correct object is the

one that matches the cue object along dimension 2. The organism must discover which type of trial defines the actual conditions for reinforcement.

Type 1 trials require choosing the object that stands at the same position as the cue object, independent of its color. Therefore, a sequence of type 1 trials may be considered an analogue of $A\overline{B}$ or DR tasks.[3] The analogue of DR is to present type 1 trials with a randomly varying position for the cue object. The analogue of $A\overline{B}$ is to present, first, type 1 trials with the cue always at position 1, until a success criterion is reached, then to present the cue always at position 2.

Similarly, type 2 trials require the choice of the object whose color matches the color of the cue, independent of position. A sequence of type 2 trials can thus be considered an analogue of the DMS task. The organism can be trained with only two colors for the cue and later tested for generalization to a third color.

V. Conditions for Reinforcement: During the cue period, when the organism orients toward the cue object, small positive reinforcement is given.[4] During the choice period, when the correct test object is selected, medium positive reinforcement is given; strong punishment is provided if the wrong object is chosen.

Limitation:
The asymmetry between correct and wrong choices is required in order to ensure that the organism will always seek an optimal choice strategy. When we treat correct and wrong choices symmetrically, we find that the network sometimes stabilizes into a suboptimal strategy (e.g., a correct choice on 75% of trials only).
Formalization: During reinforcement, the parameter r takes one of three possible values $r_{\text{small}+}$, $r_{\text{medium}+}$, $r_{\text{large}-}$.

Components of the Neural Network

We use the formalism developed by Dehaene, Changeux, and Nadal (1987). At the lowest level, the network is composed of formal neurons (threshold units) linked by synapses. However, individual neurons and synapses are not explicitly modeled. The functional units are clusters of synergic neurons linked by bundles of synapses.

Units: The building blocks of the network are clusters of neurons that are densely interconnected by excitatory synapses. The activity of a cluster is defined as the fraction of active neurons in the cluster. Due to the positive feedback established by synapses within a cluster, each cluster possesses two stable states of activity (most neurons active or most neurons resting). Thus, in contrast to isolated neurons, clusters can maintain a self-sustained activity, providing an elementary form of memory or remanence in the network (Dehaene et al. 1987).
Formalization: The activity of a cluster i is represented by a continuous variable $S_i(t)$ between 0 (resting) and 1 (active).

Connections: Individual synapses are grouped into bundles which propagate activity from a given anterior cluster to another posterior cluster. We distinguish two types of such bundles. *Static bundles* have a fixed (excitatory or inhibitory) efficacy which is not modified during learning. *Modulated bundles* are bundles whose efficacy can be modified, up to a given maximum, by the activity of a third neuronal cluster called modulator; their maximum efficacy is modified through learning. The mechanism postulated for the modification of efficacy is an heterosynaptic modulation: a given synapse of neuron A on neuron B can be influenced by the activity of a second, neighboring synapse upon neuron B, originating from the modulator neuron C; the triplet A, B, C is called a synaptic triad (Dehaene et al. 1987). Allosteric transitions of post-synaptic receptor molecules at the A-B synapse provide a plausible molecular implementation of synaptic triads (Heidmann and Changeux 1982; Changeux and Heidmann 1987).
Formalization: Static connections are represented by a matrix $(V_{ij}(t))$, giving the efficacy of each connection from cluster j to cluster i. Self-connections within clusters are represented by the terms V_{ii}, which have a fixed positive value. Synaptic triads b are triplets of clusters (anterior, posterior, and modulator); they have an instantaneous efficacy $W_b(t)$, which varies between 0 and a maximum value $W_b^m(t)$. The cluster activities are updated according to

$$S_i(t+1) = F\Big[\sum_j V_{ij}(t)S_j(t) + \sum_{b=(a,i,m)} W_b(t)S_a(t) + N\Big]$$

where F is the sigmoid function

$$F(x) = \frac{1}{1+e^{-x}}$$

and N is a noise term with uniform distribution over $[-n, n]$.[5] The instantaneous efficacies $W_b(t)$ of each bundle are themselves updated according to the level of activity of the modulator cluster m:

$$W_b(t+1) = \begin{cases} \alpha_p W_b(t) + (1-\alpha_p)W_b^m(t), & \text{if } S_m(t) > 0.5 \\ \alpha_d W_b(t), & \text{if } S_m(t) < 0.5 \end{cases}$$

where α_p and α_d are constants between 0 and 1.

Architecture of the Network

We distinguish two levels of organization in the network (Figure 1). Each level is further divided into two

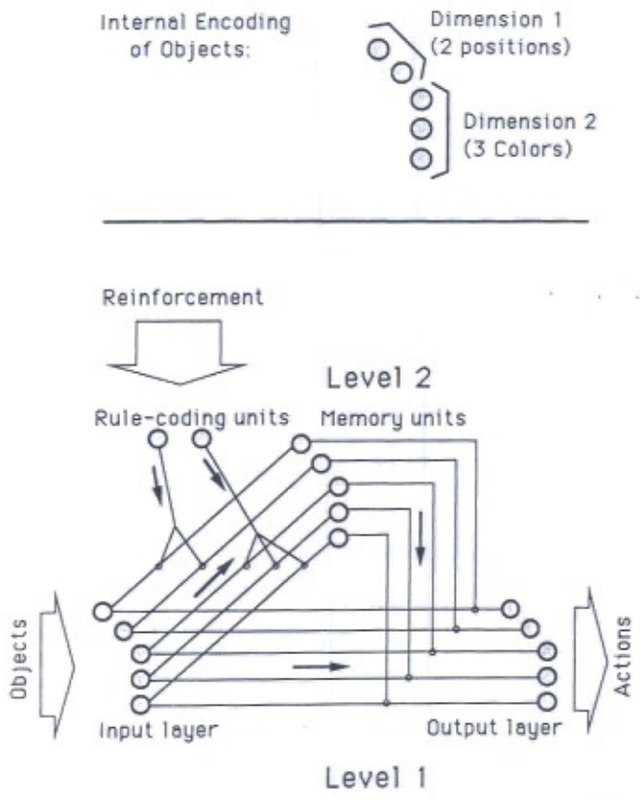

Figure 1. Outline of the formal neural network which learns analogue of the $A\bar{B}$, DR and DMS tasks. At the input, the external world is coded along only two dimensions: position and color (top). The output of the network indicates toward which object the network orients. Level 1 input-output connections provide the model with a capacity for grasping any isolated object. Level 2 activity modulates this grasping behavior according to specific rules. Learning is governed by a reinforcement signal which governs both the occurrence of random modification of rule clusters activity and the modifications of synaptic efficacies.

layers of neuronal clusters; level 1 is composed of the input and output layers, and level 2 contains a memory layer and a rule-coding layer. By analogy with primate cortex, level 1 would correspond to a visuomotor loop linking secondary visual areas to motor or premotor cortex, and level 2 to the prefrontal cortex.

Architecture of Level 1

Input Layer: Objects are decomposed into their features at the level of the input layer. Each possible feature along each dimension is coded by an individual neuronal cluster.

Limitations:

1. In the present model there are only 2 clusters for dimension 1 and 3 clusters for dimension 2. The presentation of one object activates two clusters, one along each dimension.
2. No attempt is made to solve the feature-conjunction problem. Thus when two objects are simultaneously presented, the network does not know how to associate dimension 1 features with dimension 2 features.
3. The issue of functioning and development of these feature detectors is not addressed in this paper.

Output Layer: Activity among output units commands the orientation of the organism toward one of the objects presented. Output code is isomorphic to input code: direct commands exist for orienting toward objects with any given feature. Output clusters inhibit each other in a "winner-take-all" network.

Limitations:

1. At the present abstract level of characterization, we do not distinguish between covert orientation of attention and overt orientation with the eyes or with the hand.
2. No distinction is made between orienting toward a given position (dimension 1 feature) and orienting toward a given color (dimension 2 feature), even though the distinction appears psychologically relevant (e.g., Tsal and Lavie 1988). However, our hypothesis of a direct command for orienting towards objects of a given color is consistent with psychological observations of "pop-out" from color (Treisman and Gelade 1980).
3. The mechanisms that convert such "orient-toward-color" commands into a motor code are not considered here.

Connectivity of Level 1:

Input and output layers are connected by topographic projections preserving the isomorphism of input and output codes.

Limitation:

A simple one-to-one mapping is used in this minimal model.

Architecture of Level 2

Memory Layer: This layer of self-excitatory, mutually inhibitory, neuronal clusters is organized isomorphically to input and outputs.

Rule-coding Layer: In this layer, neuronal clusters code each for a dimension of the input (why these clusters may reasonably be called "rule-coding" will become clear in the next section, where the behavior of the network is analyzed).

Connectivity of Level 2

1. The memory layer receives topographic projections from the input layer. These projections consist of bundles which are modulated by rule-coding layer activity. Each rule-coding cluster, coding for dimension x, modulates all the connections corresponding to the different possible features of dimension x (see Figure 1).
2. The memory layer projects topographically and modulates the input-output connections in a one-to-one fashion.

Learning

The reinforcement signal is internalized and plays two roles: first, it enters into a Hebbian rule for the modification of synaptic triads; second, it increases the production of spontaneous random changes of activity in rule-coding clusters.

Internal Representation of Reinforcement: The discrete input reinforcement r is internally averaged into a parameter R representing the satisfaction of the organism.

Formalization:

R is a continuous variable ranging from -1 (fully dissatisfied) to $+1$ (fully satisfied). R is updated after each phase (e.g., cue, delay) of each trial depending on the value of the input reinforcement r:

$$R(t+1) = \begin{cases} (1+r)R(t) + r & \text{if } r \in [-1, 0] \\ (1-\alpha)R(t) - \alpha & \text{if } r = 0 \\ (1-r)R(t) + r & \text{if } r \in [0, 1] \end{cases}$$

where α is a small positive factor regulating the decay of R toward -1 in the absence of positive reinforcement ($\alpha = 0.02$ in the simulations).

Learning Rule for Triads: We use a Hebbian rule, which is modulated by reinforcement. When R is positive, active triads are enhanced when the postsynaptic neuron is active; their efficacy is reduced when the postsynaptic neuron remains silent. When R is negative, the rule reverses: active triads are enhanced when the postsynaptic neuron is inactive and reduced in case of postsynaptic activity. The net effect of this rule is a stabilization of ongoing activity when R is positive and a destabilization when R is negative.

Formalization:

The maximum efficacy $W_b^m(t)$ of each modulated bundle (a, p, m) is updated according to

$$\delta W_b^m(t) = \beta R \frac{W_b(t)}{W_b^m(t)} S_a(t) [2 S_p(t) - 1]$$

Activation of Rule-coding Clusters: Another form of learning in the network relies on modifications of activity in the rule-coding layer. Modifications occur only when the organism is highly dissatisfied. Each modification resets the self-sustained activity of rule-coding clusters to a random value.[6]

Formalization: At the end of each trial, a modification of the rule-coding layer activity occurs with probability

$$P_{\text{modification}} = \begin{cases} 0 & \text{if } R > -0.5 \\ -0.25 - R & \text{if } R < -0.5 \end{cases}$$

A modification consists in resetting independently the activity of the clusters equiprobably to 0 or to 1.

Functional Properties of the Model

The network was implemented in a Pascal computer program. We simulated its behavior under two conditions: with level 1 units only, and with level 1 plus level 2 units.

Simulation with Level 1 Only

The behavior of the network is simulated assuming that the efficacy of input-output bundles is not modulated by the activity of other units, but can still be modified during learning via the Hebbian rule. The model then reduces to a mapping of visual features onto orientation responses, with the following functional properties:

Grasping: The input-output connections have initially the same efficacy. They provide a capacity for grasping an object: the network correctly orients toward any object presented in isolation, whatever the value of its features.

Behavior with Two Objects: Due to lateral inhibition among output units, the network randomly orients toward one of the two objects.

Systematic $A\overline{B}$ Error: When type 1 trials are presented with the cue always at the same position, the network rapidly abandons its random behavior and always successfully reaches this particular position. For example, a criterion of success in five successive trials (with the cue at location A) is reached in five to eight trials. This is because the Hebbian rule progressively enhances the efficacy of the corresponding connection. As a result, the network develops a strong bias toward always choosing the same position when two objects are presented simultaneously (although it remains able to orient toward an isolated object at any given location). A systematic error of the $A\overline{B}$ type is then observed when the location of the cue is changed to B: the organism continues to orient toward A. It takes about five trials to stop making this systematic error and return to random behavior and another five to ten trials to reach criterion with the new location B (Figures 2 and 3).[7]

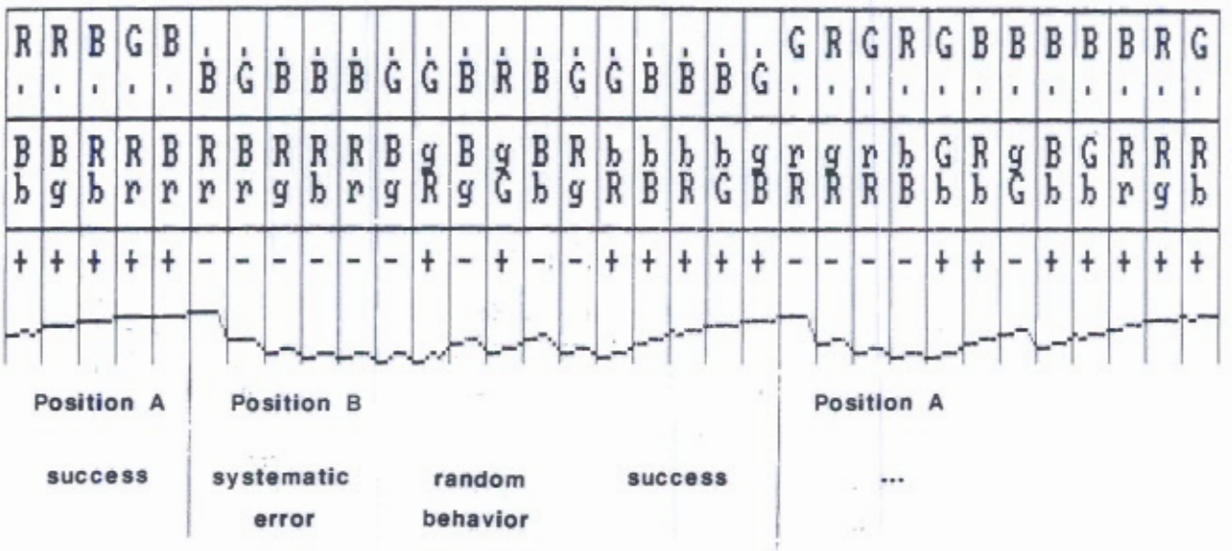

Figure 2. Sequence of actions taken by the network with level 1 only during a $A\overline{B}$ task. Each column represents a trial. The top two lines indicate the objects that were presented at each position, respectively during the cue and choice periods. Letters code for color (R = red, G = green, B = blue). In each row, the uppercase letter, if any, indicates the object toward which the network oriented. The two bottom rows give the sign of the reinforcement obtained at the end of the trial and the evolution of the satisfaction of the organism (parameter R).

In this sequence of trials, the cue object was first presented always at the same location A. After a criterion of five successful trials, location of the cue was switched to B. Examination of the second row reveals that the network continued to orient toward location A, even though it was strongly negatively reinforced. Eventually, this systematic error stopped. Following a period of random behavior, the criterion of success at location B was reached. Location of the cue was again switched, and the systematic error resumed. Without level 2, the network is never able to succeed at both locations at the same time.

Table 1

ACTIVITY IN MEMORY UNITS PREDICTS PERFORMANCE IN $A\overline{B}$

State of the Memory Unit at the End of the Delay Period*	Performance after Delay: Correct**	Performance after Delay: Failure**
active	60	6
inactive	17	17

Notes: 100 trials of $A\overline{B}$ were passed while the network was stabilized in the "location" rule. Activity of the memory unit coding for the location of the cue was recorded. Due to internal noise, activity could drop (Figure 4) or remain stable (Figure 4) during the delay. The table gives the number of correct and failed tests as a function of internal activity.

*The unit was scored as active if its activity parameter (ranging between 0 and 1) was above 0.5 and inactive otherwise.

**A correct trial was scored if the output cluster coding for the correct location was the most active and its activity was above 0.5. Otherwise, a failure was scored.

In brief, the model accounts for the $A\overline{B}$ error by supposing that during initial learning, the reinforced location A comes to elicit quasi-automatic responses via a low-level, slow-to-reverse association. According to this account, location does not play a critical role in the systematic $A\overline{B}$ error. Indeed, a systematic error of perseveration also obtains in the DMS task, where the relevant dimension is color. When the network is trained with a single color for the cue, correct performance is rapidly reached, but if the color of the cue is changed, the network systematically errs and continues to choose the initially reinforced color.

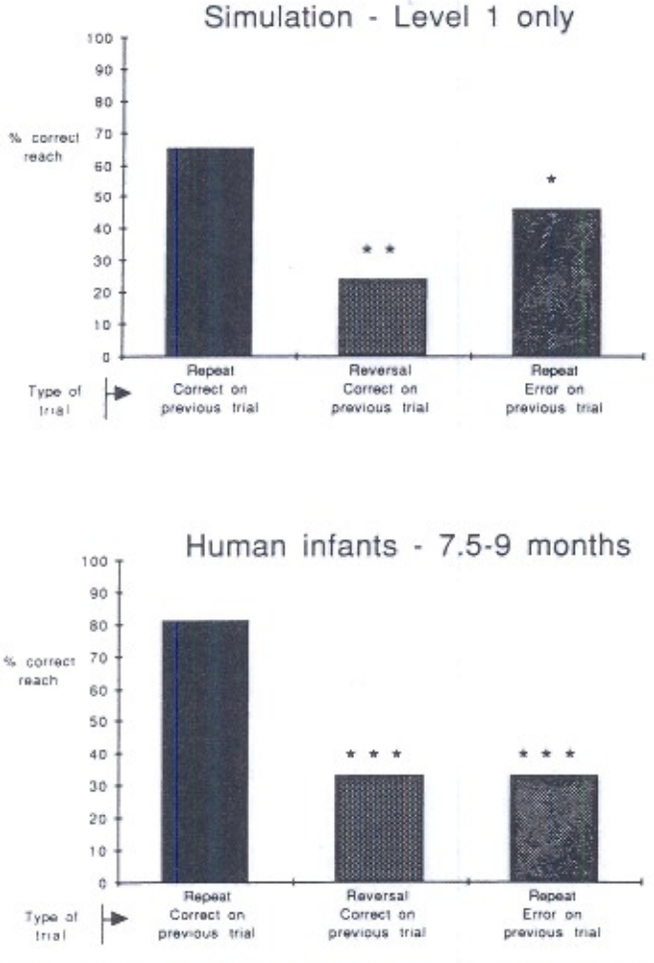

Figure 3. Performance of the network with level 1 only in $A\overline{B}$, compared with the performance of 7 1/2–9-month-old human infants (from Diamond 1985). Trials are sorted depending on (a) whether the previous trial was successful or not, and (b) whether location was changed between the previous trial and the present one. In both human infants and the simulation, the percentage of errors is significantly higher following a change in cue location than when the previous trial was correct and at the same location (*: $p < .01$; **: $p < .001$; ***: $p < .0001$). Following an error, trials at the same location also yield a high percentage of error. This shows that errors tend to occur in rows following a switch in location; however, both the infants and the network eventually succeed at the new location.

Performance in DR and DMS Tasks: In general, the network without level 2 units is simply unable to reach a performance above chance level in DR and DMS. Temporary response biases may appear, but they are rapidly erased by the anti-Hebbian rule as the internal state parameter R becomes more and more negative.

Simulation with Levels 1 and 2

The full network exhibits a much more complex behavior, the characteristics of which are described below.

Selective Extraction of Features: The input-to-memory connections are modulated by the activity of the rule-coding layer. Thus, at a given time, only a fraction of these connections will be potentiated and functional. Depending on the state of rule-coding units, only some features of the cue will be transmitted to the memory layer. For example, if only the rule-coding unit corresponding to color is active, then only the color of the target will be transmitted, not its position.

Memorization During the Delay: Via the potentiated input-to-memory connections, some memory units become activated during the cue period. During the delay, the activated units are able to maintain their activation thanks to excitatory self-connections. Thus, some features of the cue are memorized during the delay. The duration of retention will depend on the level of noise in the system (Figure 4).

Performance During Choice Period: Again, only some of the input-output connections are potentiated when the choice period begins. Their potentiation is commanded by the active memory units. Thus, the organism orients toward the object, if any, which possesses the feature which has been memorized during the delay.

As Table 1 shows, the activity of memory units predicts which object will be chosen. One hundred type 1 trials were simulated, with values of noise and delay such that in one third of the trials, the relevant memory unit for location became inactive before the end of the delay period. When the unit remained active (66% of trials; Figure 4A), performance was correct 90% of the time. However, when the unit became inactive (34% of trials; Figure 4B), performance was at chance level (50% correct).

Rule-coding Layer Activity: The sequence of events just described shows that the object which will be selected during the choice period depends eventually on the activity of rule-coding units. If only the rule-coding unit corresponding to position is active, the organism will always select the object that stands at the same position as the cue; this will ensure correct choice in type 1 trials (i.e., in DR and $A\overline{B}$ tasks). Similarly, activation of the rule-coding unit for color will yield correct performance in type 2 trials (DMS task). The population activity of the

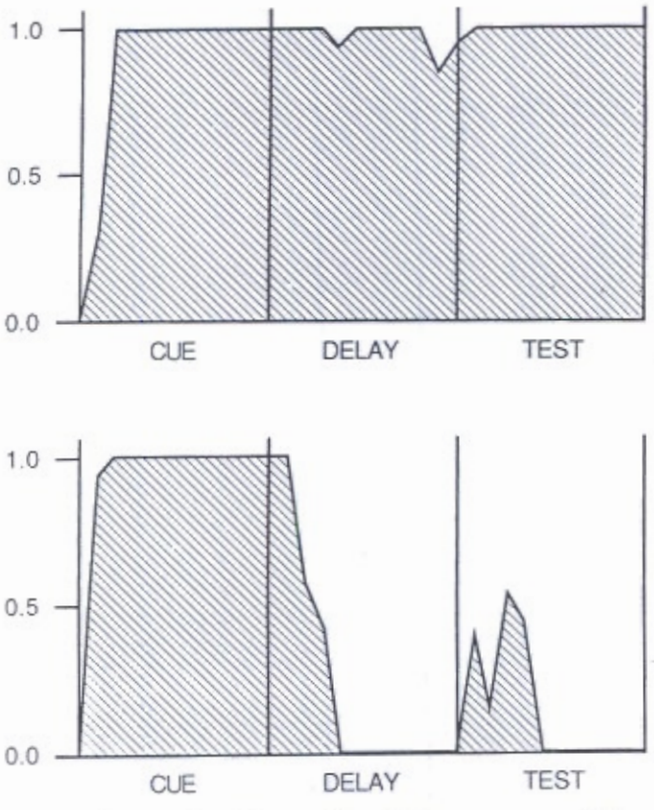

Figure 4. Records of the activity of the memory unit for location B during $A\overline{B}$. Top: The unit detects the location of the cue, memorizes it through the delay period, and command orientation toward that location during the test period. In this trial, performance was correct. Bottom: The unit detects the location of the cue, but activity drops during the delay due to internal noise. In this trial, the network failed to reach toward the correct location.

rule-coding layer thus determines the rule followed by the organism during the choice period.

Learning a rule consists in a search for the correct pattern of activation of the rule-coding layer. This search is implemented by the frequent random modifications of activity which occur as long as the satisfaction of the organism remains low. This allows the network to escape wrong patterns of rule-coding layer activation, which are signaled by punishment. Indeed, in simulations, the network always reaches perfect performance in $A\overline{B}$, DR, and DMS tasks (see Figures 5 and 6).

Fast Relearning: In the network, rules of behavior are coded in patterns of activity, which allows for a fast reaction to a change in the reinforcement schedule. With the set of parameters chosen, two errors in a row suffice to change the current strategy (Figure 5). The network is immediately ready to adapt to the new situation, without going through a long series of systematic errors.

Systematicity: Systematicity is defined as the capacity of a learning system to immediately generalize an acquired rule to a whole class of objects. According to Fodor and Pylyshyn (1988), human thought processes are systematic, since a human mind does not exist that can represent "John loves Mary" but not "Mary loves John." A system cannot be said to acquire a *rule* if it does not exhibit systematicity.

The issue of systematicity arises in the present model because the organism must respond in a regular way to a whole class of configurations. For instance in the DMS task, on each type 2 trial, the cue and test objects can be presented at two different locations and can take on three different colors, resulting in a large variety of experimental configurations. It would have been possible to learn which response is appropriate separately for each configuration. However, such an input-output associationist learning scheme would not be systematic, since the network would learn the correct response to each configuration at different moments in time. In contrast, systematicity requires the network to learn the high-level *rule* that "the colors of the cue and test objects must match" and to apply this rule correctly to all possible configurations, even those it encounters for the first time.

Our network exhibits systematicity in $A\overline{B}$, DR, or DMS rule acquisition. As soon as the activity of the rule-coding layer settles into the correct pattern, all instances of position or color are treated in the same regular fashion. Systematicity arises from the capacity of this network to represent elementary variables: rules are expressed in terms of *the color* or *the position* of objects, but without reference to particular instances.

Generalization: Patterns of activation of the rule-coding layer where only one unit is active correspond to "meaningful" rules. Another pattern, where both rule units are active, is spurious; in the simulations, it yields a random choice between "position" and "color" rules. Unexpectedly, the last pattern, where both units are inactive (null rule), is also meaningful. In this pattern, the memory layer is not affected any more by inputs. Objects are thus invariably chosen by the same criterion. For example, if the memory unit for location A is active, then the organism always orients toward location A. Such a rule may yield perfect choices in the first stage of the $A\overline{B}$ task. However, it is not systematic, since generalization to location B fails to occur. In this situation, the organism still does not make the systematic $A\overline{B}$ error; rather, learning resumes after two errors.

The coexistence of both systematic and nonsystematic rules yields an interesting and desirable property: generalization to new instances of a variable depends on how many different instances have been presented during training. For example, in a DMS task with two possible colors for the cue, the "color" rule is always stabilized, and generalization to a third color is perfect. If,

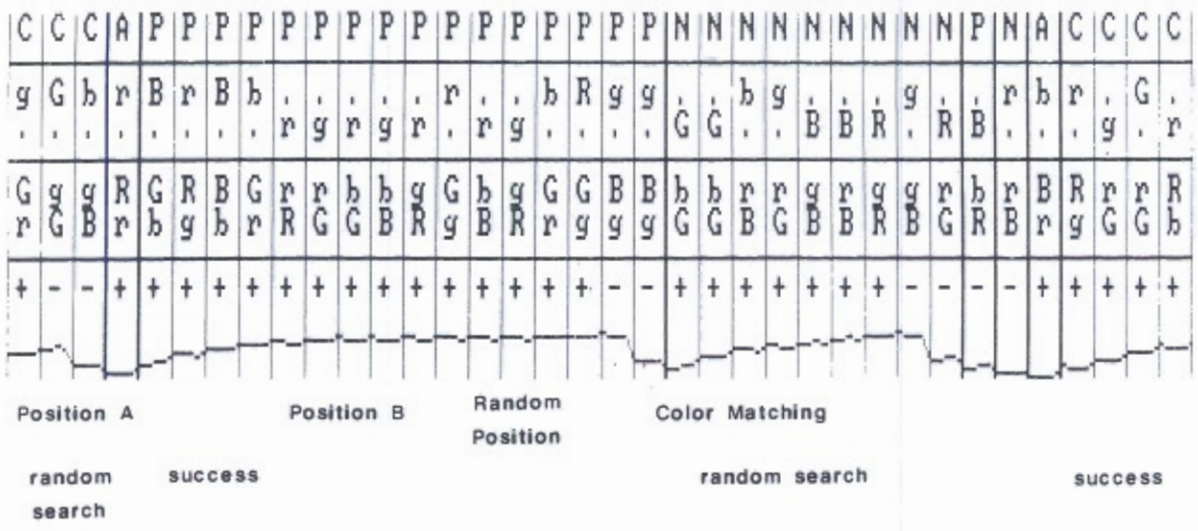

Figure 5. Sequence of actions taken by the network with both levels 1 and 2 during an $A\overline{B}$ task and after switching to a DMS task. The notation is identical to Figure 2, except for the addition of the top row which represents the configuration of rule-coding clusters (C = color rule; P = position rule; A = all rule-coding clusters active; N = no rule-coding cluster active). In the first part, the network received reinforcement if it oriented toward the object standing at the position of the cue. The cue was first presented always at the same location A. The top row indicates that the network started with the wrong rule, but settled by chance in the correct position rule following two consecutive errors. Generalization to position B was then perfect. At this stage (vertical line), the reinforcement schedule was suddenly changed to a delayed matching-to-sample task: the network received reinforcement if it oriented toward the object bearing the color of the cue, independent of position. The network made two consecutive errors, after which a modification of the activity of the rule-coding layer occurred. The subsequent random search eventually converged to the correct color rule.

however, the cue can take only one color during training, in 10–20% of cases the "null rule" is stabilized; then, generalization to the two other colors fails and learning must resume.

When testing for generalization, one should remember that what the experimenter defines as a correct generalization may not always be the optimal strategy for the network. Thus, a surprising phenomenon occurs when the system is trained at length with a DMS task where the cue can take one of two colors, but where any of the three colors can appear during the choice period. When training is short, generalization extends to trials where the cue takes the third color. However, when training is long enough, generalization to the third color fails, as the slow Hebbian learning rule progressively reduces the efficacy of the connections that are associated to the third color. We may call this a failure of generalization. However, a closer look suggests that this behavior is quite appropriate. Indeed, the third color is never a correct choice in the test phase; the network has progressively internalized the appropriate rule that "whatever you do, never choose the third color."

Selection of Rules: As mentioned above, there is one meaningless pattern of activity among rule-coding units (all units active). An interesting extension of the network is to add complete connectivity between rule-coding units, in the spirit of the Hopfield (1982) model. Modification of these connections during learning may then destabilize the spurious patterns of activity, which would become inaccessible. We simulated the network with two connections linking the two rule units, during alternated training in the $A\overline{B}$ and DMS tasks. The Hebbian rule provoked mutual inhibition between rule units, because the state with both units active was generally associated with punishment. As a result, only the three meaningful rules remained accessible; learning was accelerated because the random search was restricted to only three rules instead of four. In larger models, this process may be crucial in reducing the combinatorial explosion in the number of rules.

Comparison with the Performance of Other Neural Networks

It is not straightforward to imagine how other, more classical neural networks might perform in delayed-response tasks. First, most neural networks use supervised learning, whereas we require learning by reinforcement, which is more realistic from a neural point of view (there is no teacher to tell the neurons explicitly what they should do). Second, correct behavior requires the apprehension of relations between inputs that appear at different periods in time; neural networks such as the Hopfield (1982) model or multilayer feedforward networks (Rumelhart and McClelland 1986) accept only simultaneous inputs (at least in their classical versions).[8]

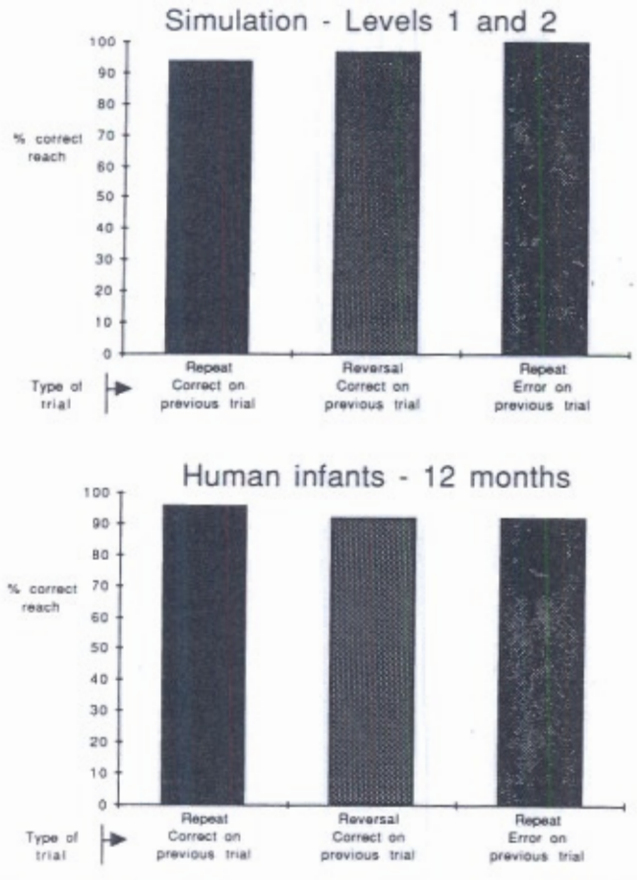

Figure 6. Performance of the network with levels 1 and 2 in $A\overline{B}$, compared with the performance of 12-month-old human infants (from Diamond 1985). Performance is close to 100% correct, irrespective of a change in the location of the cue.

For these reasons, one may say that delayed-response tasks are beyond what classical neural networks can do. This is not to say that these networks could not be adapted to cope with these tasks. Indeed, what we have done is simply to take the very same elements used in most networks — threshold units — and design a specific architecture adapted to the task. The architecture that we selected is clearly not unique. For example, all the properties of our level 1 depend only on using a correlational learning rule; any learning rule that can detect the statistical correlation that exists between the choice of a location and the obtention of reinforcement could replace the particular Hebbian rule that we have used.

To rephrase this, the problem solved by our level 1 ("grasping") is extremely simple and requires only a mapping of each input to the appropriate output. In contrast, the full task is not a simple input-output association but requires a comparison of inputs over time. Thus, an internal representation of inputs is necessary. In that sense, the clusters of level 2 are similar to the "hidden units" of the classical layered networks (Rumelhart and McClelland 1986), but the latter do not usually incorporate time as a relevant dimension.

Interpretation of the Model in Terms of Neurobiological and Psychological Data

In this section, we compare the architecture and the behavior of the network with neurobiological and psychological observations available from the literature.

Comparison with Neurobiological Data

Representational Units: In all layers of the network, the representational units are clusters of neurons; these may be identified with the *groups* or *columns* of neurons which have been found in several areas of the cortex (Mountcastle 1978; Edelman 1978, 1987; see Goldman-Rakic 1984 for the prefrontal cortex). This need not imply that all neurons in a cluster respond in exactly the same conditions. Rather, it is likely that each neuron in a cluster keeps some functional "singularity" (Changeux 1983). Similarly, it is not absolutely necessary that any given neuron belong to only one cluster, although this was found easier to model. Totally distributed or holographic representations (Pribram et al. 1974; Hopfield 1982; Rumelhart and McClelland 1986), where single neuron activity encode little information, are not very plausible. However, recent experimental evidence (Heit et al. 1988; Lee et al. 1988) supports limited "coarse-coding" (Hinton 1981; Ballard 1986) better than absolute selectivity.

Cortical Areas: In the network, units coding for similar features or events are grouped in the same layer. Separate parallel pathways are assumed for processing identity and localization of visual objects. These aspects are consistent with current views of cortical organization (Ungerleider and Mishkin 1982; Rakic and Singer 1988; Goldman-Rakic 1987, 1988). Tentatively, one may try to identify the layers of the network with real cortical areas. The input layer would stand for secondary visual areas and/or association areas such as temporal areas for visual identification of objects and parietal areas for localization. Level 2, with its memory and rule-coding layers, would stand for prefrontal areas (centered on the principal sulcus for spatial delayed-response tasks, and on the inferior convexity and/or the orbital prefrontal cortex for delayed identity-matching tasks; see Goldman-Rakic 1988). Finally, the output layer might stand for premotor areas, the basal ganglia and the other areas involved in motor control, to which the prefrontal cortex has numerous projections.

Topological Projections: In the network, ordered projections between layers play a critical role in the pro-

cess of choosing the correct object. This aspect is consistent with the finding of topological maps (see Woolsey 1981 and the references therein) and ordered projections throughout the cortex, even in prefrontal cortex (e.g., Goldman-Rakic and Schwartz 1982). Obviously, these features are very grossly encoded in the present model, where we have used a perfect isomorphism between sensory and motor codes.

Modulation and the Prefrontal Cortex: We have postulated the existence of synaptic triads, in which the efficacy of a given synapse from neuron A to neuron B is modulated by the activity of a modulator neuron C, which sends a neighboring synapse onto B. A biologically plausible implementation of this feature is a change in the configuration of postsynaptic allosteric receptor molecules at synapse A-B, triggered by chemical and/or electrical messages originating from synapse C-B. Such heterosynaptic modulation has been modeled by Heidmann and Changeux (1982), Changeux and Heidmann (1987), and Dehaene et al. (1987).

Modulation of the efficacy of connections plays two critical roles in the model. First, level 2 units do not directly command behavior, but merely modulate or select actions from the repertoire of the lower level, in agreement with old stimulation experiments (Stuss and Benson 1986). Second, modulation allows level 2 of the network to select its own inputs. As a result, only features that are relevant to the current task are memorized during the delay period. This is consistent with Fuster's (1984) conclusion that "what determines the reaction of many prefrontal cells to a sensory stimulus is the value that stimulus may have as a cue for behavior" (p. 409).

Electrophysiology of Prefrontal Neurons: Activity in the memory layer reproduces the observed pattern of firing of some prefrontal neurons during delayed-response tasks:

1. As already mentioned, such cells respond to sensory features of the cue (Niki and Watanabe 1976a); they code only for features that are relevant to behavior (location in the $A\overline{B}$ task, color in the DMS task).
2. They remain active during the delay period (Figure 4). The mechanism that we propose for this feature is excitatory connections within each cluster.
3. Their activity predicts the future behavior of the monkey (Niki and Watanabe 1976a; Watanabe 1986) and the occasional errors that it makes (see Table 1), thus coding for its "intention."

Connections with the Limbic System: We assume that reinforcement is internally represented and commands random modifications of the rule-coding units. The entry of the reinforcement signal could be mediated by the limbic system, which projects massively to the prefrontal cortex (Stuss and Benson 1986; Goldman-Rakic 1987).

Intervention of the Hippocampus: We have not explicitly included hippocampal neurons in our model, yet it is possible that some components of the network, e.g., the rule-coding neurons, are located within the hippocampus (Gray 1982). Experimental data show that hippocampal monkeys reach randomly in the $A\overline{B}$ task, rather than making the systematic $A\overline{B}$ error (Diamond et al. 1989). These data can be accounted for by the model if one assumes that the integrity of the hippocampus is necessary for learning in the network. Random behavior is also observed in the simulation in the absence of learning.

Comparison with Psychological Data

Initial Performance: From a behavioral point of view, the performance of the network with level 1 only is analogous to the performance of 7 1/2–9-month-old babies, 1 1/2–2 1/2-month-old monkeys, or prefrontal monkeys (Diamond 1988). Grasping an isolated object is possible, but the organism cannot learn DR and DMS tasks, and makes systematic errors in the $A\overline{B}$ task.

Diamond (1985) gives the percentage of errors made by 7 1/2–9-month-old babies in an $A\overline{B}$ task depending on (a) whether the previous trial was successful or not, and (b) whether location was changed between the previous trial and the present one. In order to provide a direct comparison of network and babies performance, we collected similar statistics for network performance with level 1 only (Figure 3). Our criterion for changing the location of the cue was three successes in a row (Diamond used a criterion of 2 correct reaches with infant and adult monkeys and 1–3 correct reaches with human infants). In both human infants and the simulation, the percentage of errors was significantly lower following a change in cue location than when the previous trial was at the same location ($A\overline{B}$ error).[9]

Performance at a Later Stage: With level 1 plus level 2, the network becomes able to learn the $A\overline{B}$ and DMS tasks, in a manner analogous to 12-month-old babies, 4-month-old rhesus monkeys, and adults (Diamond 1988). Figure 6 again shows a comparison of network and 12-month-old babies performance in $A\overline{B}$, as reported by Diamond (1985). There is no significant effect of changing the location of the cue; performance is close to 100% correct in all circumstances.

The network shows an ability to switch rapidly from one task to another (Figure 5); this ability crucially depends on level 2 activity. Switching between tasks may be viewed as an analog of the Wisconsin Card Sorting Test (Milner 1963). As mentioned, this test requires the subject to sort cards according to a given criterion, which the experimenter may change without telling the subject; it is considered a critical test of prefrontal

damage (see also Lhermitte 1983; Luria 1966; Shallice 1982).

Continuity in the Developmental Stages: We have simulated two different states of the network: with level 1 only, and with levels 1 and 2. Yet success in delayed-response tasks does not appear suddenly during development; rather, performance improves continuously. At least two psychological phenomena need to be discussed in that respect.

(1) In $A\overline{B}$ tests Diamond (1985) reports that the baby's gaze sometimes, although rarely, orients toward the correct location, even though the baby reaches incorrectly with the arm. Similarly, by measuring the surprise of babies viewing diverse violations of the physical properties of objects, Baillargeon (1986; Baillargeon et al. 1985) has shown that even at five months of age, babies "know" that an object remains at its location after it has been occluded. Of particular relevance is a recent experiment by Baillargeon and Graber (1988). They used a nonsearch $A\overline{B}$ task, where the babies simply watch a toy being hidden at one of two locations and being retrieved after a delay of 15 seconds. According to Baillargeon and Graber, 8-month-old babies look reliably longer when the toy is retrieved at a location different from the one it was initially hidden in than when the toy is retrieved normally at its initial location. This longer looking time is thought to indicate surprise, showing that the infants remembered where the object was and expected it to be retrieved there.

These observations suggest the existence of a stage in which the location of the cue is remembered throughout the delay, but this knowledge cannot be used to govern motor behavior. In our network, this phenomenon can be accounted for by postulating an intermediate stage of connectivity between level 1 and levels 1 plus 2. In this intermediate stage, the connections from input units to level 2 would be functional, but the connections from level 2, modulating the input-output lines, would not be available. Simulations of this situation show that correct information about the cue is indeed memorized, but cannot be used to guide behavior. The interpretation of Baillargeon's experiments in our model thus leads to a nontrivial neural prediction: afferent pathways to the prefrontal cortex may become functional before the efferent pathways from prefrontal cortex to premotor and motor areas.

(2) Human infants do not suddenly succeed in $A\overline{B}$ tasks; rather, a given infant will or will not make the $A\overline{B}$ error depending on the duration of the delay period. At short delays, performance is correct; at long delays, performance is random and the subjects may fail to cooperate and show overt signs of distress, such as crying. It is only at an intermediate value of delay that a systematic $A\overline{B}$ error is observed. The critical delay increases with the age of the baby (Diamond 1985).

In the network, the continuity of developmental changes in $A\overline{B}$ may be accounted for by supposing that the noise level in memory units slowly decreases with age. In the simulations reported above, this noise level was always very low; as a consequence, network performance was close to 100% correct independent of the length of the delay. The results of simulations with a large noise level are reported in Figure 7. A dependency in the duration of the delay period is now observed. Performance is essentially 100% for short delays and drops toward 50% at very long delays; at that time, memory unit activity no longer reflects the characteristics of the cue: it is simply random. At an intermediate value for the delay, a significant difference appears between reversal and repeat trials. Performance is significantly worse following a change in the location of the cue. Thus, there is a trend in the direction of an $A\overline{B}$ error, although it is not as impressive as in simulations with level 1 only.[10]

Learnable Rules and Systematicity: There are serious limitations on what the network can learn. Essentially, only *identity rules* are learnable, where color or position of the cue and the test objects have to be matched. Many other tasks are not accessible, for example those that involve planning and controlling a sequential behavior (Shallice 1982; Joseph and Barone 1987). In fact, the model gives an extreme example of systematicity. On the one hand, this has the important advantage (Fodor and Pylyshyn 1988) that not all arbitrary stimulus-response associations are equally learnable; for example, the rule "choose the green object only if the cue was red and on the right" cannot be learned, and this matches our intuitions. On the other hand however, the extreme regularity of the architecture impairs the learning of more "natural," yet still arbitrary rules, such as "go right if the cue is green, and left if the cue is red." In that respect, the network is clearly an extremely simplified and incomplete model of prefrontal function.

Extension to Larger Networks

The two main limitations of our network, on neurobiological and psychological grounds, are (1) its oversimplified architecture, and (2) the small size of its repertoire of learnable rules.

With 17 nodes, our network represents only a suggestive analog for prefrontal organization. Thus, a crucial remaining problem is its extension to a much larger network. Further work should include a more precise account of the known connectivity of prefrontal cortex and related brain structures. Also, the extremely regular connectivity of the proposed model should sooner or later be replaced by a more plausible organization resulting from complex processes of epigenetic development. An initial variable set of connections would be pruned according to mechanisms of selective stabilization (Changeux et al. 1973; Toulouse et al. 1986), leading to the emergence of complex clusters that the model postulates (see also Dehaene et al. 1987). Such diversification of the initial

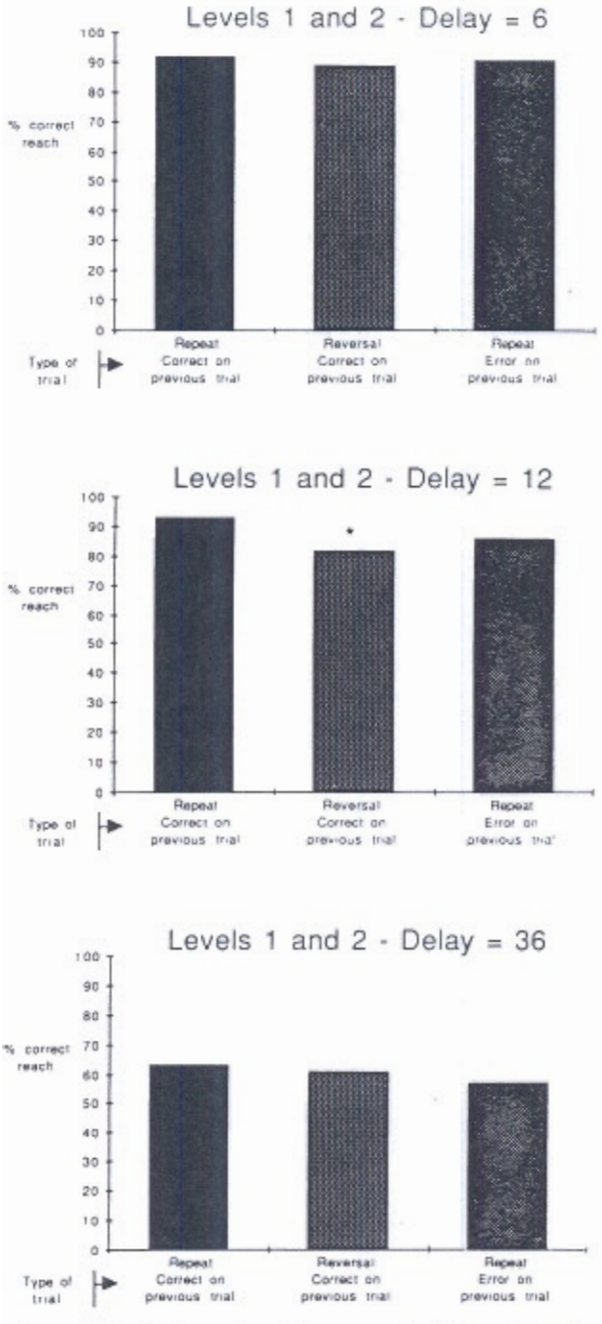

Figure 7. Evolution of performance in $A\overline{B}$ as a function of delay length. At short delays, performance is perfect. At long delays, it is essentially random. A significant trend toward the $A\overline{B}$ error is perceptible at an intermediate value of delay, reproducing the experimental observations of Diamond (1985).

connectivity, accompanied by an increase in the number of neurons, may extend the capacities of the network to the learning of less systematic, more complex rules. However, among the problems that render this extension difficult, the most important one is the combinatorial explosion of the number of possible rules, which would grow exponentially with the number of rule-coding clusters. The learning by random search through this space would soon become very slow.

Such combinatorial explosion might be alleviated by adding connections among rule units (and from input units to rule units), which would destabilize patterns of activity of the rule-coding neurons, thus reducing the search to progressively fewer and fewer rules (this may also lead to the storage of the repertoires of rules in long-term memory). Another possibility would be to implement a hierarchy of rules. Each layer of rule-coding clusters would itself be controlled by a higher-level layer containing rules for the choice of rules at the lower level ("metarules"). Finally, starting from a small set of rules, more and more complex rules could be generated by recombination of fragments of older ones. Holland, Holyoak, Nisbett, and Thagard (1986) have developed a program where initial rules are chosen at random. The most successful ones are progressively selected, and new rules are synthesized by mixing fragments of the former ones. This feature accelerates the induction process by focusing the search on some subspace of the enormous space of possible rules.

The implementation of these ideas in large neuronal networks is not straightforward and is in progress. Nevertheless, the aim of the present network was to give a minimal description of prefrontal function, and its main architectural principles will have to be preserved for future extension to larger networks:

1. Separation into two levels of complexity is required to describe the performance of infants or lesioned patients. Prefrontal cortex merely selects or modulates actions performed by a lower, more automatic level.
2. The higher level must be able to memorize neuronal activity by reverberant loops. In our network, self-excitatory connections within clusters play this role.
3. To acquire systematic rules, the higher level must be able to treat all instances of a variable (e.g., color) in the same way. Thus, some units (the rule-coding clusters in our network) must gate all units corresponding to a given feature dimension.
4. The whole system must learn by reinforcement, with no explicit teacher. Feedback about the accuracy of the unfolding of a plan of actions must be able to modify the on-line neuronal activities, not just the synaptic efficacies.

Predictions and Perspectives

In addition to reproducing known psychological and neurobiological facts, the model enlightens a number of points and leads to some specific predictions.

Interpretation of Prefrontal Function

What insight into the functions of the frontal lobes can we gain from the model? Reviewing her own experimental

work, Diamond (1988) rejects the classical explanations of frontal patients performance in terms of perseveration or forgetting. She proposes instead that two abilities may crucially depend on frontal cortex integrity: "(a) relating information over space or time (...), and (b) inhibition of predominant action tendencies" (p. 360). These two aspects are indeed the key properties of our network. To summarize them, one may say that the frontal cortex enables *inductions*[11] to be performed *over time and space* and that it competes with a lower level capable only of learning stimulus-response associations. The first part of this statement implies that the frontal cortex embodies and updates representations of the environment (Goldman-Rakic 1987) and is involved in planning the interaction of the organism with the environment (Shallice 1982; Luria 1966). The notion of competition implies that if the frontal cortex is lesioned or not fully functional (as in babies), then a lower level, "associationist," system takes over. The working of this lower-level system would be revealed by the $A\overline{B}$ error, the Wisconsin test (Milner 1963), or the "utilization behavior" described by Lhermitte (1983).

Rule-coding Neurons

Despite its extreme simplicity, the model makes specific predictions. We suggest the existence of a direct link between the rules of behavior followed by the animal and the firing of some prefrontal cells. We call these predicted cells *rule-coding neurons*. Patterns of activity among them would code for hypotheses about the rule that governs the stimuli. Their activity would correlate with the "expectancy" that is entertained by the monkey at a given time during learning. Thus, changes in their firing rate should be linked to errors (for instance, a change is likely to follow two consecutive errors).

This prediction can be tested empirically. Two experimental paradigms of neuronal recordings in the awake monkey may reveal the predicted rule-coding neurons. The first possibility would be to train a monkey to accomplish several different tasks on the very same material, under the same conditions as our model. After training, when recording any given unit, the tasks may be alternated (as in the Wisconsin Card Sorting Test) and neuronal responses examined during the search for the correct rule. Rule-coding neurons should respond only when a given rule is tested by the monkey. The second possibility would be to train a monkey with only one task (e.g., DR) but to record during the training phase. Published reports on electrophysiological correlates of DR hardly ever give details about training. It becomes of importance to determine the various behavioral phases through which the monkey passes and try to correlate them with variations in the activity of neurons from prefrontal cortex and/or hippocampus. Predicted phenomena included the progressive tuning of neurons to specific events of the task (Dehaene et al. 1987), the disappearance of responses to irrelevant features of the stimuli and the selection of patterns of activity that are often associated with success in the task.

Induction, Inference, Systematicity, and Mental Darwinism

With respect to the more philosophical issue of learning theories, our model introduces an original distinction between induction and inference. Although the network clearly induces systematic rules, it does not perform any inference on the various situations that it encounters. It does not try to extract regularities from its environment and synthesize the corresponding rule. Rather, it selects rules produced at random until one is correct.

The induction of rules by random search illustrates the notion of "mental Darwinism" (Changeux and Dehaene in press). According to this view, the brain spontaneously produces transient and labile prerepresentations, which constitute hypotheses about the world; they are implemented here as the spontaneous activity of rule-coding clusters which operate as "generators of diversity" (Changeux et al. 1984). The prerepresentations may be selectively stabilized or eliminated according to their matching to the environment and/or to their adequacy with the current goals of the organism. Learning by selection is stimulating interest in cognitive science (Piatelli-Palmarini in press), and our simulations further stress its plausibility at the higher levels of cognition (see also Edelman 1987).

Finally, our model shows that systematic rules of behavior may be acquired by reinforcement in a neural network. As correctly predicted by Fodor and Pylyshyn (1988), the price to pay for systematicity is a structured architecture in the initial state. It is a well-taken point for neurobiologists that learning is very unlikely to take place from a tabula rasa or a fully connected network and most certainly requires highly structured neuronal architectures that are laid down during development under stringent genetic control (see Changeux 1983; Changeux and Dehaene in press). However, we do not agree with Fodor and Pylyshyn that the biological structures that provide systematicity will necessarily be isomorphic with classical computer architectures, and will bring no novel insight into rational behavior. On the contrary, we feel that simple neuronal parameters and architectures may affect the higher functional levels (Changeux and Dehaene in press). Conjoint electrophysiological, behavioral, and modeling studies now have the potential to reveal such neuropsychological links.

Acknowledgments

We thank Jean-Pierre Bourgeois, Emmanuel Dupoux, and Jacques Mehler for helpful support and comments. This work was supported by grants from the CNRS (ATP "Aspects cognitifs et neurobiologiques du Langage"), the Ministère de la Recherche et de l'Enseignement Supérieur (decision N 84C1390), the INSERM (contract N 872 004) and the Collège de France.

Correspondence and reprint requests should be addressed to:

Stanislas Dehaene and Jean-Pierre Changeux, Institut Pasteur, Laboratoire de Neurobiologie Moléculaire, 25 rue du Dr Roux, F-75724 Paris Cédex 15, France.

Notes

1. We use the term "delayed-response tasks" as a generic name. Specific tasks such as $A\overline{B}$, delayed response (DR), delayed matching-to-sample (DMS), and delayed alternation (DA) are referred to by their abbreviations.
2. Delayed-response tasks have been implemented in a variety of ways. In $A\overline{B}$ with human infants, the experimenter often hides a toy in one of two wells facing the infant; after the infant has been distracted away, he is allowed to reach one of them. With infant monkeys, the same implementation has been used, except that food replaces the toy. With trained adult monkeys, the tasks are often more abstract: two lights of different colors are used; after the delay, the monkey must press one of them, basing its response either upon the color or upon the location of the cue; reinforcement consists in a drop of juice. The differences owing to these various paradigms are not taken into consideration in this paper; indeed, the convergence of the results seems striking.
3. In classical $A\overline{B}$ and DR tasks, location is the only distinguishing feature of the two choice objects; in particular, color is kept constant. Our type 1 trials are slightly different in that respect: the two choice objects may differ in color, and the network has to detect that this variation is irrelevant for the task. This feature was adopted so that type 1 and type 2 trials became symmetrical. We checked that the qualitative behavior of the network was not affected when only one color was used for all type 1 trials.
4. Delivery of small positive reinforcement during the cue period is not absolutely necessary, but we found that it slightly accelerates learning.
5. The noise term N plays no role in information processing. It is introduced solely to demonstrate the robustness of the simulation to small perturbations. Also, variations in the noise level may account in part for behavioral modifications in the course of development (see the section *Comparison with Psychological Data*).
6. Burnod and Korn (1989) give evidence for an activity-dependent modulation of the synaptic noise level in the Mautner cell of the fish. This may provide a plausible mechanism for the reinforcement-dependent random modifications of rule-coding layer activity that we have postulated. However, similar evidence in the cortex is still lacking.
7. The durations of the phases of systematic error and random behavior are only indicative. They are highly dependent upon the choice of parameters. For instance, when we use a large value for the learning parameter β, with boundary conditions ensuring that the synaptic efficacies do not diverge, the phase of random behavior may even disappear. Following a reversal in cue location, the network then errs for 2 or 3 trials in a row, and suddenly reaches correctly. Because of this dependency upon the details of the model, real data may easily be fitted by the model, but such attempts would not necessarily lead to a better understanding of the underlying processes. In addition, our network lacks mechanisms that may contribute to real performance, such as the memory of previous errors.
8. Barto, Sutton, and Anderson (1983) have successfully applied a powerful algorithm of learning by reinforcement to the problem of dynamically balancing a pole standing on its tip, a problem in which time is a crucial dimension. Their formalism is quite powerful and it is likely that some version of their algorithm can succeed in our tasks. However, extending the pole-balancing network to our problem is not straightforward.
9. The slight difference in performance between network and infants in trials repeated at the same location following an error can be attributed to the noisier performance of the network; thus, a row of systematic $A\overline{B}$ errors is less frequent in the simulations than in real infant performance. Again, this is highly dependent upon the choice of parameters.
10. Another complementary way to allow for a continuous change between the network with level 1 only and the network with levels 1 and 2 is to suppose that the memory units come to modulate the input-output connections in a progressive manner. In the above simulations, a given input-output connection is 100% efficient if its corresponding memory unit is active; otherwise, its efficacy drops to 0. Intermediate modulation may be modeled by the following equation:

$$W_b(t+1) = \begin{cases} \alpha_p W_b(t) \\ \quad +(1-\alpha_p)W_b^m(t) & \text{if } S_m(t) > 0.5 \\ \alpha_d W_b(t) \\ \quad +\min(1-\alpha_d)W_b^m(t) & \text{if } S_m(t) < 0.5 \end{cases}$$

where the min parameter is assumed to vary from 1 to 0 in the course of development. This equation implies that the efficacy of input-output lines never drops lower than a certain percentage of the maximum efficacy. Simulations were run using this equation with min = 0.80 and a large noise level. At short delays, performance is perfect, and indeed, the network is functionally identical to the normal network with levels 1 and 2. At longer delays, noise erases the activity of memory units; the network becomes equivalent to level 1 only, and a systematic $A\overline{B}$ error is observed.
11. Very young infants and prefrontal monkeys succeed in $A\overline{B}$ and DR when no delay is used, so it may be argued that prefrontal cortex may not be required for the induction of behavioral rules, but only for their memorization. Yet, we think that in the absence of a delay, $A\overline{B}$ and DR become simple grasping tasks and do not require the learning of a specific rule. On the opposite, the performance of prefrontal patients in tests like the Wisconsin Card Sorting Test or the Towers of Hanoi clearly demonstrates deficiencies in rule induction and/or application.

References

Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants. *Cognition*, 23, 21–41.

Baillargeon, R. & Graber, M. (1988). Evidence for location memory in 8-month-old infants in a nonsearch $A\overline{B}$ task. *Developmental Psychology*, 24, 502–511.

Baillargeon, R., Spelke, E.S. & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20, 191–208.

Ballard, D.H. Cortical connections and parallel processing: Structure and function. *Behavioral and Brain Sciences*, 1986, 9, 67–120.

Barto, A.G., Sutton, R.S. & Anderson, C.W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 1983, *SMC-13*, 834–846.

Burnod, Y. & Korn, H. (1989). Consequences of stochastic release of neurotransmitters for network computation in the central nervous system. *Proceedings of the National Academy of Sciences USA*, *86*, 352–356.

Changeux, J.P. *L'homme neuronal*. Paris: Fayard, 1983.

Changeux, J.P. & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, 1976, *264*, 705–712.

Changeux, J.P. & Dehaene, S. (in press). Neuronal models of cognitive functions. *Cognition*.

Changeux, J.P. & Heidmann, T. (1987). Allosteric receptors and molecular models of learning. In G. Edelman, W.E. Gall, & W.M. Cowan, (Eds.), *Synaptic function*. New York: Wiley, 549–601.

Changeux, J.P., Courrège, P., & Danchin, A. (1973). A theory of the epigenesis of neural networks by selective stabilization of synapses. *Proceedings of the National Academy of Sciences USA, 70*, 2974–2978.

Changeux, J.P., Heidmann, T., & Patte, P. (1984). Learning by selection. In P. Marler & H. Terrace, (Eds.), *The biology of learning*. Berlin: Springer-Verlag, 115–139.

Cummings, E.M. & Bjork, E.L. (1983). Search behavior on multichoice hiding tasks: Evidence for an objective conception of space in infancy. *International Journal of Behavioral Development, 1*, 71–88.

Dehaene, S., Changeux, J.P., & Nadal, J.P. (1987). Neural networks that learn temporal sequences by selection. *Proceedings of the National Academy of Sciences USA, 84*, 2727–2731.

Diamond, A. (1985). The development of the ability to use recall to guide action, as indicated by infant's performance on $A\overline{B}$. *Child Development, 56*, 868–883.

Diamond, A. (1988). Differences between adult and infant cognition: is the crucial variable presence or absence of language? In L. Weiskrantz, (Ed.), *Thought without language*. Oxford: Clarendon Press.

Diamond, A. & Doar, B. (1989). The performance of human infants on a measure of frontal cortex function, the delayed response task. *Developmental Psychobiology, 22(3)*.

Diamond, A. & Goldman-Rakic, P.S. (1989). Comparison of human infants and rhesus monkeys on Piaget's $A\overline{B}$ task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74*, 24–40.

Diamond, A., Zola-Morgan, S., & Squire, L.R. (1989). Successful performance by monkeys with lesions of the hippocampal formation on $A\overline{B}$ and object retrieval, two tasks that mark developmental changes in human infants. *Behavioral Neuroscience, 103(3)*.

Edelman, G. (1978). *The mindful brain. Cortical organization and the group-selective theory of higher brain function*. Cambridge: The MIT Press.

Edelman, G. (1987). *Neural Darwinism*. New York: Basic Books.

Fodor, J.A. & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

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

Fuster, J.M. (1984). Electrophysiology of the prefrontal cortex. *Trends in Neuroscience, 7*, 408–414.

Fuster, J.M., Bauer, R.H., & Jervey, J.P. (1982). Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. *Experimental Neurology, 77*, 679–694.

Goldman-Rakic, P.S. (1984). Modular organization of prefrontal cortex. *Trends in Neuroscience, 7*, 419–424.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In F. Plum & V. Mountcastle, (Eds.), *Handbook of Physiology, 5*, 373–417.

Goldman-Rakic, P.S. (1988). Topography of cognition: Parallel distributed networks in primate association cortex. *Annual Review of Neuroscience, 11*, 137–156.

Goldman-Rakic, P.S. & Schwartz, M.L. (1982). Interdigitation of contralateral and ipsilateral columnar projections to frontal association cortex. *Science, 216*, 755–757.

Gouzé, J.L., Lasry, J.M., & Changeux, J.P. (1983). Selective stabilization of muscle innervation during development: a mathematical model. *Biological Cybernetics, 46*, 207–215.

Gray, J.A. (1982). Précis of the neural psychology of anxiety; and inquiries into the functions of the septal-hippocampal system. *Behavioral and Brain Sciences, 5*, 469–534.

Heidmann, T. & Changeux, J.P. (1982). Un modèle moléculaire de régulation d'efficacité d'une synapse chimique au niveau post-synaptique. *Comptes-rendus de l'Académie des Sciences (Paris)*, série 3, *295*, 665–670.

Heit, G., Smith, M.E., & Halgren, E. (1988). Neural encoding of individual words and faces by the human hippocampus and amygdala. *Nature, 333*, 773–775.

Hinton, G.E. (1981). Shape representation in parallel systems. In A. Drina, (Ed.), *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, IJCAI.

Holland, J.H., Holyoak, K., Nisbett, R.E., & Thagard, P.R. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: The MIT Press.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA, 79*, 2554–2558.

Jacobsen, C.F. (1935). Functions of frontal association areas in primates. *Archives of Neurology and Psychiatry, 33*, 558–560.

Jerne, N. (1967). Antibodies and learning: selection versus instruction. In G.C. Quarton, T. Melnechuck, & F.O. Schmidt, (Eds.), *The Neurosciences*. New York: Rockfeller University Press, 200–205.

Joseph, J.P. & Barone, P. (1987). Prefrontal unit activity during a delayed oculomotor task in the monkey. *Experimental Brain Research, 67*, 460–468.

Lee, C., Rohrer, W.H., & Sparks, D.L. (1988). Population coding of saccadic eye movement by neurons in the superior colliculus. *Nature, 332*, 357–360.

Lhermitte, F. (1983). "Utilization behaviour" and its relation to lesions of the frontal lobe. *Brain, 106*, 237–255.

Luria, A.R. (1966). The higher cortical functions in man. New York: Basic Books.

Mariani, J. & Changeux, J.P. (1981a). Ontogenesis of olivocerebellar relationships: I. Studies by intracellular recordings of the multiple innervation of Purkinje cells by climbing fibers in the developing rat cerebellum. *Journal of Neuroscience, 1*, 696–702.

Mariani, J. & Changeux, J.P. (1981b). Ontogenesis of olivocerebellar relationships: II. Spontaneous activity of inferior olivary neurons and climbing fiber-mediated activity of cerebellar Purkinje cells in developing rats and in adult cerebellar mutant mice. *Journal of Neuroscience, 1*, 703–709

Merzenich, M., Recanzone, G., Jenkins, W.M., Hallard, T.T., & Nudo, R.J. (1988). Cortical representational plasticity. In P. Rakic & W. Singer, (Eds.), *Neurobiology of neocortex*. Dahlem workshop reports. New York: Wiley, 41–67.

Milner, B. (1963). Effects of brain lesions on card sorting. *Archives of Neurology, 9*, 90–100.

Milner, B. & Petrides, M. (1984). Behavioural effects of frontal lobe lesions in man. *Trends in Neuroscience, 7*, 403–407.

Mountcastle, V.B. (1978). An organizing principle for cerebral function: The unit module and the distributed system. In G.M. Edelman & V.B. Mountcastle, (Eds.), *The Mindful Brain*. Cambridge, MA: The MIT Press.

Niki, H. (1974). Differential activity of prefrontal units during right and left delayed response trials. *Brain Research, 70*, 346–349.

Niki, H. & Watanabe, M. (1976a). Cingulate unit activity and delayed response. *Brain Research, 110*, 381–386.

Niki, H. & Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Research, 171*, 213–224.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Piatelli-Palmarini, M. (in press). Evolution, selection and cognition: From learning to parameter setting in biology and the study of language. *Cognition*.

Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed model of language acquisition. *Cognition, 28*, 73–193.

Pribram, K.H., Nuwer, M., & Baron, R. (1974). The holographic hypothesis of memory structure in brain function. In R.C. Atkinson, D.H. Krantz, R.C. Luce, & P. Suppes, (Eds.), *Contemporary developments in mathematical psychology*. San Francisco: W.H. Freeman.

Purves, D. & Lichtman, J.W. (1980). Elimination of synapses in the developing nervous system. *Science, 210*, 153–157.

Rakic, P. & Singer, W. (Eds.) (1988). *Neurobiology of neocortex*. Dahlem workshop reports. New York: Wiley.

Rumelhart, D.E., McClelland, J.L., & the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: Bradford Books.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B, 298*, 199–209.

Stuss, D. & Benson, F. *The frontal lobes*. New York: Raven.

Taine, H. (1870). *De l'intelligence*. Paris: Hachette.

Toulouse, G., Dehaene, S., & Changeux, J.P. (1986). Spin-glass models of learning by selection. *Proceedings of the National Academy of Sciences USA, 83*, 1695–1698.

Treisman, A. & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology, 12*, 97–136.

Tsal, Y. & Lavie, N. (1988). Attending to color and shape: the special role of location in selective visual processing. *Perception and Psychophysics, 44*, 15–21.

Ungerleider, L.G. & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale & R.J. Mansfield, (Eds.), *Analysis of visual behavior*. Cambridge, MA: The MIT Press, 549–586.

Watanabe, M. (1986). Prefrontal unit activity during delayed conditional Go/No-Go discrimination in the monkey. II. Relation to Go and No-Go responses. *Brain Research, 382*, 15–27.

Woolsey, C.N. (Ed.) (1981). *Cortical sensory organization*, Volume 1: *Multiple sensory areas*. Clifton, NJ: Humana Press.

Young, J.Z. (1973). Memory as a selective process. *Australian Academy of Science Report: Symposium on Biology of Memory*, 25–45.