The Wisconsin Card Sorting Test: Theoretical Analysis and Modeling in a Neuronal Network

Stanislas Dehaene and
Jean-Pierre Changeux

URA CNRS D1284, Neurobiologie Moléculaire,
Institut Pasteur, Paris, France

Neuropsychologists commonly use the Wisconsin Card Sorting Test as a test of the integrity of frontal lobe functions. However, an account of its range of validity and of the neuronal mechanisms involved is lacking. We analyze the test at 3 different levels. First, the different versions of the test are described, and the results obtained with normal subjects and brain-lesioned patients are reviewed. Second, a computational analysis is used to reveal what algorithms may pass the test, and to predict their respective performances. At this stage, 3 cognitive components are isolated that may critically contribute to performance: the ability to change the current rule when negative reward occurs, the capacity to memorize previously tested rules in order to avoid testing them twice, and the possibility of rejecting some rules a priori by reasoning. Third, a model neuronal network embodying these 3 components is described. The coding units are clusters of neurons organized in layers, or assemblies. A sensorimotor loop enables the network to sort the input cards according to several criteria (color, form, etc.). A higher-level assembly of rule-coding clusters codes for the currently tested rule, which shifts when negative reward is received. Internal testing of the possible rules, analogous to a reasoning process, also occurs, by means of an endogenous autoevaluation loop. When lesioned, the model reproduces the behavior of frontal lobe patients. Plausible biological or molecular implementations are presented for several of its components.

Studies of the frontal cortex have benefited, in past years, from the concerted efforts of neuropsychologists (Luria, 1966; Shallice, 1982, 1988; Stuss and Benson, 1986; Damasio et al., 1990) and neurobiologists (Goldman-Rakic, 1987, 1988; Fuster, 1989) to an extent that justifies the elaboration of plausible models for the neural bases of specific frontal cortex functions (e.g., Dehaene and Changeux, 1989; Levine and Prueitt, 1989). Recently, we proposed a simple neural network that accounts for the contribution of frontal cortex in delayed-response tasks (Dehaene and Changeux, 1989). The present paper extends this initial model to a more complex, yet classical, psychological task tapping frontal cortex: the Wisconsin Card Sorting Test. We tackle the problem through 3 convergent modes of analysis. First, we briefly review the experimental findings in normal subjects and brain-lesioned patients. Then, using a computational analysis, we identify cognitive components required for an artificial machine to pass the test. Finally, we describe how these components can be implemented within a plausible neuronal architecture.

Description of the Test

The Wisconsin Card Sorting Test (Grant and Berg, 1948) requires subjects to discover the principle according to which a deck of cards must be sorted. The standard material consists of cards bearing geometric figures that vary in color (red, green, blue, or yellow), shape (triangle, star, cross, or circle) and number (1, 2, 3, or 4 items). Four reference cards, shown in Figure 1, are aligned in front of the subject throughout the test. Another deck of cards serves as response cards. The subject is instructed to place each response card in front of 1 of the 4 reference cards, wherever he thinks it should go. After each response, he is told whether the response was "right" or "wrong," but not where the card should have gone. The goal for the subject is to get as many "right" responses as possible. Initially, cards must be sorted according to, say, color. When performance is successful, the sorting rule is changed, for example from color to shape; the subject must notice the change and find the new correct rule.

Neuropsychologists commonly employ 2 versions of the test that differ only in details. In Milner's (1963) original test, 2 decks of all possible 64 cards are used as response cards. The criterion for a change of rule is 10 successes in a row. The subject is not told when

Cerebral Cortex Jan/Feb 1991;1:62–79; 1047–3211/91/$2.00

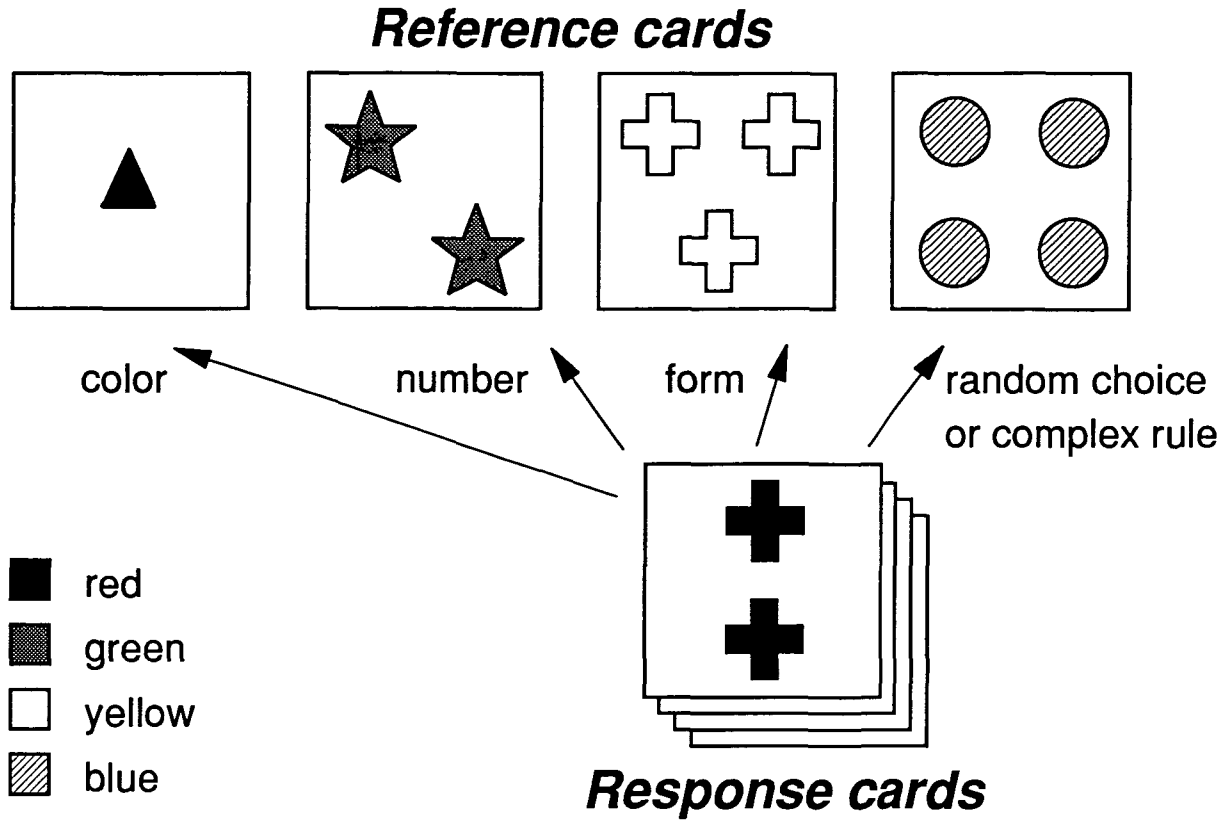

Figure 1. Material used in the Wisconsin Card Sorting Test (adapted from Milner, 1963). The patient must place each response card under 1 of the 4 reference cards, and is then told by the experimenter whether the choice was right or wrong. On the basis of this feedback, the patient must discover the correct sorting rule: *color, number,* or *form.*

the rule changes. The test ends either when the subject has reached 6 criteria, namely color, shape, number, color, shape, number, or when all 128 cards have been used.

Nelson (1976) introduced a number of modifications to the test.

(1) Among the 64 cards of a complete deck, 40 are ambiguous because 2 or more of their attributes are shared with the same reference card. For example, the card with two red triangles is similar to the first reference card in both color and shape. Such cards are ambiguous both for the experimenter, who cannot infer which rule the subject is following, and for the subject, who cannot determine for which rule he was reinforced. Nelson eliminated those 40 cards and used only the 24 cards that shared only one attribute with each of the reference cards. Each card is used twice, for a total of 48 cards in the response deck.

(2) Nelson's version employs a criterion of only 6 successes in a row before the rule is changed.

(3) The subject is warned before each change of rules.

(4) Finally, the correct rules are not defined a priori by the experimenter. Rather, for the first trial, whichever sorting strategy the patient uses is considered correct. Subsequently, whichever new strategy is employed defines the second rule. The third rule is determined in the same way. The 3 rules are then repeated in the same order.

Performance of Normal and Lesioned Subjects

Milner (1963), Drewe (1974), and Nelson (1976) have examined the performance of normal subjects and of patients with different lesion types. Their observations are summarized below.

Normal subjects versus patients. As expected, normal subjects generally obtain more success criteria than lesioned subjects, regardless of lesion site.

Global performance level is not discriminative. Many normal subjects fail to complete the test. For example, in Nelson's (1976) study, 6.5% of normal subjects learned only 1 rule correctly out of the 6 that the full test comprises. Age is a good predictor of success or failure, so the oldest nonfrontal patients make approximately as many errors as the frontal ones. Hence, frontal and nonfrontal old patients cannot be reliably separated on the basis of the number of learned rules, or of the total number of errors.

Perseveration characterizes frontal patients. An error is classified as perseverative if the subject continues to use a rule that was previously correct, even after negative feedback is provided. Frontal patients make significantly more perseverative errors than normal subjects or nonfrontal patients; they fail to shift from one sorting rule to another.

Sites of disruptive lesions. Milner (1963) compared pre- and postoperative performances in patients with a dorsolateral frontal lobectomy, and found a significant postsurgery increase in perseveration. By con-

trast, no excessive perseverations were observed following temporal, bilateral hippocampal, or even inferior frontal lobectomies. According to Drewe (1974), medial frontal lesions may be even more disruptive than dorsolateral frontal lesions. Finally, left frontal lesions generally yield more total errors, though not necessarily more perseverations, than right frontal lesions (Drewe, 1974).

Drewe (1974) correctly emphasizes the "non-unitary nature of the impairments shown" (p. 168). Several factors probably contribute to the complexity of the test. This complexity may be useful, or even necessary in order for the test to be sensitive to frontal lesions. Yet it also impedes our comprehension of the cognitive components involved. In the following section, we undertake a computational or functional analysis of the components that are necessary in order to pass the test. We aim at better understanding the "computational constraints" (Marr, 1982) that govern performance: what are the minimal abilities required, what types of performance are to be expected from a given cognitive architecture, and what are the possible sources of failure in the test?

A Functional Analysis of the Wisconsin Card Sorting Test

Let p be the number of dimensions (color, form, number, etc.) along which the cards may vary, and let q be the number of features along each dimension ($p = 3$, $q = 4$ for the standard test). There are thus q reference cards defining the q possible responses, and q^p different response cards. We assume that a cognitive system, C, is confronted with an infinite series of response cards. For each card, C selects a response according to its current selection rule, receives a positive or negative reward, and decides to change its current rule or not. Our goal is to calculate the mean number of trials required for the system to converge to the correct rule specified by the experimenter, as a function of p, q, the version of the test used (Milner or Nelson), and the architecture of C.

The definition of the cognitive system C must specify the number and type of available selection rules, and the conditions for the rejection of the current rule and the adoption of a new one.

Available rules. Let r be the total number of available rules. It is necessary that the p base rules that define selection according to the p dimensions of the cards (color, form, number, etc.) be available; hence $r \geq p$. The number r is not necessarily equal to p, because the test instructions do not specify the set of possible rules. A subject can envisage complex rules, for example, "Choose whichever reference card is most similar to the response card," or "Choose the triangle if the response card is green, otherwise choose the circle." Obviously, it is not possible to list all such rules. We shall only assume that the $r - p$ additional rules are independent of the base rules and of each other (2 rules are independent if the probability that they specify the same response to a given card, averaged over all possible cards, is $1/q$, i.e., at chance level). This assumption will generally not be satisfied, especially for large r, but it may serve as a first-order approximation.

Conditions for changing the current rule. Initially, the current rule is chosen at random. To ensure the stability of the correct rule once it has been found, it must be assumed that no change of current rule can occur on positively reinforced trials. We further assume that on each trial, there is a probability, P, of ignoring the reward. Thus, a new current rule is selected only when the reward for the preceding trial was negative and was not ignored.

Six different cognitive systems, or machines, may be distinguished according to the manner in which they select a new rule (Fig. 2). The first 3 are random machines that simply draw a new rule at random from the repertoire of available rules, without any reasoning. The simplest possible machine (random) draws a new rule with replacement of the preceding rule; hence, there is a finite probability, $1/r$, of not changing the rule at all. A second, more complex machine (random + context) avoids drawing again the rule that was just rejected. Finally the third machine (random + memory) keeps an episodic memory[1] of previously rejected rules, and draws only among the remaining possible rules, thus progressively reducing the possibilities.

The next 3 machines we consider have the faculty of rejecting some rules not only by trial and error, but also by reasoning. The reasoning is quite simple: on a negatively reinforced trial, all the rules that would have led to the same (wrong) response are necessarily incorrect. The fourth machine (reasoning, no memory) uses such reasoning only locally in time: it avoids choosing any such necessarily incorrect rule, but it does not keep a memory of the rules that were rejected on previous trials. The fifth machine (reasoning + memory) keeps such an episodic memory. Hence, rules that are rejected, either by trial and error or by reasoning, are definitely labeled as incorrect and are never tried again.

Finally, the sixth machine (optimal) guarantees the shortest learning time. In addition to reasoning on negative trials and memorizing rejected rules, it also reasons on positive trials. In such trials, all the rules that would not have led to the same response are rejected as incorrect and memorized as such: they will not be tried on later negative trials. Schematic diagrams of the operation of the 6 machines appear in Figure 2.

Theoretical Results

Given the above specifications, we mathematically derived the mean number of trials before a given machine converges to the correct rule, for both Milner's and Nelson's versions of the tests (an outline of the calculations is provided in Appendix A). We then calculated the mean test duration, i.e., the mean number of trials before 6 criteria are reached. If this mean test duration is lower than 128 trials in Milner's version, or 48 trials in Nelson's version, the machine can be said to pass the test successfully.

Sensitivity of the test. Figure 3 gives plots of mean

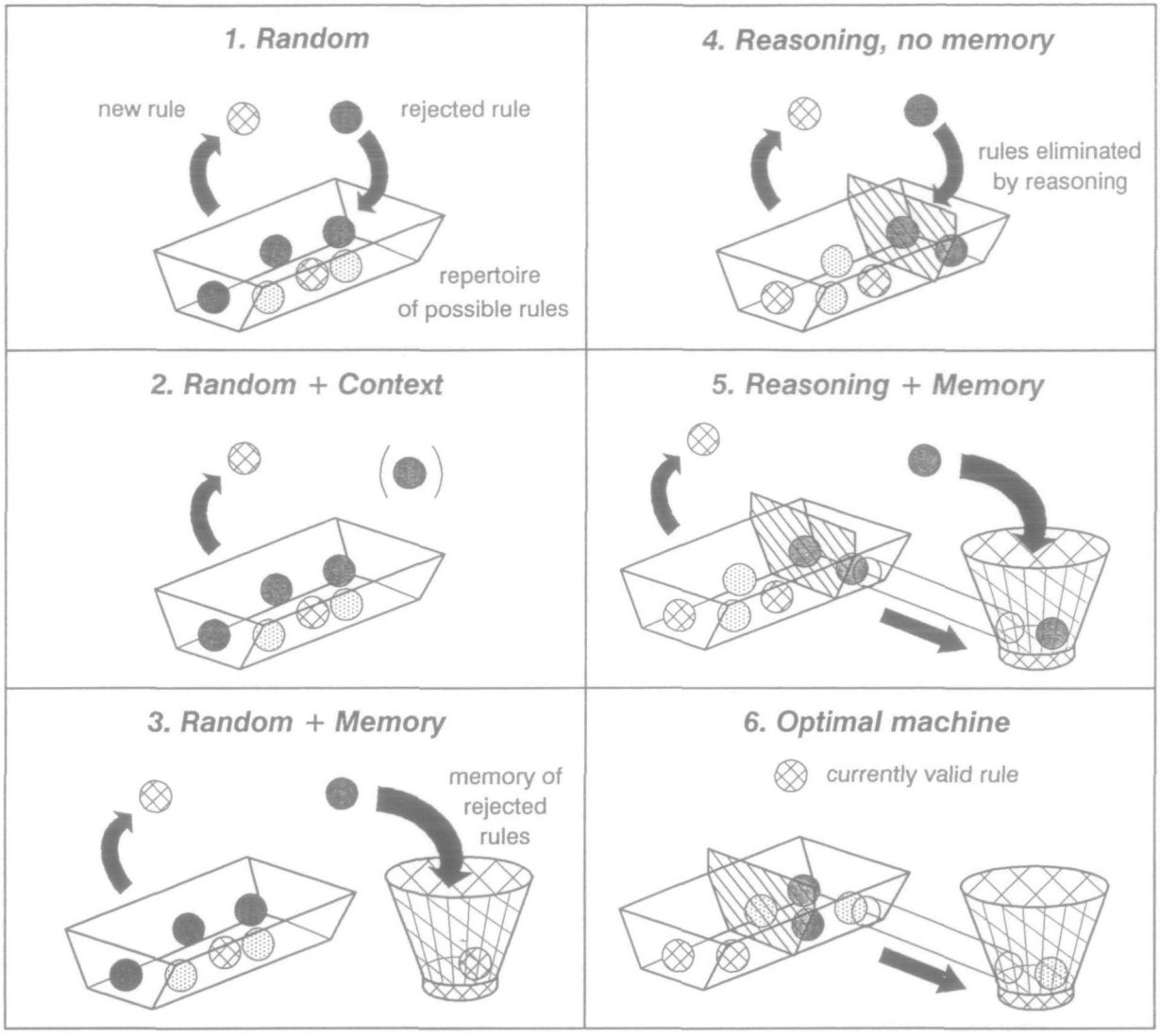

Figure 2. Schematic representation of the learning processes carried out by the 6 machines described in the text. Balls represent the possible rules among which the correct one must be found. The diagrams illustrate the behavior of each machine when its current rule was found to be incorrect, and a new one had to be selected. The *random* machine merely draws a new rule at random from the whole set, including the current rule that was just rejected. For the *random* + *context* machine, the set from which the new rule is drawn excludes the rule that was just rejected, but still includes rules that were rejected on previous trials. The *random* + *memory* machine keeps a memory of previously rejected rules, which are never tried a second time. The machines with *reasoning* reject in one trial not only the rule that was just found incorrect, but also rules that would have lead to the same incorrect response and thus must be incorrect, too. For the machine with *reasoning, no memory,* these rules are just set aside for the current trial only. For the machine with *reasoning* + *memory,* they are definitely labeled as rejected. Finally, the *optimal* machine behaves similarly to the *reasoning* + *memory* machine during incorrect trials, but it also uses reasoning during correct trials. In such trials, the current rule is, of course, kept unchanged, but all rules that would have led to a different response are eliminated from the pool of remaining rules.

test duration as a function of machine type and the number of rules available in the repertoire, when the reward is never ignored ($P = 0$). Note first that when the repertoire contains only the 3 base rules, the performances of the 6 machines are quite similar, and even the most simple random machine passes the test (though Nelson's version appears more selective). Thus, we expect the test to be only weakly sensitive to interindividual differences in memory or reasoning abilities. This is important for clinical purposes, because it allows the testing of subjects with a relatively low level of education (Nelson, 1976). However, as far as neuropsychological research is concerned, it means that the test will not permit the discrimination of subtle lesion types that would affect only one cognitive component, for example, the reasoning process or the episodic memory for previously rejected rules.

Influence of repertoire size. As shown in Figure 3, performance in the test is highly dependent on the total number of available rules (parameter r). The formal analysis thus suggests that an important source of intersubject variability is the range of rules that a given subject will consider. The reliability of the test might be improved by explicitly instructing the subjects that only 3 sorting rules—color, form, and number—are possible.

Cognitive architecture and the combinatorial explosion. Most machines cannot handle a large number of rules and still pass the test; the random search for the correct rule rapidly becomes too slow if the search

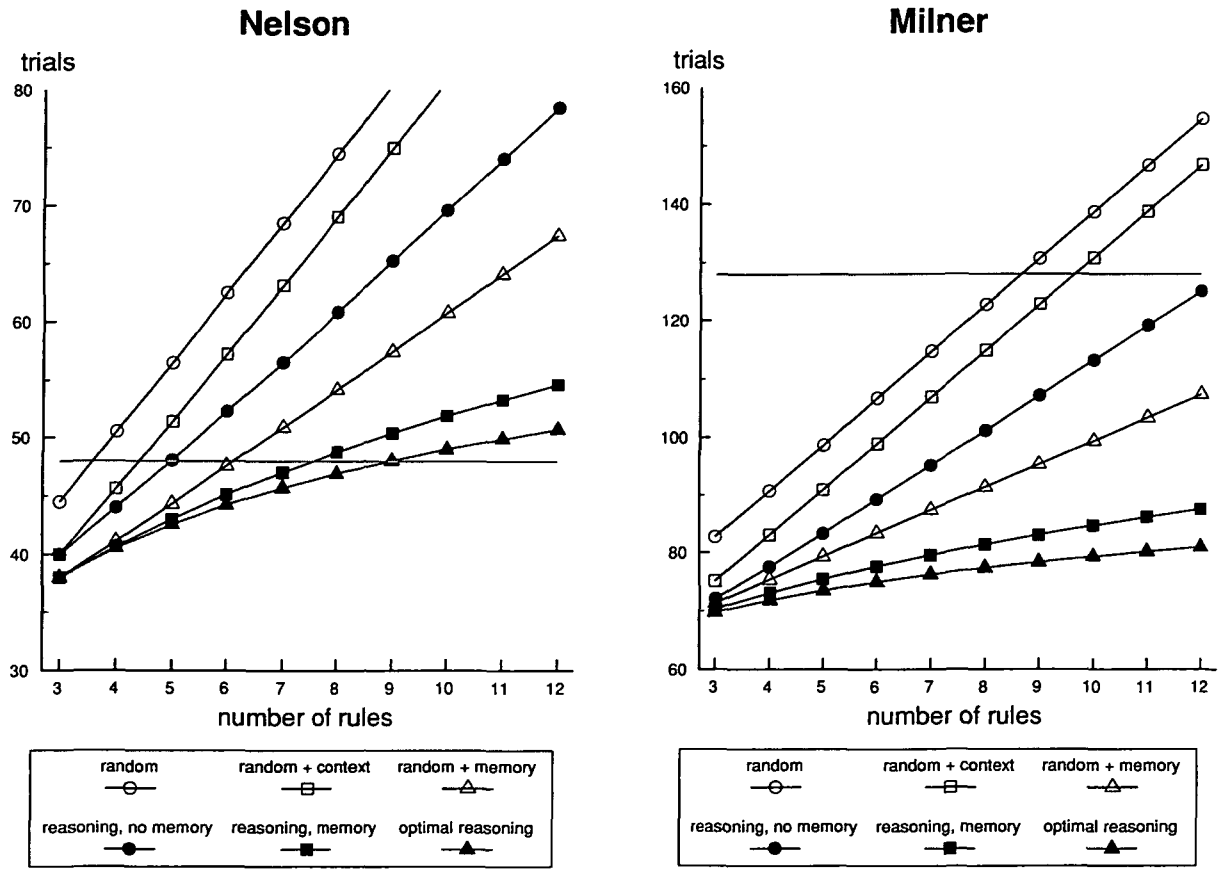

Figure 3. Theoretical test durations for the different machines as a function of the number of rules available, for Milner's and Nelson's versions of the Wisconsin Card Sorting Test. The *horizontal line* gives the upper limit for success in the test.

space is large. However, not all machines are equally sensitive to this effect. In particular, the two machines that possess both episodic memory and reasoning abilities are able to reduce seriously the combinatorial explosion. For them, mean test duration increases only logarithmically, not linearly, with the number of rules.

Importance of correct reward processing. How does performance evolve when the probability of ignoring reward is not 0? The effect of this variable is shown in Figure 4. Ignoring reward on some trials yields a dramatic increase in the mean test duration. The increase is similar in nature for all 6 machines. The data in Figure 4 correspond to machines mastering only the 3 base rules. Of course, the increase would be even more important for machines using a large number of rules.

Sources of failure in the test. Figures 3 and 4 show that there are essentially 2 sources of failure in the test. A subject may fail either because he/she is evaluating too many rules, or because he/she neglects the reward signal. The existence of these 2 sources can readily explain the paradoxical result that many normal subjects fail to reach the required number of criteria in the test. Normal subjects probably fail for the first reason, while frontal patients probably fail for the second one. Note that both types of failures, although confounded by total test duration or total number of errors, can be distinguished by a trial-by-trial examination of performance. Subjects that imagine overly complex rules will appear to perform orthogonally to the 3 base rules, and in particular will make errors inconsistent altogether with the color, form, and number criteria (out-of-class errors). Subjects that tend to ignore the reward will make more numerous perseverative errors, continuing to use a rule after it has been shown to be incorrect. Consistent with this analysis, Nelson (1976) argued that the percentage of perseverative errors discriminates frontal patients from others better than the total number of errors or the number of criteria reached does.

According to this account, the main effect of frontal lesions would be to disrupt the appropriate integration of reward signals into a purposive behavior. In addition, however, one need not assume that the machine type describing a given patient remains invariant following a lesion. Lesions may alter not only the parameters of any given machine (i.e., the number of rules and the probability of ignoring reward), but also the machine type itself, by disrupting reasoning abilities, for example. Such qualitative alterations of processing may not be identifiable using the Wisconsin Card Sorting Test alone (see *Sensitivity of the test,* above), but they may worsen the deficit observed. This point will be further discussed below.

Conclusions of the Functional Analysis

The functional analysis has isolated at least 3 cognitive components involved in the Wisconsin Card Sorting Test: the ability to change the current rule rapidly

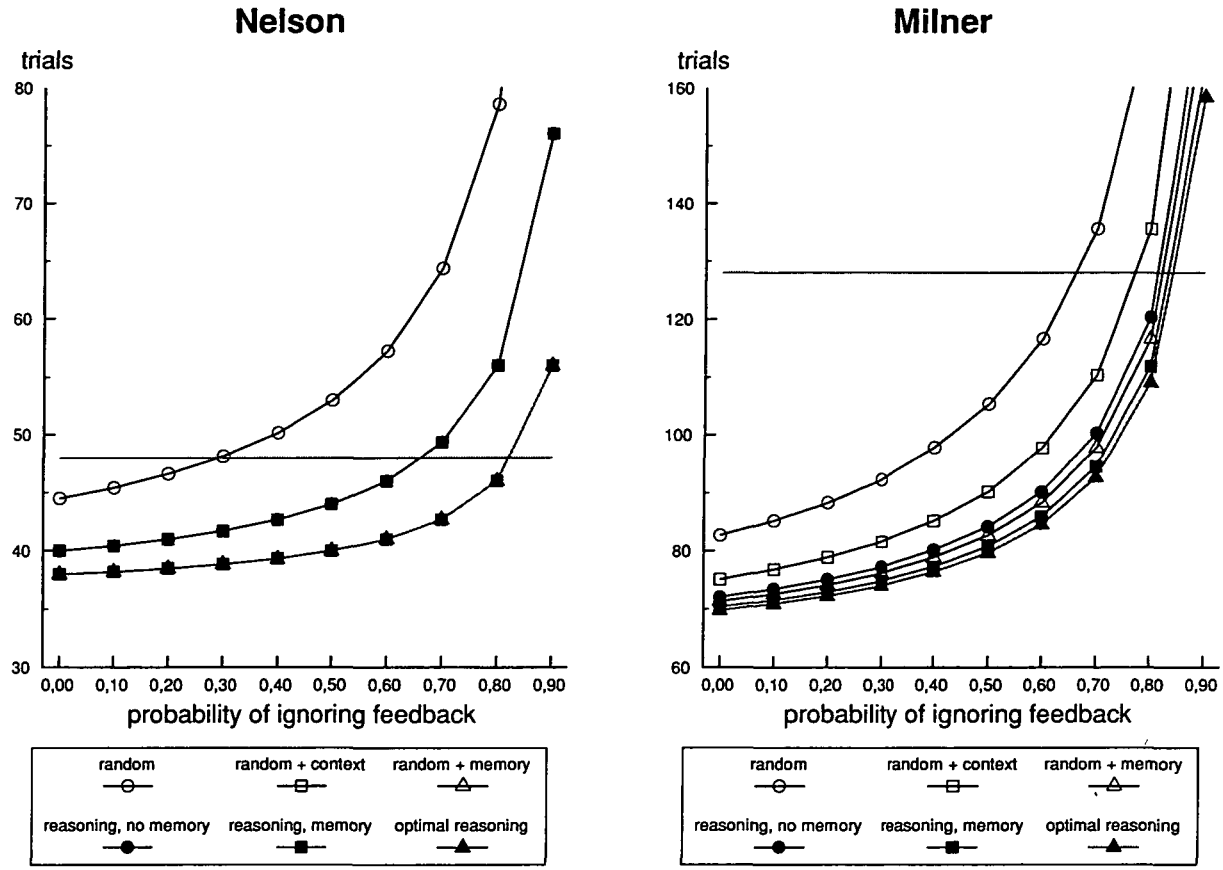

Figure 4. Theoretical test durations for the different machines as a function of the probability of ignoring the reward. Performance deteriorates rapidly, regardless of cognitive architecture, if a machine fails to change its current rule when it receives a negative reward.

when a negative reward occurs, the ability to memorize previously tested rules and to avoid testing them twice, and the ability to reject some rules a priori by reasoning on the possible outcomes of using one rule or the other. It has often been proposed that the frontal cortex is involved in these 3 functions, so it is perhaps not surprising that the Wisconsin Card Sorting Test is so sensitive to frontal damage. On the other hand, our analysis suggests that only ability 1—correct processing of negative rewards—is critically assessed by the test. Assessing episodic memory and reasoning abilities requires more sensitive tests.

The value of such a formal analysis is to provide boundary conditions for the implementation of the above-defined processes or machines. However, many such implementations are compatible with the computational description. At this stage, the actual psychological mechanisms by which rule selection and reasoning take place are left unspecified, and at the neural level, the precise neuronal circuitry of each cognitive component remains to be identified. The next step, then, is to discover, or at least hypothesize, the particular implementation used in the human brain. In the following section, we introduce a model neuronal network resting on biologically plausible principles and able to pass the Wisconsin Card Sorting Test. The purpose of this modeling is not to attempt a better quantitative fit to psychological data on the test, since the presently available data are at best qualitative and are adequately captured by the above formal analysis. Rather, we aim at introducing experimentally testable hypotheses about the neuronal and molecular mechanisms for rule selection, episodic memory, and reasoning.

A Neuronal Architecture that Passes the Wisconsin Card Sorting Test

We shall first describe the anatomical architecture of the network, then its dynamics and its normal functioning. Finally, we shall examine the behavioral consequences of lesioning parts of the model.

Neuronal Clusters and Synaptic Bundles

The units of our model neuronal network are neuronal clusters linked by bundles of excitatory and/or inhibitory synapses, as described in Dehaene et al. (1987). The clusters are viewed as modeling the widespread columnar organization of cerebral cortex (Edelman, 1978, 1987, 1989; Mountcastle, 1978; Goldman-Rakic and Schwartz, 1982; Goldman-Rakic, 1984). Each cluster is composed of hundreds of neurons densely interconnected by excitatory synapses. The detailed connectivity of neuronal clusters is not explicitly formalized (but see below). Rather, each cluster is modeled as a unit with sigmoid response curve and a strong auto-excitatory connection. Unlike single neurons, individual neuronal clusters may possess, depending on their threshold and their auto-

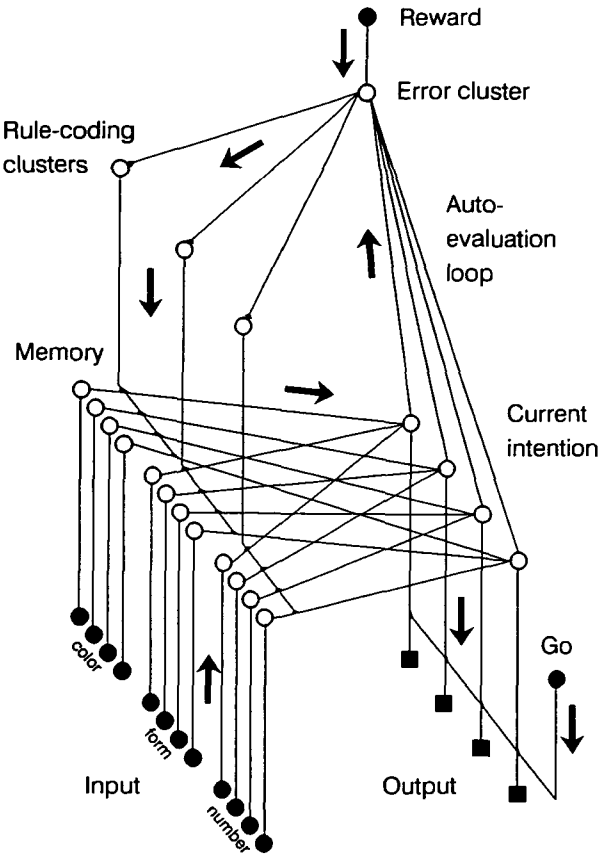

Figure 5. Schematic architecture of a model neuronal network able to pass the Wisconsin Card Sorting Test. *Filled circles* represent the input clusters, *filled squares* represent the output clusters, and *empty circles* represent the internal clusters. Each *line* represents a bundle of synapses; auto-excitatory connections and lateral inhibitory connections within each assembly are not shown. On the input side, cards are coded along the dimensions of color, form, and number, and their features are stored in the short term as a pattern of activation in memory clusters. Memory clusters in turn activate the clusters defining the current intention for output. Rule-coding clusters modulate this transmission between memory clusters and current intention clusters, thus effectively deciding on the sorting rule. When the *go* unit is activated, activity coding for the current intention is in turn transmitted to output units. The subsequent entry of positive or negative reward (*top*) selects among the possible states of activity of the rule-coding layer, until the appropriate rule-coding cluster is activated. In the absence of an external reward, an auto-evaluation loop enables the system to reject rules autonomously, by evaluating the current intention with respect to memorized situations.

excitation strength, two stable states of activity: low-level firing and high-level firing. Once activated, they may thus keep, through self-excitation, a sustained level of firing long after the input has ceased.

Clusters with similar coding properties (e.g., those coding for the color of stimuli) are grouped in neuronal assemblies. By this term we mean an ensemble of neuronal clusters that inhibit each other. The level of inhibition ensures that only one cluster will be active within each assembly at any given moment.

The model assumes that individual pieces of information relevant to behavior are coded by individual neuronal clusters ("grandmother" coding). Indeed, neurophysiological experiments have invariably found that a given neuron in prefrontal cortex repeatedly codes the same parameter value, such as a given location (Funahashi et al., 1989, 1990), with a high specificity. Our cluster representation also implies that several hundred neurons encode similar, though perhaps not strictly identical, information in parallel. Thus the representation may resist degradation. Nevertheless, "grandmother cluster" encoding was adopted in the model only for the sake of simplicity. Functioning depends mainly on different parameters, such as location or color, being coded in separate neuronal populations, but individual parameter values (e.g., a particular location) might have been coded by distributed activity patterns without much affecting the simulation.

Architecture and Functioning of the Model

The organization and connectivity of neuronal assemblies in the model is shown in Figure 5. We shall describe the assemblies in the order in which they are activated in the course of a test trial (Fig. 6).

At the input, response cards are coded according to their features along the dimensions of color, form, and number, by specific feature-coding clusters. This coding scheme is inspired by the known anatomical segregation of dimensions in the visual system (Ungerleider and Mishkin, 1982; Livingstone and Hubel, 1988; Zeki and Shipp, 1988). Input activations are transmitted topographically to memory clusters, which maintain a sustained level of activity even when the input that activated them is suppressed. Cells with such prolonged firing have been observed in dorsolateral prefrontal cortex (e.g., Fuster, 1973, 1989; Niki, 1974) and are assumed to maintain in short-term memory a representation of the external world (Niki, 1974; Goldman-Rakic, 1987). In the model, all clusters corresponding to an input dimension (e.g., color) compete and inhibit each other, ensuring that only 1 feature is memorized for each dimension. Thus, new memories erase the preceding ones.

Memory clusters project topographically to a layer of units coding for the current "intention" of response. The term intention is used here in the restricted operational sense of a prolonged activation predicting subsequent motor outputs. "Intention cells" have been recorded electrophysiologically in prefrontal, premotor, and motor cortex (e.g., Niki, 1974; Niki and Watanabe, 1976; Georgopoulos et al., 1989). These cells fire long before a movement occurs and respond differentially for different movements. Some show a progressive increase in firing rate (response anticipation; e.g., Kojima and Goldman-Rakic, 1982), and their activity predicts the occasional errors that the animal makes (e.g., Fuster, 1973).

The model includes 4 intention clusters, each coding for the choice of a particular reference card in Figure 1. The activation of any such cluster signals an intention of choosing the corresponding reference card. Again, competition ensures that only 1 intention cluster is activated at a given moment. The intention is actualized only when activation is propagated to the actual output clusters. Propagation is gated by the activity of a "go" unit. Hence, a response is given only when an external go signal is received. The go cluster might also be activated by endogenous decision pro-

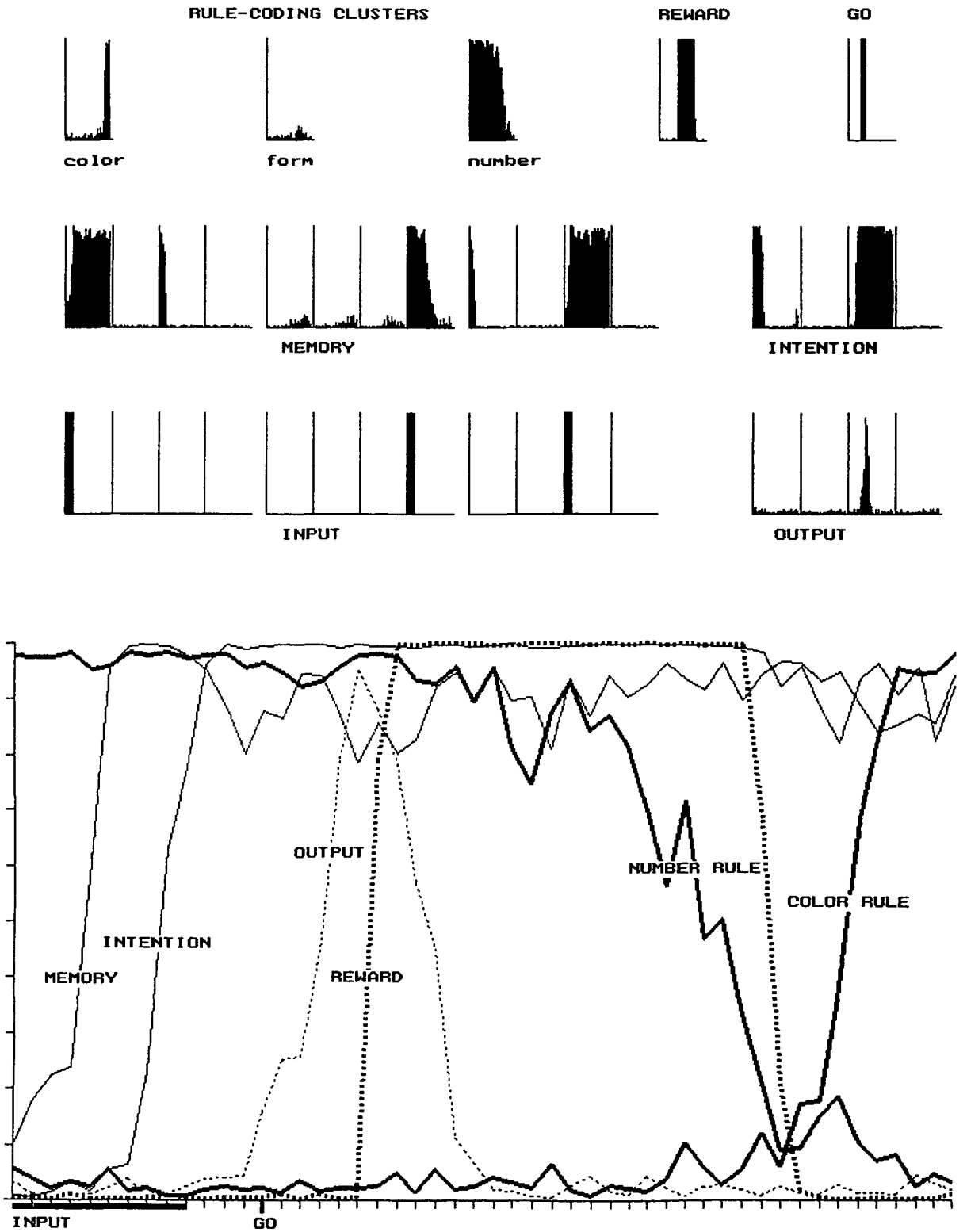

Figure 6. Time course of activation of the different clusters during a typical test trial. *Top*, Each box represents the activity of one neuronal cluster over time. Three input clusters, coding for the stimulus card features along the dimensions of *color*, *form*, and *number*, are activated initially. Their activation is transmitted to the corresponding memory units, whose activity level remains high even after stimulus extinction (except for 1 memory failure in the form-coding assembly). Previous memories are erased. Since in this particular trial the rule-coding cluster corresponding to number is active, the activity pattern among number-coding memory clusters is transmitted to the clusters coding for the intended output (again, replacing any previous activation). Later, the *go* unit is activated by an external signal, potentiating the transmission of activation from intention to actual output clusters. A response is recorded, which is scored as incorrect by the experimenter. Upon reception of the negative *reward* signal, the number rule becomes destabilized, and the cluster coding for the color rule is eventually activated: the next card will be sorted according to color, not number. *Bottom*: Superimposed activation curves for relevant clusters in the same trial, showing the relative onsets of the various trial and neuronal events.

cesses, but we do not use this possibility in the following simulations.

The critical function of the network—to sort the cards according to 1 out of 3 possible criteria—is performed by the rule-coding clusters. Each rule-coding cluster codes for a particular sorting rule (color, form, number, etc.) and gates the corresponding subset of memory-to-intention connections (Fig. 5). Memorized information about color, for instance, influences the response of the network only if the connections for color are potentiated by the color rule-coding cluster. Hence, the pattern of activity over rule-coding clusters determines what sorting rule will be used. In most of the following simulations, only the 3 base rules of color, form, and number are coded in the network. However, additional rule-coding clusters are sometimes included. These clusters then modulate a randomly chosen subset of the memory-to-intention connections.

The existence of rule-coding neurons is an original prediction of the model. These neurons would keep a sustained activity across several experimental trials in which the animal uses the same rule of behavior, even if the stimuli and the action taken on each particular trial vary. They would change their firing only when the animal changes its rule of behavior. Since in most experiments of neuronal recordings the animal is trained to perform a single task, it is perhaps not surprising that such neurons have not yet been unambiguously identified in prefrontal cortex. However, Thorpe et al. (1983), recording in orbitofrontal cortex, found several classes of units, among which 1 might be classified as a rule-coding neuron. This neuron fired on each trial of a go–no go task when the response contingency (the rule) was blue = go, green = no go. It remained silent when the contingency was reversed (blue = no go, green = go). This isolated finding remains to be replicated and extended to more complex situations, by recording during the training phase or while the animal is switching between tasks.

For our network, finding the correct rule consists of selecting 1 of the possible states of activity of rule-coding clusters. This selection is achieved by the reward signal. Each time a wrong response is chosen, the network receives a negative reward, i.e., an "error" cluster is activated. The biological inspiration for the error cluster is the existence of error-coding cells in prefrontal cortex, which fire selectively after the animal makes an error or fails to receive juice (e.g., Niki and Watanabe, 1979).

Activation of the error cluster transforms the rule-coding network into a "generator of diversity." Under normal principles of operation, reception of a negative reward destabilizes the currently active rule-coding cluster; the rule-coding clusters then enter into a competition until 1 of them wins and becomes active again. Hence, the reception of negative reward generally yields the choice of a new current rule. Figure 6 shows this process occurring during a negatively reinforced trial: when the reward is received, the color rule is replaced by the form rule. We now describe, in some detail, the operating principles of this generator of diversity.

The Generator of Diversity and Episodic Memory

We assume that the error cluster modulates the connections of rule-coding clusters in such a way that negative reward leads to a short-term depression of currently active excitatory connections among rule-coding clusters. The exact equations simulating this process are given in Appendix B. The depression mechanism ensures that the auto-excitatory connections of the currently active rule-coding clusters weaken. If connection strength drops below a certain threshold, then auto-excitation is no longer sufficient to maintain a sustained level of activity, and the cluster becomes inactive. The other rule-coding clusters are thus disinhibited. Noise in the simulation ensures that one cluster will eventually reactivate and inhibit all the others. Functionally, the net result is a generator of diversity: the current rule changes at random when sufficient negative reward is received.

Depressed connections spontaneously recover their normal strength. The speed of recovery is a crucial parameter that governs the memory span of the generator. If recovery is very fast, then the rule-coding cluster that was just eliminated immediately re-enters in competition with the other clusters. Hence, there is a finite probability of not changing rules at all. This provides an implementation of the random machine analyzed formally in our functional analysis above. In contrast, if recovery is slower, then the competition takes place only between the remaining rule-coding clusters, leaving aside the 1 just eliminated. This corresponds to our random + context machine. Finally, if recovery is very slow and extends over several consecutive trials, the network exhibits an episodic memory and behaves as the random + memory machine described above. Eliminated rules do not enter the competition again; the generator of diversity keeps track of previously rejected rules and does not activate them any longer.

Functionally, we observe that a network with slow recovery serially tries the 3 rules of color, form, and number until the correct rule is found. In contrast, a network with fast recovery may try sequences such as color, color, form, color before eventually settling into the correct number rule. Figure 7 provides a quantitative comparison of networks with three different values for recovery rate ($\sigma = 0.95$, fast recovery = random machine; $\sigma = 0.97$, intermediate recovery = random + context machine; $\sigma = 0.99$, slow recovery = random + memory machine). In agreement with our functional analysis of the test (above), performance increases with memory span. Compared to the other two machines, the random + memory machine learns faster, is more likely to find the correct rule in a single trial, hardly ever perseveres in an incorrect rule, and avoids retesting a rule once it was shown to be incorrect. Our network thus offers an elementary implementation of episodic memory for the previous

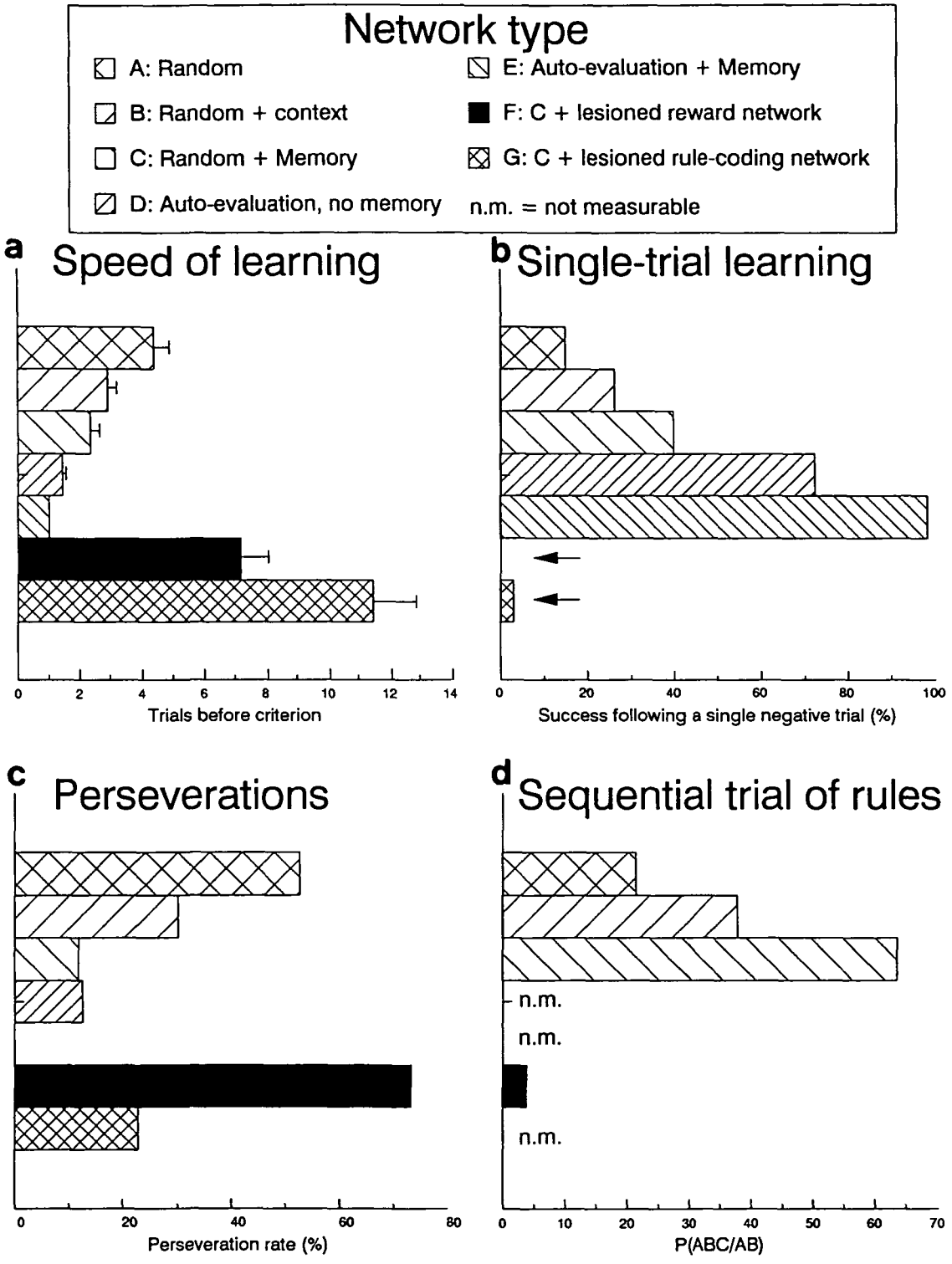

Figure 7. Comparison of the performances of several intact or lesioned networks. For each network type, 500 trials were simulated using the modified version of the test described in the text, with a criterion of 3 successes in a row. *a*, Number of trials necessary before reaching criterion. Note the improvement in performance with machine complexity, from the basic random machine *A* to the most complex machine *E* with auto-evaluation and episodic memory. Lesioned networks perform poorly, especially the one with lesioned rule-coding layer (*G*), which must reconstruct the sorting rule by slow synaptic modifications. *b*, Percentage of immediate success following a single negative trial. This display is essentially the reciprocal of the preceding one. Note the quasi-optimal performance of machine *E* (98.4% immediate success), and the floor effect for lesioned networks. *c*, Perseveration rate, measured as the percentage of negatively rewarded trials that were not followed by a change of rules. Lesioned networks perform much worse than the comparable intact machine *C*. Also note a nonnegligible perseveration rate for machine *A*, which lacks a memory of the previous rule tried. *d*, Evaluation of episodic memory. We measured the probability *P(ABC/AB)* that the third rule *C* be tried following 2 consecutive negative trials with different rules *A* and *B*. This measure estimates the capacity of a network to try consecutively all 3 available rules, and to avoid testing the same rule twice. Note the improvement with increasing memory capacity for the first 3 machines. Measurements for machines with auto-evaluation (*D* and *E*) were not measurable, since 2 consecutive failures occurred only rarely, and the sequential trial of rules occurred covertly between trials. For the lesioned network *F*, failure on this measure of memorization merely results from the high level of perseverations. Performance of network *G* could not be measured because this network often made choices altogether inconsistent with the 3 base rules.

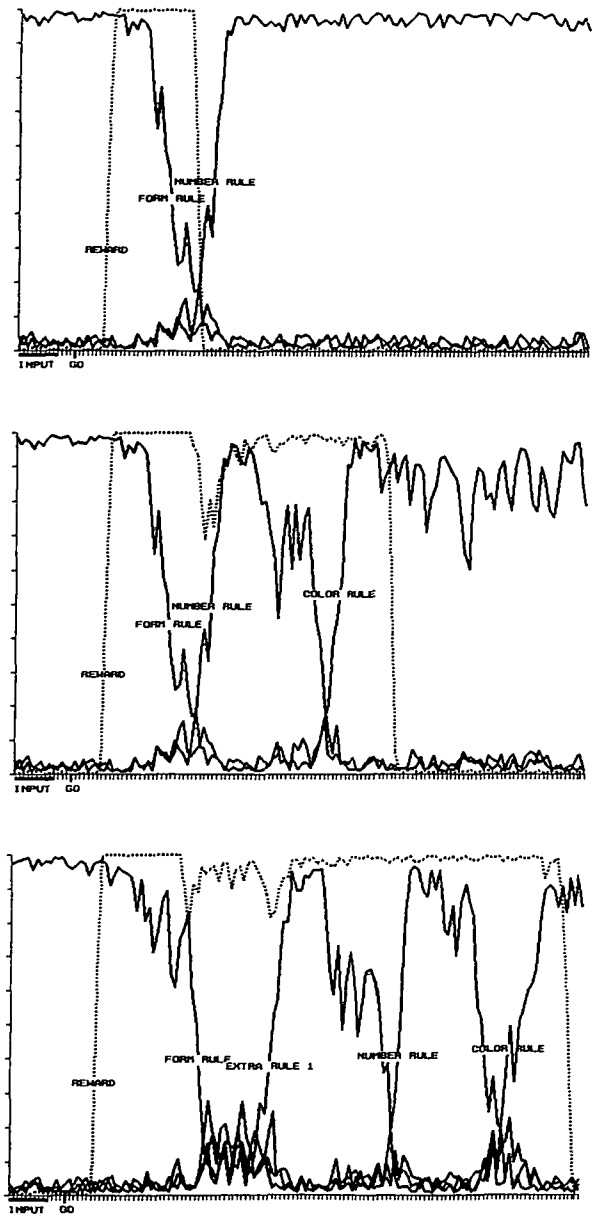

Figure 8. Comparison of the activity of rule-coding clusters in networks with or without the auto-evaluation loop. *Top,* The simple transition from the form rule to the number rule following negative reward in a network without auto-evaluation. *Middle,* Subsequent events in a network with auto-evaluation. Because on this particular trial the number rule led to the same incorrect choice as the form rule, the reward cluster remains activated by the intention-to-error connections, until the number rule eventually gets replaced by the correct color rule. *Bottom,* What happens when additional rules are added to the repertoire of the network: the sequential trial of rules continues until a satisfactory rule is found—here the color rule.

choices of the network and their associated reward values.

Reasoning and the Auto-Evaluation Circuit

Up to this point, we have described the implementation of two cognitive components: the generator of diversity, responsible for trying new rules, and the episodic memory, responsible for ensuring that rules are not tried again once they have proven to be incorrect. We now propose an implementation of the third component, the reasoning process, by which rules may be eliminated a priori by evaluating in advance their possible outcomes. In the model, this role is subsumed by the auto-evaluation loop. As shown in Figure 5, this loop short-circuits the external entry of reward and permits an endogenous activation of the error cluster. The network then functions as a critic of its own choices, and corrects itself until a plausible rule is found, one that is compatible with the previously stored knowledge of positively or negatively rewarded situations.

In detail, the auto-evaluation loop consists of direct intention-to-error connections whose efficacies change rapidly according to a classical Hebbian rule (equations in Appendix B). When no reward is given, these connections relax toward 0 and, hence do not significantly affect the simulation. However, when negative reward is received and the error cluster is activated, the connection linking the currently active intention cluster to the error cluster is strengthened. Hence this intention is labeled "incorrect." As described previously, the error signal also triggers the choice of a new current rule. During the intertrial period, this new rule, applied to the memorized features of the previous input card, yields a new pattern of activation over intention units. This pattern codes for the action that would have been undertaken had the new current rule been active on the previous trial. If it is in fact the same pattern as before, this means that, placed in the same situation, the new rule would have led to the same incorrect choice as the previous rule; hence, the new rule must also be incorrect. The auto-evaluation loop then keeps the error cluster active through the potentiated intention-to-error connection, thereby preventing the new rule from stabilizing. The sequential internal evaluation of rules will continue until a plausible rule is found, that is, a rule whose consequences, positive or negative, cannot be known in advance on the basis of the previous trial.

Figure 8 compares the evolution of activity in rule-coding units when networks with or without auto-evaluation are simulated. In all cases, negative reward initially triggers a change of the current rule. However, only networks with auto-evaluation are able to predict that the new current rule also yields an incorrect choice. The auto-evaluation loop then maintains the activation of the error cluster, triggering as many rule changes as needed for a plausible rule to be found. The resulting sequential evaluation of rules constitutes an implementation of an elementary reasoning process.

Figure 7 provides a comparison of performance for networks with and without the auto-evaluation loop. As noted in the initial analysis, the Wisconsin Card Sorting Test is only weakly sensitive to differences in cognitive architecture. To render the improvements due to the auto-evaluation loop more visible, we designed a variant of the test that better separates reasoning from nonreasoning machines. This variant uses only the 36 cards for which 2 rules yield the choice of the same reference cards, and the third rule yields a different choice. For example, a card with 2 red

triangles is similar to the first reference card in both color and form, but is identical in number to the second reference card (Fig. 1). With such cards, a machine that reasons can find the correct rule in only 1 trial: as soon as negative reward is received, all but one sorting rule can be eliminated. In contrast, a non-reasoning machine should reach about 50% success following the first negative trial.

The comparisons of network types that appear in Figure 7 were performed with this new version of the test. The crucial measure to evaluate reasoning abilities is the percentage of success following a single negative trial (single-trial learning). Clearly, the network with auto-evaluation and memory scores better than the corresponding random + memory network (98.4% vs. 39.8%). Similarly the network with auto-evaluation but no memory scores better than the associated random + context network (72.3% vs. 26.2%).

Lesioning the Network

The most crucial evaluation of the model concerns its ability to reproduce the behavior of brain-lesioned patients following ablation of some of its components. We consider 3 types of lesions to our network.

Lesioning the reward network. We stimulate a disruption in the error-processing circuit, in an otherwise intact random + memory network, by a weakening of the input to the error cluster (+3 instead of +6 on incorrect trials). The net result is an increase in perseverations, similar to what is observed in frontal patients (Fig. 7). The current rule is often not changed following a single negative reward. Rather, it takes several negative rewards in a row before the rule is destabilized. In essence, the generator of diversity still works, but with abnormal inertia. As a consequence, single-trial learning totally disappears, and it now takes much longer for the network to reach criterion.

Lesioning rule-coding clusters. A more radical lesion consists of totally eliminating the rule-coding clusters. We then assume that the memory-to-intention connections have a fixed short-term component $S_{ij} = 0.5$, and that their efficacy varies in the long term according to a Hebbian rule whose sign is modulated by the sign of the reward (see Appendix B for technical details). Slow learning by correlation of input and output is then the only means of adaptation to the task. The main behavioral consequence is lack of systematicity (Fodor and Pylyshyn, 1988). The lesioned network can no longer treat all instances of a variable like color in the same regular way. Rather, the responses to each of the 4 possible colors have to be learned separately at different times. Hence, learning is slowed down by about a factor of 10 (Fig. 7). Performance in the course of learning appears chaotic. The network sorts in a quasirandom manner by color, form, or number.

Lesioning the auto-evaluation loop. The above lesions have a dramatic effect on behavior, with increased perseverations and slowed learning. In contrast, our model predicts that some focal lesions may have almost undetectable effects on performance. Such is the case with a lesion to the auto-evaluation loop only. The only consequence of such a lesion is the loss of the reasoning ability. In Figure 7, this corresponds to going from network *E* to network *C*. Clearly, neither learning speed nor perseverations are much affected by such a lesion. As discussed earlier, the only variable that is importantly affected is single-trial learning, and this should be detectable only with our modified version of the Wisconsin Test. Of course, more complex tests such as the Tower of London (Shallice, 1982) would be more apt at detecting the difficulties in planification and self-monitoring of behavior that are expected following disruption of the auto-evaluation loop.[2]

Finer-Scale Implementations

Despite its biological inspiration, the above model is rather abstract. It is therefore of importance to show that its mechanisms are not purely ad hoc, but may be implemented quite naturally in biological hardware. In this section, finer-scale implementations are considered for 3 crucial components of the network: the bistable neuronal clusters, the generator of diversity, and the reward-dependent synaptic depression rule for rule-coding clusters.

Bistable Neuronal Clusters

In the above model, neuronal clusters were modeled as a single threshold unit with strong auto-excitatory connection. A more realistic simulation of a single cluster composed of 49 individual neurons was performed. Each neuron was modeled as a McCulloch and Pitts (1943) unit with sigmoidal response characteristic. Neurons could be either excitatory (about 75%) or inhibitory (about 25%), but not both. Each neuron was randomly connected to 10 others with strength ± 1.5, and a random threshold between 2 and 4. About half of the neurons sometimes also received an external excitatory input of strength 3. Noise level was ± 1.5.

Figure 9 shows the individual histograms of the 49 neurons before, during, and after external stimulation of the cluster. Different types of units were observed. Some responded only during stimulation, or did not respond at all. Most units, however, kept a sustained level of activation for a long time following the external stimulation. Finally some units responded only after the stimulation, but not during it. It is thus possible to reproduce in this elementary model the full variety of unit types that are found in single-cell recordings in frontal cortex during delayed-response tasks (e.g., Funahashi et al., 1989, 1990).

In the simulation, the intrinsic firing characteristics of the units were identical (except for their threshold). The type of activity exhibited during and after stimulation depended only on the recurrent connection patterns established at random between neurons. Such reverberant circuits offer a plausible explanation for the sustained firing of prefrontal "memory cells." However, we cannot exclude the alternative possibility that some specific cell property

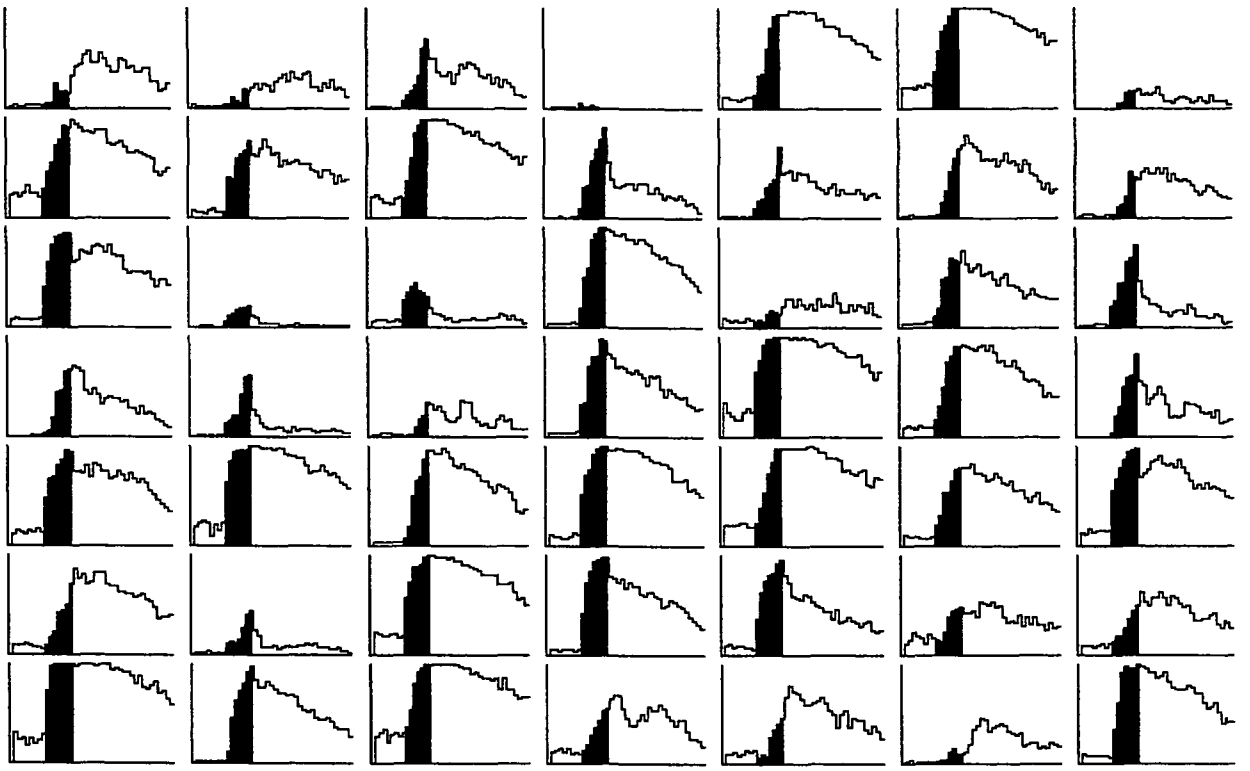

Figure 9. Fine-scale simulation of a cluster of 49 neurons. Histograms show the activity of each simulated neuron before, during (*shaded area*), and after external stimulation of the cluster. Most neurons are activated by the stimulation and remain active for a long period thereafter. Some neurons show no activation at all (*top row, 4th from left*), activation only during stimulation (*3rd row, 3rd from left*), or only after stimulation stopped (*bottom row, 6th from left*).

is also involved that would distinguish cells with long-lasting firing from simple reactive cells. Isolated cells with bistable firing characteristics have been studied, for instance, in the spinal cord (Eken et al., 1989).

Generator of Diversity

Several clusters were assembled to model the generator of diversity at the level of individual neurons and synapses (Fig. 10*A*). Again, the neurons, modeled as McCulloch and Pitts (1943) units, were either excitatory or inhibitory, but not both. Excitatory neurons were organized in clusters with strong intracluster excitation and weak intercluster synapses (see M. Kerzsberg, S. Dehaene, and J. Changeux, unpublished observations, for a biologically plausible account of the epigenesis of such clusters). Excitatory neurons also contacted inhibitory interneurons, which in turn distributed inhibition randomly to many excitatory neurons (for simplicity, there were no synapses between inhibitory neurons, though this did not significantly affect the behavior of the simulation). Finally, synaptic efficacy was assumed to vary according to the rules defined in the previous section: the efficacy decreased for active synapses when negative reward was received, and later recovered slowly.

We expected that in the absence of negative reward, the network would settle into a stable state with one cluster active and the others silent. However, this simple property was not obtained immediately. Rather, the network was unstable, oscillating between hyperactive and silent states. Oscillations arose because of the temporal lag between excitation and inhibition in the network. As is visible in Figure 10*A*, it takes only one synapse for excitatory neurons to excite each other, but two synapses are necessary for excitatory neurons to inhibit each other via any inhibitory interneuron. As a result, the network showed synchronized bursts of "epileptic" activity, followed by silent periods of variable length. Subsequent experimentation with the network showed that to obtain stability, the disynaptic inhibitory pathway had to be at least as fast as the monosynaptic excitatory pathway. Indeed, electrophysiological recordings of cortical neurons reveal that inhibitory cells may have smaller membrane time constants than excitatory ones (McCormick et al., 1985). Our simulation points to the importance of this differential speed for network stability. Similar bursts of activity following blockade of fast inhibition have been observed independently in a simulation of the hippocampus (Traub et al., 1988).

With appropriate connection strengths and transmission speeds, the network initially stabilized into a state with one neuronal cluster active and the others silent. We then observed that negative reward of sufficient duration triggered transitions of activity from one cluster to the other. As in the larger-scale simulation, the randomness of the sequence of transitions was determined by the time for a depressed synapse to recover its initial efficacy. If recovery was fast, the sequences were random. If recovery was slow, the sequences tended not to repeat the same activations in close succession. Hence, the properties that were

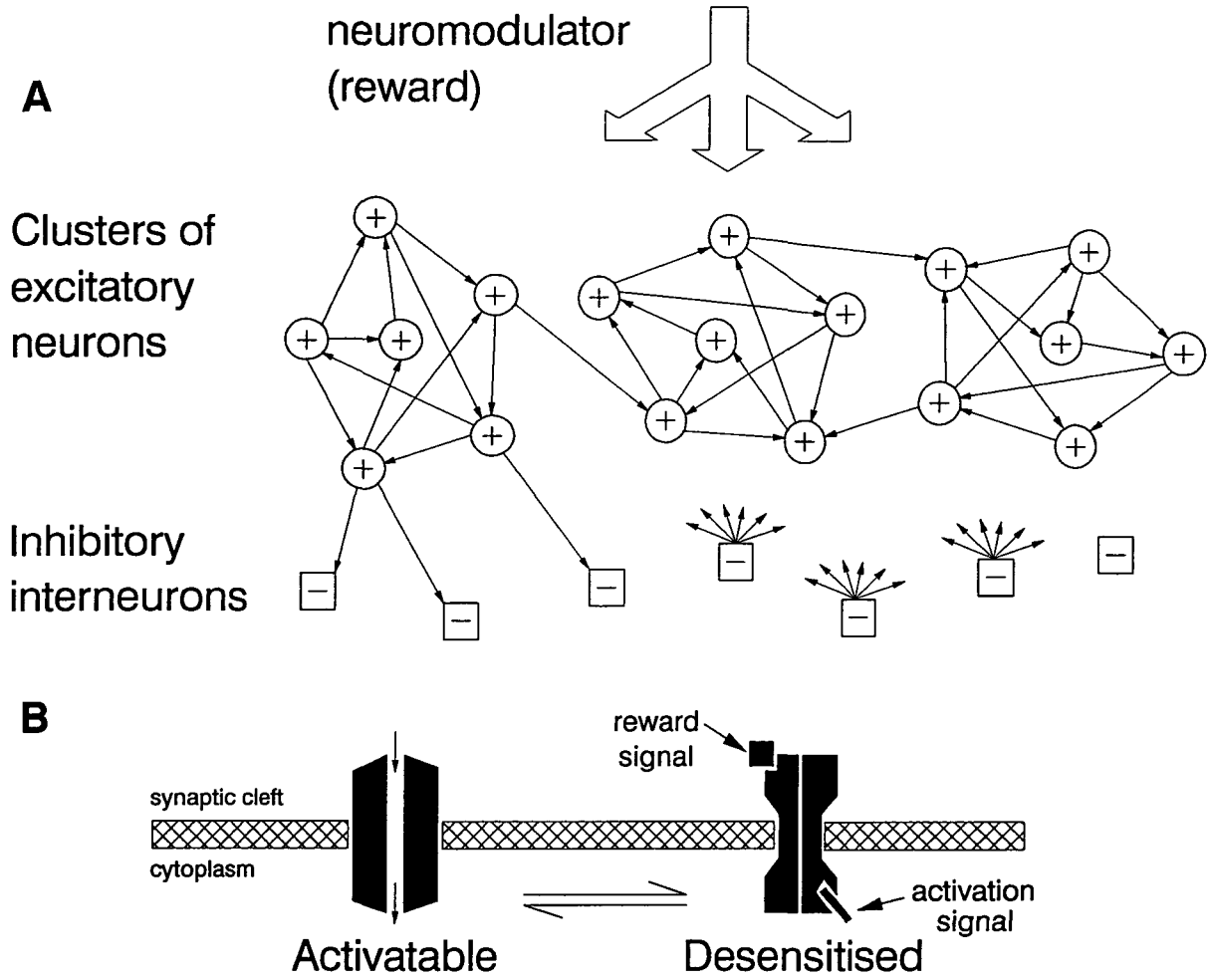

Figure 10. Plausible finer-scale implementations for the generator of diversity. *A*, Fine-scale description of the architecture of the generator. Neuronal clusters consist of strongly interconnected excitatory neurons. Clusters inhibit each other *via* inhibitory interneurons. The reward signal enters as a diffuse, modulatory input to all neurons. *B*, Possible molecular implementation of the modification of synaptic efficacy responsible for rule destabilization following negative rewards. Postsynaptic receptor molecules normally exist in an activatable state where the ion channel opens upon fixation of the neurotransmitter. Receptor molecules switch to a desensitized state, with ion channel blocked, upon simultaneous reception of the 2 messengers signaling negative reward and recent activation of the synapse. The activation signal is tentatively shown as a postsynaptic intracellular event (e.g., calcium concentration), but it might as well be extracellular (e.g., neurotransmitter or coexisting peptide concentration).

crucial to the generator of diversity, with or without episodic memory, were confirmed in this more elementary simulation.

Reward-Dependent Synaptic Depression Rule

We may then move one step further and consider which molecular implementations may account for the key property of the above network, namely, the fast depression of active synapses when negative reward is received, followed by medium to slow recovery of normal synaptic strength. Our hypothesis is that fast synaptic depression may result from the desensitization of receptor molecules, mediated by their allosteric transitions in the postsynaptic membrane. The receptors for neurotransmitters indeed are allosteric proteins (Changeux, 1981, 1990) that may carry multiple topographically distinct binding sites and exist under several conformational states. For example, the nicotinic acetylcholine receptor is known to exist under at least 4 interconvertible states: resting (*R*), active (*A*), rapidly desensitized (*I*), and slowly desensitized (*D*), where the ion channel is open only in state *A*. The transitions towards states *I* and *D* take place in time scales from 10 msec to several minutes, and accordingly may regulate synaptic efficacy. Transitions towards the desensitized states *I* and *D* yield a depression of synaptic strength, whereas the transitions towards the resting state *R* potentiate synaptic strength. The signals that may regulate transitions between allosteric states are of several kinds: electrical potential, local concentration of calcium, second messengers, covalent modifications, or neuromodulatory substances from outside the postsynaptic cell (for review see Changeux, 1990). These signals may originate from a second, neighboring synapse on the same postsynaptic cell, opening the possibility of regulation of the efficacy of a given synapse by the activity of another synapse (heterosynaptic regulation; see Heidmann and Changeux, 1982; Changeux and Heidmann, 1987; Dehaene et al., 1987; Finkel and Edelman, 1987).

For our generator of diversity, we may assume that the desensitization reaction is enhanced by the co-occurrence of two signals converging on the same

postsynaptic receptor molecule (Fig. 10*B*). The first one, endogenous to the postsynaptic cell, signals the recent activation of the synapse. This role may be assumed, for instance, by the high local intracellular concentration of calcium or a high extracellular concentration of neurotransmitter or of a coexisting messenger (Hökfelt et al., 1986). The second signal, diffused to all synapses throughout the relevant network, indicates a recent negative reward. Such gating of synaptic modifications by reward may be achieved, for instance, by diffuse neuromodulatory projections of catecholamine neurons from the mesencephalon to the prefrontal cortex (Bloom, 1988; Fuxe et al., 1989; Fuxe and Agnati, 1990). The simultaneous reception of these 2 converging signals would trigger a conformational change of receptor molecules into a state where the ion channel is closed (and the synapse is thus depressed). Recovery by the reverse reaction would occur on the 0.1–1 sec time scale. Such values have been observed in the case of the cerebellum parallel fiber–Purkinje cell synapse by Ito (1989).

It should be stressed again that several alternatives exist to our proposed molecular mechanism. For instance, NMDA receptors also permit the integration of convergent molecular signals in a manner consistent with our assumptions (Gustafsson and Wigström, 1988; Zador et al., 1990).

Discussion

We have provided a theoretical analysis of the Wisconsin Card Sorting Test, defined a hierarchy of machines able to pass the test, and compared their relative efficiency. A plausible neuronal implementation for the machines has been described, the hierarchical architecture of which is compatible with the organization and specialization of cortical areas. On the finer scale of individual neuronal circuits, the units of our neuronal model are neuronal clusters plausibly homologous to cortical columns, and their function derives directly from the collective behavior of their component neurons. At the molecular level, a plausible implementation for the synaptic modification rules used in the model is presented in terms of allosteric transitions of postsynaptic receptor molecules.

Another network model for the Wisconsin Card Sorting Test has been proposed by Levine and Prueitt (1989). The architecture of their network has similarity with ours and incorporates an interesting distinction between a level of habits and another level of biases that may modulate these habits. However, we have introduced 2 novel components, episodic memory and reasoning abilities. Lacking these, Levine and Prueitt's model performs no better than our simplest machines ("random" or "random + context").

An important issue is the extent to which the present model may handle more complex tasks. A network with an architecture close to the present one was previously shown to account for behavioral and electrophysiological data in delayed-response tasks (Dehaene and Changeux, 1989). The episodic memory and the reasoning component introduced in the present model are likely to be crucial in several other complex tasks tapping frontal cortex functions, such as the Tower of London Test (Shallice, 1982). This work thus represents an additional step towards the neuronal implementation of the frontal supervisory system that supports planning and task monitoring abilities (Shallice, 1988).

The architecture of our network, in its present formulation, imposes severe limitations on what can be learned: only the 3 base rules of sorting by color, form, or number are learnable. While this may be somewhat extreme, the network, like the brain, does not behave as a tabula rasa on which any possible environmental regularity can be imprinted. Learning by selection, as implemented here or in other models (Barto et al., 1983; Edelman, 1987, 1989) implies that storage of external regularities can only take place through stabilization or elimination of predefined forms called prerepresentations (see Changeux et al., 1984; Changeux and Dehaene, 1989). The performance is thus limited by the initial repertoire of learnable rules. Learning power is, however, traded against systematicity: in our model, the correct rule may be found in very few trials and immediately generalizes in a systematic fashion to subsequent trials, a feature that is lacking in classical associationist neural nets (Fodor and Pylyshyn, 1988).

Cohen et al. (1990) have proposed a PDP model of the Stroop effect that combines a fixed, constrained processing architecture and the powerful learning abilities of the back-propagation algorithm. Their network includes modulatory "task demand" units that bear similarity to our rule-coding clusters. Thus, like our color cluster, one unit labeled "color naming" can systematically attenuate or amplify transmission in the color processing pathway (see also the "modifiers" and "conjunctive connections" of Feldman and Ballard, 1982). But, unlike our model, the networks controlled by such units are initially unstructured and are trained instructively by back-propagation to map colors or color names to verbal responses. Thus, the color-to-name mapping is captured by an unconstrained learning algorithm, whereas the systematic task requirements are hard-wired in the connections of the "task-demand" units.

Cohen et al.'s (1990) model applies when subjects are taught the task beforehand. In contrast, our model deals with situations in which the subject must discover what task is asked for. The original mechanism that we introduced for this purpose is the auto-evaluation loop. It enables the network to select rules, in the absence of new inputs, by reasoning on the possible outcomes of each rule and comparing them with stored knowledge of previously rewarded situations. Critics often argue that learning will be unmanageably slow if rules are simply picked at random from a large set until one proves to be correct. The auto-evaluation loop accelerates learning by selection by permitting the elimination of implausible rules by mere reasoning, without waiting for an external correction signal. It thus provides a tentative solution to

the combinatorial explosion expected from a simple-minded scheme of learning by selection. On the other hand, auto-evaluation is studied here in a very restricted context where the repertoire of accessible rules is limited from the start by prewiring. Future work will have to examine its adequacy in more open learning situations.

That learning may proceed purely by "mental experiment" is, of course, not new from a cognitive perspective, but our neurally plausible implementation of this ability opens the possibility of probing its neuronal bases electrophysiologically in awake animals. The Wisconsin Card Sorting Test may be difficult to adapt to nonhuman species. However, our model generalizes easily to simpler experimental paradigms such as serial visual search (Treisman and Gelade, 1980) or memory scanning (Sternberg, 1966) that also imply covert sequencing of mental objects in the absence of overt behavioral cues. If animals could be trained with analogs of these tasks, mental processing would then become directly accessible to neurobiological experimentations.

Appendix A: Outline of the Mathematical Calculations

Milner's Version

A first step in the calculations of convergence times for each machine is the determination of $\theta(n)$, which is the convergence time conditioned by the fact that the current rule is incorrect, and that n rules remain to be tried. When $\theta(n)$ is known, convergence time is simply given by the following equations:

$$T_f = (1 - 1/r)\Theta(r),$$

for the time to reach the first criterion;

$$T_n = \Theta(r),$$

for the time to reach subsequent criteria. (The difference between T_f and T_n is due to the probability of finding the correct rule by chance on the very first trial.) Then the total time to complete 6 criteria is given by

$$T_t = 60 + T_f + 5T_n$$

(this is the variable plotted in Figs. 3 and 4).

$\Theta(n)$ itself obeys the following equation, obtained by considering the events that may occur on a given trial:

$$\Theta(n) = \frac{1}{1 - P} + P_c\Theta(n) + (1 - P_c)\sum_{i=0}^{n-1} P_e(i) \cdot [1 - P_f(n - i)]\Theta[\Phi(n, i)],$$

where P is the probability of ignoring reward, P_c is the probability that the current rule, even though it is incorrect, will accidentally yield a correct response and thus will not be eliminated ($P_c = 1/q$); $P_e(i)$ is the probability that i rules will be eliminated from the pool of possible rules for the current trial, given that the current rule was rejected; $P_f(n - i)$ is the probability of finding the correct rule when drawing at random from the $n - i$ remaining rules [$P_f(n - i) = 1/(n - i)$]; and $\Phi(n, i)$ is the number of rules remaining plausible for the next trial [$\Phi(n, i) = n - i$, for machines with episodic memory; n, otherwise].

The functions $P_e(i)$ and $\Phi(n, i)$ vary depending on the cognitive architecture of the machine under study. Let us treat, for example, the case of the random + context machine. Because this machine uses context, but not reasoning, only the current rule can be rejected on a negative trial. Hence $P_e(1) = 1$, $P_e(i) = 0$ for $i \neq 1$. Because the machine has no episodic memory, the rules that have been rejected on a previous trial reenter the pool of possible rules on the next trial, hence $\Phi(n, i) = n$. With these values, the fundamental equation becomes

$$\Theta(n) = \frac{1}{1 - P} + \frac{1}{q}\Theta(n) + \left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{n - 1}\right)\Theta(n).$$

From this we derive

$$\Theta(r) = \frac{q(r - 1)}{(q - 1)(1 - P)}$$

and the corresponding value for T_r.

Similar analytical results were obtained in all cases, except for machines with both reasoning and memory. In the latter cases the fundamental equation was solved numerically by recurrence.

Nelson's Version

The analytical treatment of Nelson's version is more complex. Since the base rules are not independent of each other any more, it is necessary to distinguish whether the current rule is a base rule. Without entering into details, one must calculate $\Theta_b(n_1, n_2)$ and $\Theta_a(n_1, n_2)$, namely, the convergence times conditioned by the fact that the current rule is incorrect, it is a base rule (b) or an additional rule (a), and n_1 base rules and n_2 additional rules remain to be tried.

$\Theta_b(n_1, n_2)$ and $\Theta_a(n_1, n_2)$ obey coupled recurrence equations similar to the above equations. We were able to solve these equations exactly for nonreasoning machines, and numerically for the others. Of course, the equations are much simpler, and exact expressions can be obtained when $r = p$ (that is, if the repertoire is limited to the p base rules).

Appendix B: Formalization of the Model

The dynamics obeyed by each neuronal cluster is as follows:

$$s_i(t + 1) = F\left[\sum_j W_{ij}(t)s_j(t) - T_i + N\right],$$

where $s_i(t)$ is the activity of cluster i at time t, $W_{ij}(t)$ is the efficacy of the synaptic bundle from cluster j to cluster i, T_i is a threshold, N is a noise term with uniform distribution over $[-n, n]$, and F is the sigmoid function

$$F(x) = 1/(1 + e^{-x}).$$

The diagonal terms W_{ii} represent positive intracluster

auto-excitatory connections. The off-diagonal terms W_{ij} $(i \neq j)$ represent either negative lateral inhibition within a given neuronal assembly, or positive transmission of activation between assemblies.

In general, the efficacies $W_{ij}(t)$ may vary with time. This variation is decomposed into a product of short- and long-term components:

$$W_{ij}(t) = S_{ij}(t)L_{ij}(t).$$

Short-Term Component

The short-term component $S_{ij}(t)$ varies between 0 and 1. It represents heterosynaptic influences on the $j \rightarrow i$ connections. We make the hypothesis that 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). Synaptic triads are included in the model by assuming that for some couples (i, j), there is a modulator cluster m such that

$$S_{ij}(t+1) = \begin{cases} \alpha S_{ij}(t) + 1 - \alpha, & \text{if } s_m(t) > 0.5, \\ \alpha S_{ij}(t), & \text{if } s_m(t) < 0.5. \end{cases}$$

Qualitatively, this equation implies that synaptic efficacy increases toward a maximum when the modulator is active, and decreases toward 0 when the modulator is inactive. The current model includes 2 sets of triads: memory-to-intention connections are modulated by the relevant rule-coding cluster, and intention-to-ouput connections are modulated by the go cluster. For other connections, S_{ij} is independent of t $(S_{ij} = 1)$. The auto-excitatory connections of rule-coding clusters obey a slightly different rule for short-term modifications. As described in the text, a negative reward can destabilize active rule cluster by the following mechanism:

$$S_{ii}(t+1) = [\sigma S_{ii}(t) + 1 - \sigma][1 - Q(t)] + \delta\, S_{ii}(t)Q(t),$$

where $Q(t) = [s_r(t)s_i(t)]^2$.

In the above equation, $s_r(t)$ is the activation of the error cluster. Qualitatively, $S_{ii}(t)$ rapidly drops toward 0 when both the error cluster and the presynaptic cluster i are active, and relaxes toward 1 otherwise.

Finally the following equation applies for the intention-to-error connections of the auto-evaluation loop:

$$S_{ir}(t+1) = \begin{cases} \delta S_{ir}(t) + 1 - \delta, & \text{if } s_r(t) > 0.5 \text{ and } s_i(t) > 0.5, \\ \delta S_{ir}(t), & \text{otherwise.} \end{cases}$$

Long-Term Component

For some excitatory connections $W_{ij}(t)$, namely the memory-to-intention connections, the long-term component $L_{ij}(t)$ may also vary with learning according to the following Hebbian rule:

$$L_{ij}(t+1) = L_{ij}(t) - \beta\, s_r(t)S_{ij}(t)s_j(t) \cdot [2s_i(t) - 1].$$

An additional constraint is that the $L_{ij}(t)$ must remain bounded by 0 and an absolute maximum L.

Numerical Parameters

In the simulation, the following numerical values were used: $\alpha = 0.4$, $\beta = 0.4$, $\delta = 0.97$, n = 0.7, and $0.95 < \sigma < 0.99$ depending on the machine simulated. The initial connection strengths and thresholds were also as follows:

S_{ij}	= 0;	
L_{ii}	= +6;	
L_{ij}	= −2,	for lateral inhibition,
	= +3,	for input-to-memory and memory-to-intention connections,
	= +2,	for intention to output connections,
	= +5,	for intention to error connections,
	= +6,	for the input to the error cluster;
T_i	= 3,	for memory and intention clusters,
	= 4,	for output clusters,
	= 2,	for rule-coding clusters,
	= 5.5,	for the error cluster.

Notes

1. This form of memory may be termed "episodic" (Tulving, 1972) since it registers the history of the machine's attempts at solving a problem, and thus forms part of its autobiographical record. Shallice (1988) argues that "the autobiographical record of the organism's previous environments, activities, plans, and intentions has the primary function of acting as raw material for the [frontal] Supervisory System when it is directing some non-routine activity" (page 371). Our model implements Shallice's notion in a restricted situation: the episodic memory of previously rejected rules narrows down the choice of possible rules remaining to be tried.

2. In a much more complex context such a lesion might result in the sociopathic behavior reported by Damasio et al. (1990).

This research was supported in part by INSERM.

Correspondence should be addressed to Dr. Dehaene, Neurobiologie Moleculaire, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris cédex 15, France.

References

Barto AG, Sutton RS, Anderson CW (1983) Neuronlike elements that can solve difficult learning control problems. IEEE Trans Syst Man Cyber SMC-13:834–846.

Bloom F (1988) Neurotransmitters: past, present and future directions. FASEB J 2:32–41.

Changeux JP (1981) The acetylcholine receptor: an "allosteric" membrane protein. Harvey Lect 75:85–254.

Changeux JP (1990) Functional architecture and dynamics of the nicotinic acetylcholine receptor: an allosteric ligand-gated ion channel. Fidia Research Foundation Neuroscience Award Lectures, Vol 4 (Changeux JP, Llinas RR, Purves D, and Bloom FE, eds), pp 21–168. New York: Raven.

Changeux JP, Dehaene S (1989) Neuronal models of cognitive functions. Cognition 3:63–109.

Changeux JP, Heidmann T (1987) Allosteric receptors and molecular models of learning. In: Synaptic function (Edelman G, Gall WE, and Cowan WM, eds), pp 549–601. New York: Wiley.

Changeux JP, Heidmann T, Patte P (1984) Learning by selection. In: The biology of learning (Marler P and Terrace H, eds), pp 115–139. Berlin: Springer.

Cohen JD, Dunbar K, McClelland J (1990) On the control of automatic processes: a parallel distributed processing model of the Stroop effect. Psychol Rev 97:332–361.

Damasio AR, Tranel D, Damasio H (1990) Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. Behav Brain Res 41:81–94.
Dehaene S, Changeux JP (1989) A simple model of prefrontal cortex function in delayed-response tasks. J Cognitive Neurosci 1:244–261.
Dehaene S, Changeux JP, Nadal JP (1987) Neural networks that learn temporal sequences by selection. Proc Natl Acad Sci USA 84:2727–2731.
Drewe EA (1974) The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. Cortex 10: 159–170.
Edelman G (1978) Group selection and phasic reentrant signaling: a theory of higher brain function. In: The mindful brain (Edelman GM and Mountcastle VB, eds), pp 51–100. Cambridge: MIT Press.
Edelman G (1987) Neural Darwinism. New York: Basic.
Edelman G (1989) The remembered present. New York: Basic.
Eken T, Hultborn H, Kiehn O (1989) Possible functions of transmitter-controlled plateau potentials in α motoneurones. In: Progress in brain research, Vol 80 (Allum JHJ and Hulliger M, eds), pp 257–267. New York: Elsevier.
Feldman JA, Ballard DH (1982) Connectionist models and their properties. Cognitive Sci 6:205–254.
Finkel LH, Edelman GM (1987) Population rules for synapses in networks. In: Synaptic function (Edelman GM, Gall WE, and Cowan MW, eds), New York: Wiley.
Fodor JA, Pylyshyn ZW (1988) Connectionism and cognitive architecture: a critical analysis. Cognition 28:3–71.
Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. Neurophysiol 61:331–349.
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 (1989) The prefrontal cortex (2nd ed). New York: Raven.
Fuxe K, Agnati L (1990) Volume transmission in the brain: novel mechanisms for neural transmission. In: Advances in neuroscience, Vol 1, pp 1–602. New York: Raven.
Fuxe K, Agnati L, Zoli M, Bjelke D, Zini I (1989) Some aspects of the communicational and computational organization of the brain. Acta Physiol Scand 135:203–216.
Georgopoulos AP, Crutcher MD, Schwartz AB (1989) Cognitive spatial-motor processes. 3. Motor cortical prediction of movement direction during an instructed delay period. Exp Brain Res 75:183–194.
Goldman-Rakic PS (1984) Modular organization of prefrontal cortex. Trends Neurosci 7:419–424.
Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In: Handbook of physiology, Vol 5. (Plum F and Mountcastle VB, eds), pp 373–417. Bethesda, MD: American Physiological Society.
Goldman-Rakic PS (1988) Topography of cognition: parallel distributed networks in primate association cortex. Annu Rev Neurosci 11:137–156.
Goldman-Rakic PS, Schwartz ML (1982) Interdigitation of contralateral and ipsilateral columnar projections to frontal association cortex. Science 216:755–757.
Grant DA, Berg EA (1948) A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. J Exp Psychol 38: 404–411.
Gustafsson B, Wigström H (1988) Physiological mechanisms underlying long-term potentiation. Trends Neurosci 11:156–162.
Heidmann T, Changeux JP (1982) Un modèle moléculaire de régulation d'efficacité d'une synapse chimique au niveau post-synaptique. C R Acad Sci Paris, Ser C 295:665–670.
Hökfelt T, Holets VR, Staines W, Meister B, Melander T, Schalling M, Schultzberg M, Freedman J, Björklund H, Olson L, Lindk B, Elfvin LG, Lundberg J, Lindgren J, Samuelsson B, Terenius L, Post C, Everitt B, Goldstein M (1986) Coexistence of neuronal messengers: an overview. Prog Brain Res 68:33–70.
Ito M (1989) Long term depression. Annu Rev Neurosci 12:85–102.
Kojima S, Goldman-Rakic PS (1982) Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. Brain Res 248:43–49.
Levine DS, Prueitt PS (1989) Modelling some effects of frontal lobe damage—novelty and perseveration. Neural Networks 2:103–116.
Livingstone M, Hubel D (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240:740–749.
Luria AR (1966) The higher cortical functions in man. New York: Basic.
Marr D (1982) Vision. San Francisco: Freeman.
McCormick DA, Connors BW, Lighthall JW, Prince DA (1985) Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J Neurophysiol 54:782.
McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–137.
Milner B (1963) Effects of brain lesions on card sorting. Arch Neurol 9:90–100.
Mountcastle VB (1978) An organizing principle for cerebral function: the unit module and the distributed system. In: The mindful brain (Edelman GM and Mountcastle VB, eds). Cambridge: MIT Press.
Nelson HE (1976) A modified card sorting test sensitive to frontal lobe defects. Cortex 12:313–324.
Niki H (1974) Differential activity of prefrontal units during right and left delayed response trials. Brain Res 70: 346–349.
Niki H, Watanabe M (1976) Prefrontal unit activity and delayed response: relation to cue location versus direction of response. Brain Res 105:79–88.
Niki H, Watanabe M (1979) Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Res 171:213–224.
Shallice T (1982) Specific impairments of planning. Philos Trans R Soc London Ser B 298:199–209.
Shallice T (1988) From neuropsychology to mental structure. New York: Cambridge UP.
Sternberg S (1966) High-speed scanning in human memory. Science 153:652–654.
Stuss DT, Benson DF (1986) The frontal lobes. New York: Raven.
Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49:93–115.
Traub RD, Miles R, Wong KS (1988) Large scale simulations of the hippocampus. IEEE Eng Med Biol Mag 6: 31–38.
Treisman A, Gelade G (1980) A feature-integration theory of attention. Cognitive Psych 12:97–136.
Tulving E (1972) Episodic and semantic memory. In: Organization of memory (Tulving E and Donaldson W, eds). London: Wiley.
Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Analysis of visual behavior (Ingle DJ, Goodale MA, and Mansfield RJ, eds), pp 549–586. Cambridge: MIT Press.
Zador A, Koch C, Brown TH (1990) Biophysical model of a Hebbian synapse. Proc Natl Acad Sci USA 87:6718–6722.
Zeki SM, Shipp S (1988) The functional logic of cortical connections. Nature 335:311–317.