CHAPTER

59

Thinking and Problem Solving

Stanislas Dehaene, John Jonides, Edward E. Smith, and Manfred Spitzer

The higher mental processes include mental calculation, reasoning, problem solving, and the understanding of language. Virtually all instances of these processes require the use of stored information (see Chapters 55 and 56). Consider two examples, one involving mental arithmetic and the other geometric analogies.

1. Mentally calculate the running total of the following series of numbers: 28, 17, 33, 19, 22.
2. Inspect the 3 × 3 matrix in Fig. 59.1, in which the bottom right entry is missing, and determine which of the 8 alternatives given below the matrix is the missing entry. (To do this, you have to determine the rules that specify how the forms vary across the rows and columns.)

You cannot do the mental arithmetic problem without accessing the relevant arithmetic facts from **long-term memory,** nor can you solve the geometric analogy without retrieving from memory the instructions about choosing among alternatives and analyzing the rows and columns. But more than long-term memory is involved here. In the mental arithmetic example you have to keep accessing the current running total before you can add a new number to it; the constantly changing running total is presumably active, maintained in short-term or **working memory** (in contrast to long-term memory). In the geometric analogy, in each row or column you have to determine the similarities and differences between pairs of items, and then keep these similarities and differences active in working memory (see Chapter 56).

The general point is that thought processes typically generate a number of intermediate mental products that must be held "on-line" for successful performance to occur. This on-line storage system is working memory, and earlier work has established some of its key computational properties (at least for verbal contents like digits, letters, or words). Specifically, working memory:

1. Requires on the order of a few hundred milliseconds to encode a new item.
2. Has a limited storage capacity that has been estimated at 7 ± 2 items.[1]
3. Loses information in a matter of seconds.[2]
4. Requires about 50–100 ms for retrieval of each item.[3]

The system is thus well suited to briefly holding a limited amount of material that must be rapidly accessible, just the kind of system needed to expedite thinking.

The fact that working memory is involved in many forms of thought suggests the following two-step strategy for studying the neural basis of thought:

1. Start by focusing on tasks that require just working memory, and try to characterize the neural circuitry of this system.
2. Then consider tasks that require the involvement of working memory in increasingly complex ways, and use what has been learned about working memory to bootstrap our understanding of the neural basis of these more complex tasks.

In what follows, we adopt this strategy. Specifically:

1. First we consider relatively pure memory tasks and provide an account of a neural network for working memory.
2. Then we move up one step in complexity, looking at tasks in which subjects not only store items in working memory but also code each one with respect to temporal order.
3. Next we consider tasks in which subjects perform a few operations on information stored in work-

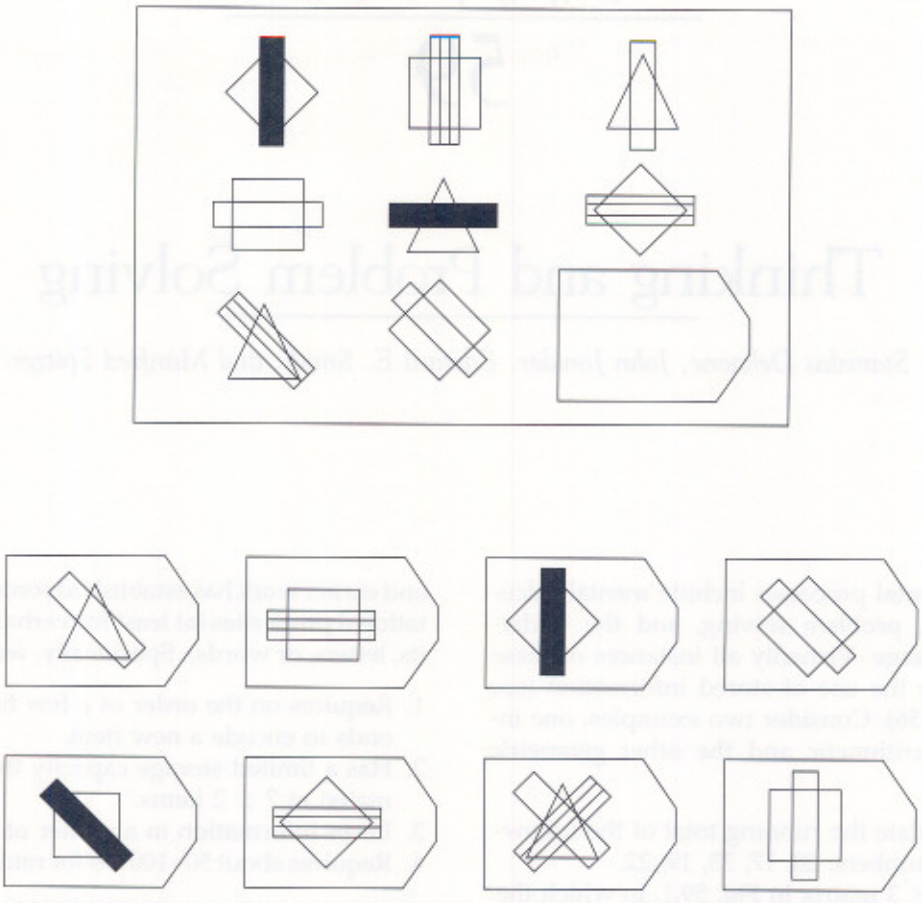

FIGURE 59.1 Example of a geometric analogies problem (from the Ravens Progressive Matrices Test). The task is to determine which of the eight alternatives presented beneath the matrix is the missing bottom-right entry in the matrix. See text for further explanation.

ing memory to accomplish a meaningful goal. Here, we consider two kinds of tasks—language understanding and mental arithmetic.

In discussing the tasks used in steps 2–3 we will be interested in whether the neural basis of performance involves the network for working memory. In addition, we will be interested in what other neural networks are recruited for task performance.

VERBAL WORKING MEMORY

Before beginning our discussion, one more introductory issue must be addressed. We have talked as if there is just one working memory, as if the same storage system is used regardless of the contents that have to be maintained. This assumption is almost certainly incorrect, as shown by single-cell studies of nonhuman primates and neuroimaging studies of humans. Recordings from the dorsolateral prefrontal cortex of adult monkeys[4] showed that some neurons responded only when spatial information had to be stored, and other neurons responded only when visual-object information had to be briefly maintained. In addition, positron emission tomography (PET) studies with humans have found different patterns of activation depending on whether the material stored briefly is spatial, visual object, or verbal in nature.[5,6] Thus, we need to specify *which* working memory we are discussing. In what follows we focus on verbal working memory because so many problem-solving and reasoning tasks have a verbal component.

Verbal Working Memory Retains Items in a Speech-Based Code

Verbal working memory is the system involved when one must briefly remember a series of numbers, letters, words, or other verbal items. For example, after reading the following letters, look away from the text for about 5 s and then report these letters in order:

E B T G V C P. Most people who do this task report that during the 5-s retention interval they implicitly spoke or "rehearsed" the names of the letters to themselves. This introspection is supported by objective behavioral data. For example, the faster one's rate of implicit speech, the better one performs the task (because it is less likely that a letter will have faded before it can be rehearsed again[7]). The upshot of this line of behavioral research is that verbal working memory represents items in a phonological code (the sounds of the items), and these phonological representations can be maintained by a rehearsal process that consists of internal speech.

Rehearsal and Storage Components of Working Memory Have Separate Anatomy

With regard to the neural basis of working memory, consider first some evidence from neuropsychology. Brain damage can result in an impairment known as the **short-term memory syndrome,**[8] in which the primary deficit is an inability to store verbal information for a period of seconds (the deficit is particularly severe for the auditory presentation of material). We can illustrate with one frequently studied patient, KF (see Box 56.3, Chapter 56). When KF is presented with a sequence of 1 to 7 digits and is required to immediately repeat them back in order, he gets only 1 digit correct (normal is 7). In contrast, when KF is given a long-term memory task—say, learning a list of 20 words over a number of trials—he performs normally. Thus, KF's memory deficit is confined to verbal working memory. Importantly, part of KF's brain damage includes the posterior parietal cortex of the left hemisphere. This region is the most frequent site of damage in patients manifesting the short-term memory syndrome.

PET scanning studies with normal subjects provide additional support for the involvement of the left-hemisphere posterior parietal cortex. In one study[9] subjects were scanned while they were performing two different tasks. One task, *item recognition,* is presented schematically in Fig. 59.2, top. On each trial, a target set of four letters is presented briefly (200 ms), followed by a blank retention interval of 3000 ms, followed by presentation of a *probe* letter. The subject's task is to decide as quickly as possible whether the probe names one of the target letters (subjects indicate their decision by pressing one of two response buttons). This is a standard test of verbal working memory,[3] and the pattern of observed activation should reveal the brain structures involved in this system. However, the task also includes components that are not part of working memory—perceiving the letters, selecting a response, and executing a response—and the activation due to these unwanted components needs to be subtracted out if a clear "picture" of working memory is to be gained. This problem is routine in PET research and there is a routine solution: scan the same subjects on a task that is thought to involve the same perceptual and response components as the memory task of interest but not the working memory component, and then subtract the activation pattern obtained in this control task from that obtained in the memory task. The control task used in this study is sketched in Fig. 59.2, bottom. In this task, a set of four letters again is presented on each trial, but now the letters remain in view while the probe is presented so that the subject need not rely on memory in making a decision.

The data obtained from the memory and control tasks consist of sets of images, each image showing the relative changes in blood flow in a particular horizontal slice of cortex. Since increases in blood flow are assumed to be monotonic with increases in neural activity, each brain image reveals which regions have relative increases in neural activity during performance of the task of interest. When the images of the control task are subtracted from those of the memory task, a number of regions are significantly active; presumably these regions mediate verbal working memory. Importantly, one of these regions in the posterior cortex of the left hemisphere is the same region implicated by the neuropsychological studies of the short-term memory syndrome. However, other regions are active as well. They include anterior left-hemisphere regions known to be involved in the production and planning of speech,[10] including Broca's area, the premotor area, and the supplementary motor area (or SMA). Given their role in overt speech, it seems plausible they may mediate covert speech (i.e., rehearsal) as well.

PET experiments in other laboratories provide corroborative evidence for this distinction between posterior storage mechanisms and anterior rehearsal mechanisms. For example, the item-recognition task was associated with activation in the left-hemisphere posterior parietal cortex as well as in anterior speech-related regions.[11] In this same study, when activation in a task requiring rehearsal but not storage was subtracted from the activation associated with the item-recognition task, significant activation remained in the posterior region but not in the anterior regions. This pattern of results gives us a picture of verbal working memory that includes a storage component in the posterior cortex and a rehearsal component in the anterior cortex, both in the left hemisphere.

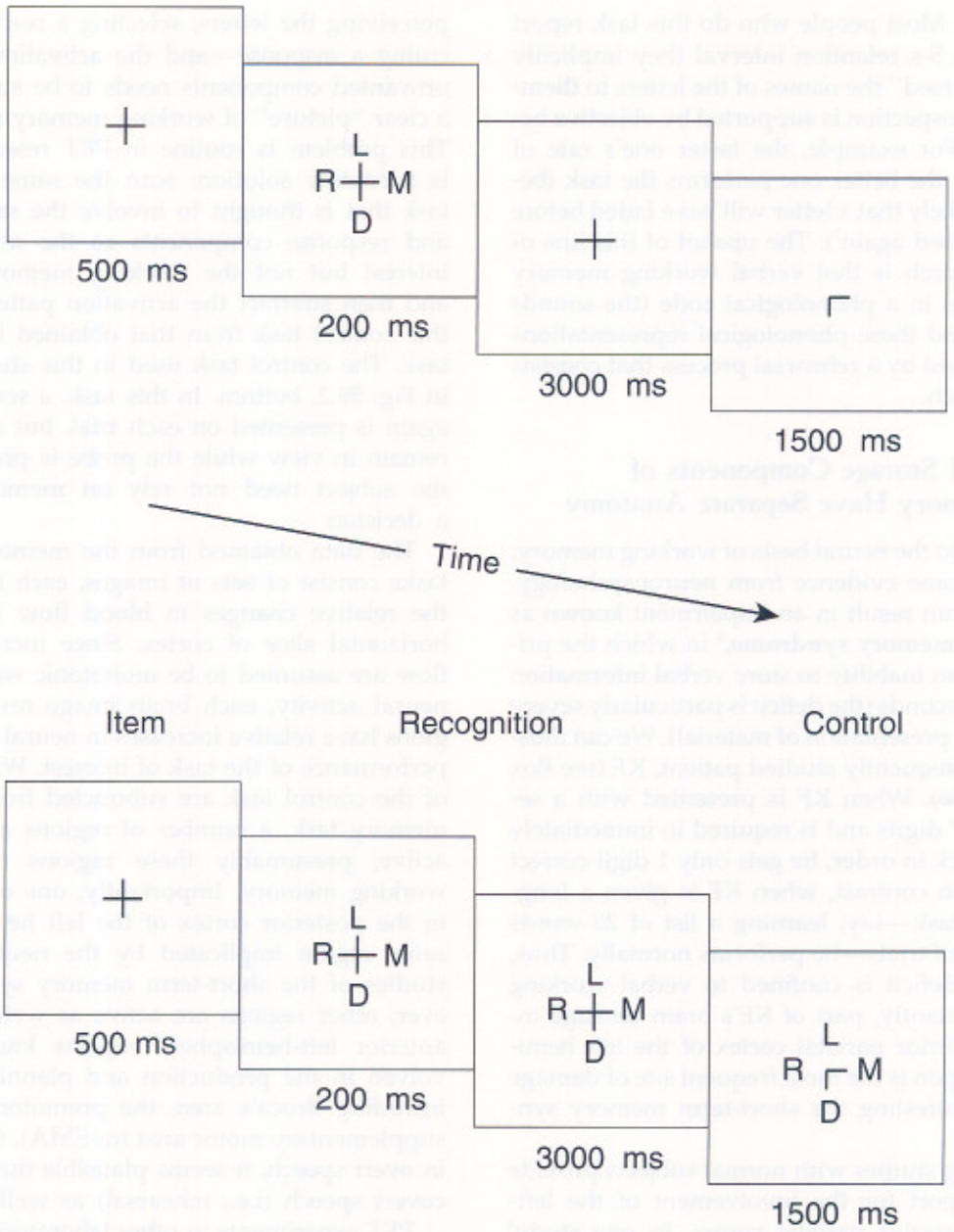

FIGURE 59.2 Schematic representations of the item recognition task (top) and its control condition (bottom).

Working Memory Can Include Temporal Tags to Order Information

Recent neuroimaging research has focused on tasks that required subjects not only to maintain items in working memory but also to code them with respect to their temporal order.[5,9,12,13] Such tasks are of interest because they are further along the continuum from pure memory tasks to tasks that involve working memory in more complex cognition (in accordance with the strategy laid out earlier).

In one such study,[9] subjects saw a continuous sequence of letters, and for each one they had to decide whether it was identical to the one presented two letters earlier or two back. This "two back" task requires a working memory load of two to three items (the last two letters presented plus the current one). It also requires subjects to temporally code the items currently in working memory—because only the letter that entered two back is a proper match—and to continually change these codes as new items enter working memory. The questions of interest are: Do PET images for

this task show evidence of the working-memory neural network described earlier, and do new regions of activation emerge that may correspond to the coding operations needed in this task? The answer to both questions is "yes." When activation under a suitable control condition is subtracted from that in the two back memory condition, the resulting PET images again show activation in the left-hemisphere posterior parietal region that presumably mediates storage, as well as the left-hemisphere anterior regions that presumably mediate inner speech (Broca's area, the premotor area, and SMA).

Importantly, some regions that are not active in pure memory tasks (like item recognition) are active in this memory-plus-coding task. Some of the additional regions are right-hemisphere homologs of left-hemisphere areas already described, including right-hemisphere SMA. These additional regions may reflect mainly the added difficulty of this task compared to the item recognition tasks; the right-hemisphere homologs are helping out in a particularly difficult version of what is normally a left-hemisphere task. Another result, however, suggests that the current task is qualitatively different from a pure memory task. When more sensitive statistical analyses are conducted, an area in the dorsolateral prefrontal cortex (area 46) also shows evidence of activation. This area is known to be involved in the temporal coding of information.[14]

Thinking Includes Working Memory and Executive Attention

So far we have considered working memory in isolation. PET studies have revealed specific brain areas that mediate components of working memory such as storage and active rehearsal.

Let us take a bigger step up the complexity scale and consider tasks in which subjects perform some operations on the information stored in working memory so as to accomplish a meaningful goal. This involves not only the problem of representing information in memory that has been considered in this chapter but also the issue of executive control discussed in Chapter 54. Two kinds of tasks will be considered: certain situations requiring an understanding of language and standard mental arithmetic problems. Both kinds of tasks clearly exemplify thinking.

Working Memory Mediates Communication

The best behavioral evidence for the involvement of working memory in language understanding comes from correlational studies.[15,16] There is a positive correlation between a measure of an individual's working-memory capacity and that individual's performance in a language-understanding task, such as answering true–false questions about previously read paragraphs. (Such correlations, though, are obtained with measures of working memory that require the subject to store material while concurrently engaging in some processing, for example, storing the last word of each of a series of sentences that the subject also has to understand.)

Some neuropsychological evidence for the involvement of working memory in language understanding comes from patients with the short-term memory syndrome (impairment in short-term memory tasks associated with damage in left posterior parietal cortex). Although such patients have no difficulty understanding most kinds of sentences, they perform poorly when they also have to carry out a mental operation based on the verbatim content of the sentence.[17] Presumably, storing the verbatim content—the exact words—requires verbal working memory, and this is why the patients' capacity to understand sentences breaks down. Thus, the patients are impaired in answering orally presented comparative questions like: "Which is green, a poppy or lettuce?" Successful performance here depends on maintaining the exact words ("green," "poppy," and "lettuce") and performing mental operations on the representations of these words.

Another kind of sentence on which the patients show impairment is an orally presented instruction with high information content.[18] An example is "Before picking up the green circle, touch the red square," where the patient is expected to carry out the instruction. In this case, understanding requires setting up a plan, and the generation of this plan requires verbatim memory of some of the information; the latter requirement presumably causes the patients' problems. Other sentence types that people with short-term memory syndrome have difficulty with show the same general characteristics.[17]

Given that we are dealing with language understanding, regions other than the left posterior parietal cortex—namely, the language regions in the perisylvian area of the left posterior temporal cortex (including Wernicke's area)—are also involved. These are the regions that are damaged in many patients who show language disturbances (e.g., Wernicke's aphasia). The left-hemisphere posterior temporal cortex apparently houses our normal language understanding system. However, when verbatim memory of language is required, the verbal working memory system in left-hemisphere parietal and frontal cortices is called into play as well.

Arithmetic Also Requires Working Memory

Extensive behavioral evidence indicates that verbal working memory is used in mental arithmetic. For instance, the longer partial sums must be held before they can be reported, the poorer mental-calculation performance will be, which suggests that the outputs of mental calculation undergo decay just as any other information in working memory. Thus, when subjects must mentally add a 3-digit and a 2-digit number, their accuracy is greater when they can report their answers in reverse order (units, tens, hundreds) than in forward order.[19] Another piece of behavioral evidence for the link between working memory and calculation comes from studies on the cognitive effects of aging. It is well known that working memory declines with age.[20] It has also been established that the ability to mentally execute a multistep numerical calculation declines with age. Neither of these two findings is surprising. What is newsworthy is that the age-based decline in mental calculation is almost totally attributable to those steps of the calculation that require storage in working memory.[21]

Turning to the neural underpinnings of mental calculation, we find that patients with the short-term memory syndrome also show impairment on arithmetic tests.[17] In addition, early neuroimaging work[22] showed that when subjects engaged in a mental arithmetic task (subtracting successive 3s from 50), the left-hemisphere parietal cortex was activated. These same imaging results also showed extensive activation in the frontal cortex, including the dorsolateral prefrontal cortex. More recent neuroimaging work has confirmed that mental calculations recruit both inferior parietal regions associated with number processing and dorsolateral prefrontal ones associated with working memory.[23] Again, working memory seems to be an important component in a thinking task, but it is only one component of the total neural system recruited for the task.

Summary

Neuroimaging tasks that look at verbal working memory show activation of left-hemisphere frontal and parietal sites. The frontal area is closely related to rehearsal of the items in working memory, whereas the posterior area is related to storing items. Things can be viewed as a combination of these working memory areas with high-level executive attention networks that were reviewed in Chapter 54.

MODELS OF PROBLEM SOLVING

Nearly all the areas of the brain can be said to be involved in some form of problem solving. For instance, some visual areas solve the problem of recovering the 3D shape of objects from their retinal projection. As already noticed by Helmholtz at the end of the 19th century, such "perceptual problems" can be extremely difficult and ambiguous and require a sophisticated apparatus, adequately called by Helmholtz an "unconscious inference," in order to solve them. However, the kind of problem solving that is performed in perception (and also in simple motor control) has the characteristic of being highly inflexible. No matter how long we look at many visual illusions, our percepts do not change. This is because our perceptual apparatus is dedicated to solving a single restricted problem, and almost always solves it in the same way.

The kind of problem solving that we consider in this chapter, however, is quite different and is characterized by a considerable amount of flexibility. Mammals (especially the higher primates and, of course, humans) can find a solution to problems that they have never encountered before and for which evolution has not developed dedicated "wired-in" solutions. For instance, a rat can find the way out of a maze, a chimpanzee can figure out how to use a stick to unhook a banana from the ceiling, and a human adult can subtract 356 from 644, play chess, or plan a trip to Mexico. Such activities involve (a) constructing an accurate mental representation of the goal to be achieved, (b) selecting appropriate means for achieving this goal, and (c) executing the planned strategy and monitoring how successful it is.

Obviously, models of problem solving can be aimed at several different levels of analysis. Mathematics (especially graph theory and game theory), computer science, and artificial intelligence have mostly been concerned with designing and evaluating different computational strategies for solving problems. Although some insights have been gained through this approach, it will only be minimally discussed here because, for the most part and until recently, it has shown little concern for the actual solutions that humans and animals use or for the neural apparatus underlying them. We shall deal more deeply with models, originating from cognitive psychology and/or neuropsychology, that have examined the mental architectures underlying simple reasoning. Finally, we shall discuss recent models that have made specific proposals about the neuronal or even the molecular systems underlying flexible problem solving.

Artificial Intelligence Studies Aid Us in Describing Problem Domains

From its early inception, artificial intelligence has been trying to copy the flexibility that humans, and

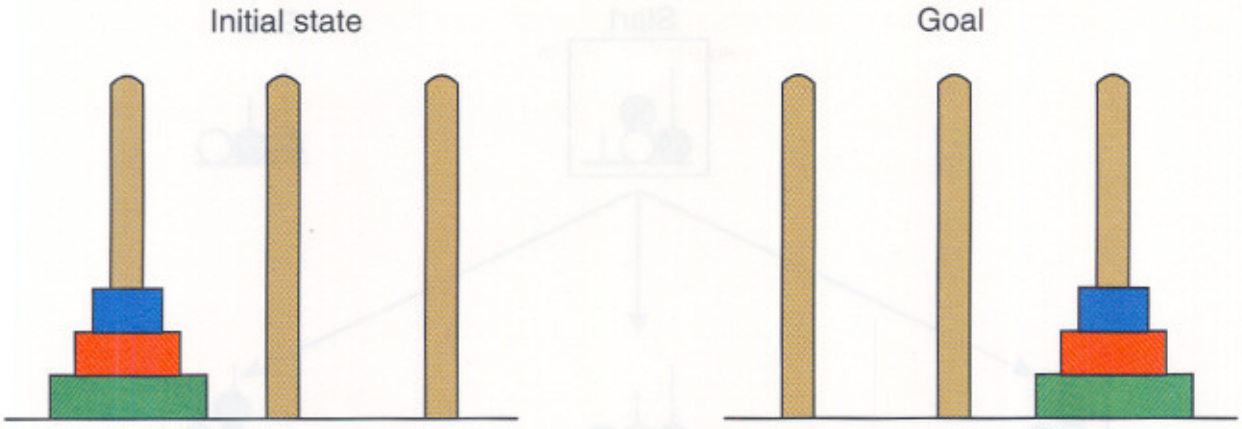

FIGURE 59.3 Example of a simple Tower of Hanoi problem. The task is to move the stack of disks over to the right hand peg, moving one at a time, and being sure not to stack a larger disk on a smaller one.

to some extent animals, show in abstract problem solving as exemplified by chess playing or theorem proving. One problem task that has been widely studied is the Tower of Hanoi, illustrated in Fig. 59.3. In this task an optimal solution can be specified, and yet the task captures some of the decision-making strategies involved in common games such as chess.

An important insight from studies of the Tower of Hanoi and of chess is that such problems can be abstractly represented by **decision trees:** at each point in time, there are many options for action, some of them leading closer to the goal and most of the others diverging from the solution (Fig. 59.4). Thus, solving a problem becomes equivalent to exploring a large tree of possibilities and trying to find the shortest path toward the goal.

In the most interesting cases, the tree of possibilities is so large that it cannot be explored, or even known, in its entirety. Hence, **heuristics** or rules of thumb must be devised to guide the search. One possibility is to reason from the goal backward, defining a chain of achievable subgoals that become progressively easier to reach from the current starting situation. Another useful device is an evaluation function that, for each given situation, computes an approximation of how remote the goal is. If the evaluation function is adequate, simply picking, whenever a decision needs to be made, the action that leads to the most valuable situation will be a successful strategy.

Another useful problem-solving trick in artificial intelligence (AI) is using learning to progressively reduce and focus the search. One of the earlier successes of AI, Samuel's checker-playing program, progressively adapted its evaluation function so that, in the end, the evaluation of a very large tree of possibilities could be predicted by a single "look" at the checker deck. Similarly, frequently used action sequences may be compiled into a more or less fixed scenario or script.[24] Upon later encounteres with a similar situation, the script can be reused. The role of the controller program is then confined to checking the execution of the script and reacting appropriately if it becomes inadapted or ends prematurely. In the course of solving more and more problems, some AI programs thus learn to compile a repertoire of strategies for solving different problems. When a relatively novel problem-solving situation occurs, and when the known strategies seem to fail, such a program still needs to resort to tree exploration heuristics as above. However, it can automatically detect recurring regularities in its exploration behavior and add them to its list of available strategies. Progressively, then, problem solving comes to rely more and more on precompiled strategies, with calls to a higher-level controller becoming less and less frequent.

Although most AI approaches are aimed at developing efficient programs rather than at simulating human problem-solving behavior, it is remarkable that in so doing, they have developed concepts that often provide adequate characterizations of the actual mental architectures for abstract thought. There is considerable evidence for a dissociation, in humans, between "routine" and "creative" problem solving. The greatest chess players, for instance, have developed highly efficient routines for "parsing" any chess situation and immediately getting excellent hunches as to what the best moves may be. Before becoming experts, however, we all have to painfully explore a tree of possibilities that taxes our working memory, error detection, and backtracking abilities. Areas of the prefrontal cortex are instrumental in the creative form of reasoning and problem solving.

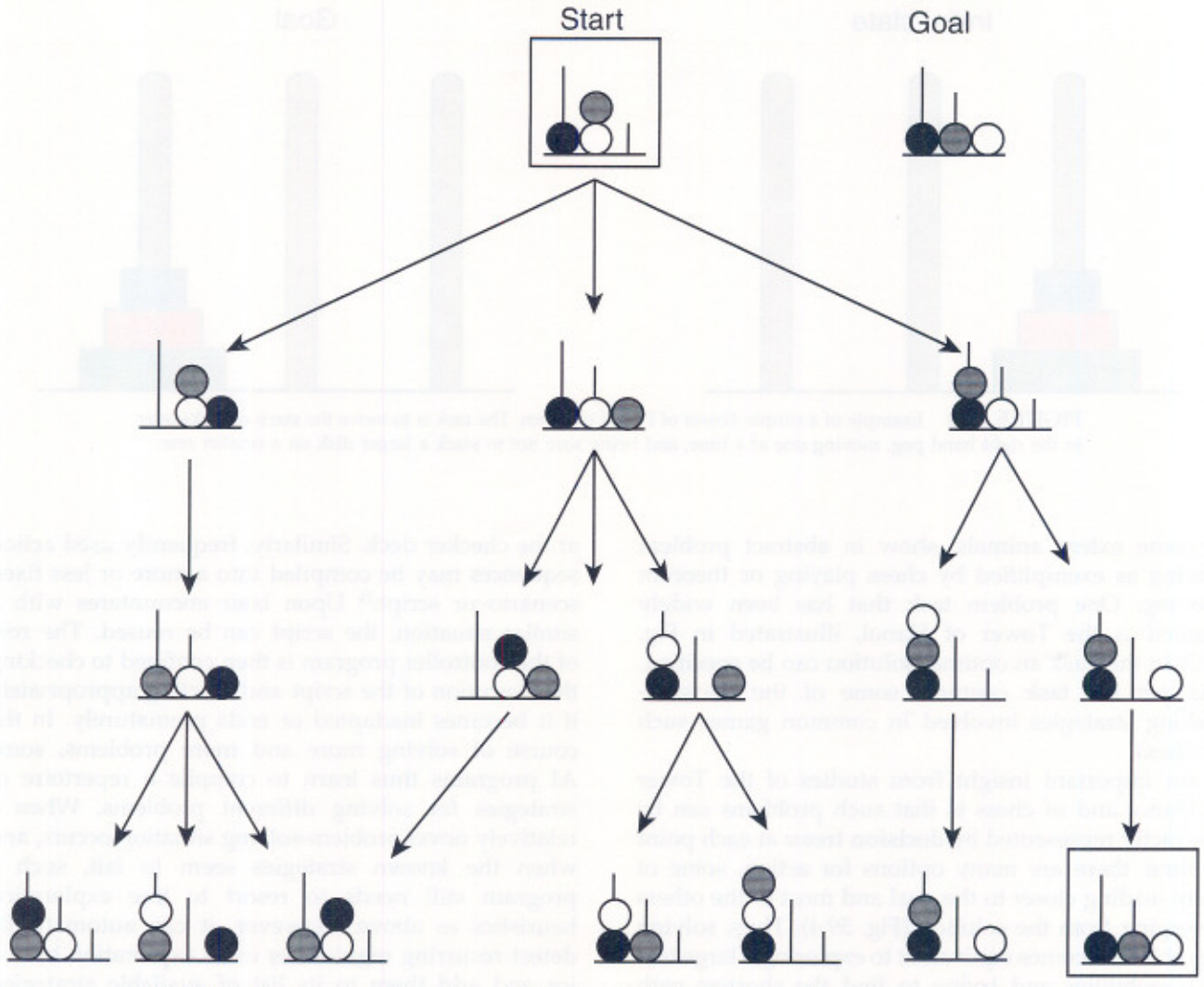

FIGURE 59.4 A small portion of the decision tree for Shallice's Tower of London test, which requires planning a series of moves of balls on three pegs[6].

There Is a Common Human Architecture for Problem Solving

One of the more influential models of the architecture of human problem solving was developed by Norman and Shallice (see also Chapter 54 for a review of this model). The Norman–Shallice model belongs to cognitive neuropsychology. Although it nicely captures the behavioral data on problem solving in normal human adults and in patients with damage to the frontal lobe, it provides no means of predicting the neuronal circuits and the activity of cells involved in problem-solving tasks. However, as described in Chapter 54, tasks involving conflict usually activate frontal areas. A number of neuronal network models have been proposed on the basis of these findings.

The Prefrontal Cortex May Play a Role in the Flexible Adaptation of an Organism to Various Tasks

A classical set of task tapping prefrontal functions in animals is the delayed-response task. Depending on the version of the task, the animal must pay attention to and store in working memory different aspects of its environment, such as the identity of an object (in delayed matching-to-sample tasks) or its location (in spatial delayed-response tasks). Following a delay, the animal is given a choice between two objects and is asked to find the one that matches the description stored in memory. Dehaene and Changeux[25,26] have proposed a simple model of how the prefrontal cortex contributes to such delayed-response tasks. The model

is based on a hierarchy of neural layers, each of which modulates the level immediately lower to it (Fig. 59.5). When the object to be memorized is presented for input, a low-level direct mapping between sensory data and corresponding actions allows it to be grasped. This lower level input–output system, which corresponds to the triggering of action schemata in the Norman–Shallice model, has no memory, however, so it cannot support delayed-response performance. A higher level mapping includes working memory units with long-lasting firing properties that can modulate and select among actions triggered at the lower level. This indirect mapping coarsely maps onto the dorsolateral prefrontal cortex, and the activity of the simulated units mimics the long-lasting, delayed related firing of actual prefrontal cells.

In the Dehaene–Changeux model, behavioral flexibility is achieved because the nature of the task to be performed (spatial delayed-response or delayed matching-to-sample) is encoded in cell activities rather than in connection strengths and can therefore be modified "on the fly" as required, without requiring slow modifications of synaptic strengths. As noted, some units in the model code for aspects of past experience stored in working memory. The activity of other units, called "rule-coding units," encodes the behavioral set of plans that the organism is currently following. The network might change, for instance, from performing spatial delayed response to performing delayed matching-to-sample by the mere switching on or off of some rule-coding units.

In human adults, a classical task tapping the ability to switch between mental sets is the Wisconsin Card Sorting Test. A deck of cards bearing colored symbols must be sorted, but at different stages of the test the sorting criterion changes abruptly from, say, color to shape or to number of stimuli on the cards. Patients with lesioned prefrontal cortext fail to change their sorting criterion and perseverate sorting by color. Such perseveration behavior can be mimicked by the Dehaene–Changeux network (Fig. 59.5B).

A group of units (color rule-coding units), when

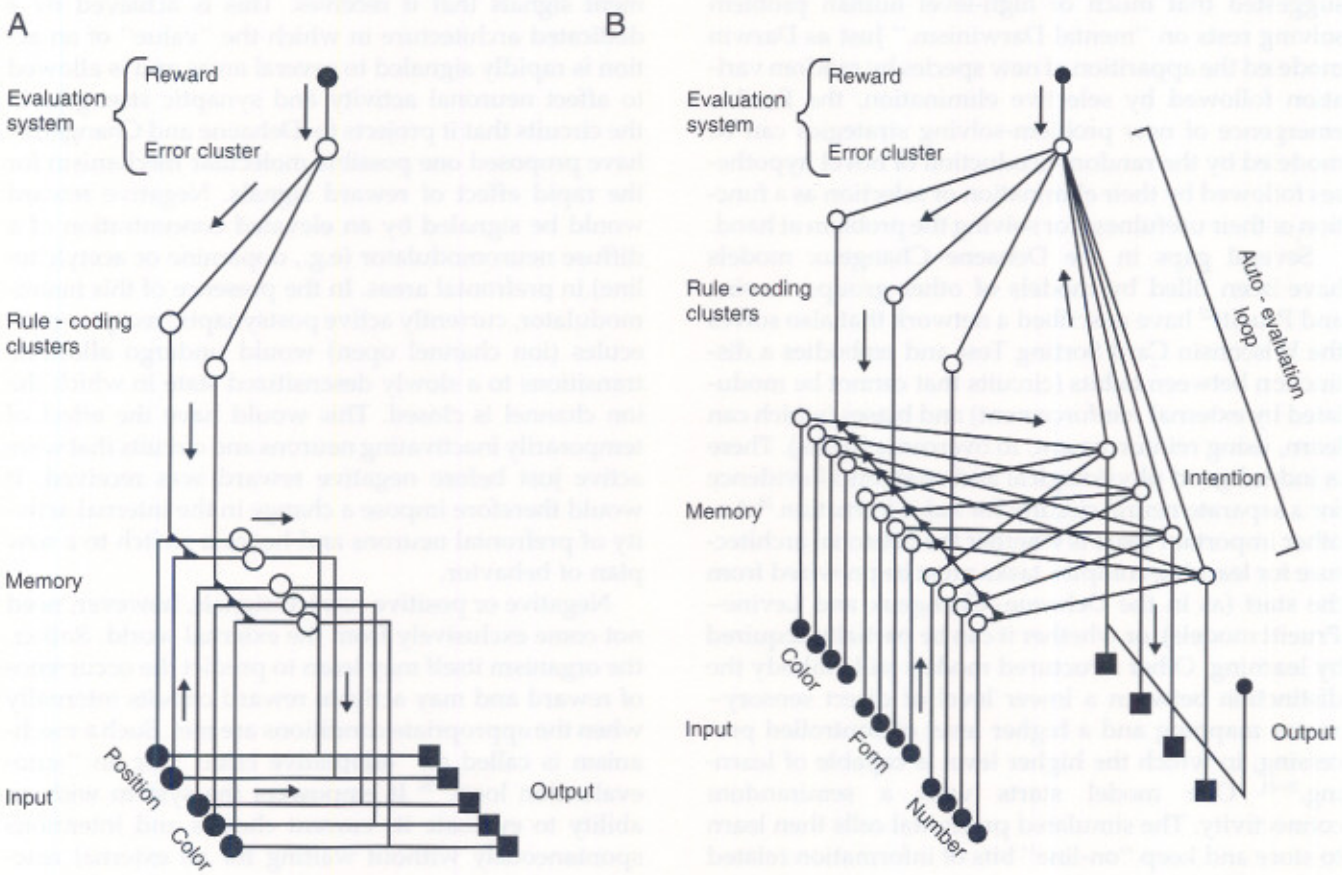

FIGURE 59.5 Architecture of two simple model neuronal networks that have been used to simulate functions of the prefrontal cortex. Left: Delayed-response tasks. Right: Wisconsin Card Sorting Task. Redrawn from Dehaene and Changeux.[26]

active, directs the lower levels of the system to orient to color and therefore to sort the stimulus cards according to their color. The overall sorting plan can be rapidly changed if needed, for instance, if negative external reward signals that the current rule is likely to be invalid. It is assumed that the different rule-coding units are organized in a specialized neuronal network that functions as a generator of diversity. That is, units coding for different rules have a strong level of spontaneous firing that enables them to spontaneously come into play and propose a new hypothetical rule or plan to be followed. Normally, the units inhibit each other so that only one can be active at a given time. Negative reward, however, has the effect of destabilizing the currently active unit and therefore of releasing all the others from inhibition. Spontaneous firing then ensures that new rules will be tried out until a fitting one is found. The model was shown to be able to pass tests of cognitive flexibility such as the Wisconsin Card Sorting Test, and impairments similar to those found in patients with damage to the frontal lobes were observed when parts of the control structure were lesioned. More generally, Dehaene and Changeux[25] have suggested that much of high-level human problem solving rests on "mental Darwinism." Just as Darwin modeled the apparition of new species by random variation followed by selective elimination, the flexible emergence of new problem-solving strategies can be modeled by the random production of novel hypotheses followed by their elimination or selection as a function of their usefulness for solving the problem at hand.

Several gaps in the Dehaene–Changeux models have been filled by models of other groups. Levine and Prueitt[27] have described a network that also solves the Wisconsin Card Sorting Test and embodies a distinction between habits (circuits that cannot be modulated by external reinforcement) and biases (which can learn, using reinforcement, to overcome habits). There is indeed good physiological and anatomical evidence for a separate neural circuity for habit formation.[28] Another important issue is whether the neuronal architecture for learning complex tasks must be prewired from the start (as in the Dehaene–Changeux and Levine–Prueitt models), or whether it can be partially acquired by learning. Other structured models still embody the distinction between a lower level of direct sensory–motor mapping and a higher level of controlled processing, in which the higher level is capable of learning.[29–31] One model starts with a semirandom connectivity. The simulated prefrontal cells then learn to store and keep "on-line" bits of information related to the task at hand.

It is important to note that most current models either are prewired to solve a specific task or require thousands of learning trials before they can master a novel task. No model to date has shown anything close to the specifically human capacity for switching tasks by mere exposure to a few seconds of instructions. The most flexible of current models can only switch between a small number of local circuits. To put such "flexibility" into perspective, brain-imaging techniques in humans have shown that the pattern of regional activation of the entire brain can be radically altered by attention, intention, and instruction. Much more research is needed to understand how the selection of active circuits is controlled in the human brain.

Emotions Are Involved in Models of Reasoning

An interesting feature of current neuronal models of problem solving and decision making is that they have led to a renewed interest in reward, evaluation, and emotion systems and to a reconsideration of their role in reasoning. In most of the above models, flexibility comes from the prompt behavioral reaction of the system to a change in the positive or negative reinforcement signals that it receives. This is achieved by a dedicated architecture in which the "value" of an action is rapidly signaled to several areas and is allowed to affect neuronal activity and synaptic strengths in the circuits that it projects to. Dehaene and Changeux[26] have proposed one possible molecular mechanism for the rapid effect of reward signals. Negative reward would be signaled by an elevated concentration of a diffuse neuromodulator (e.g., dopamine or acetylcholine) in prefrontal areas. In the presence of this neuromodulator, currently active postsynaptic receptor molecules (ion channel open) would undergo allosteric transitions to a slowly desensitized state in which the ion channel is closed. This would have the effect of temporarily inactivating neurons and circuits that were active just before negative reward was received. It would therefore impose a change in the internal activity of prefrontal neurons and hence a switch to a new plan of behavior.

Negative or positive reward signals, however, need not come exclusively from the external world. Rather, the organism itself may learn to predict the occurrence of reward and may activate reward circuits internally when the appropriate conditions are met. Such a mechanism is called an "adaptative critic"[32] or an "auto-evaluation loop."[26] It empowers the system with an ability to evaluate its current choices and intentions spontaneously without waiting for an external reaction, and in some cases, without even actually performing actions. Reasoning, in effect, becomes a mental simulation of possible actions and an internal evalua-

tion of their predicted consequences. Unlike a more basic system, a network possessing an auto-evaluation loop can use reasoning to avoid making the same error twice in the Wisconsin Card Sorting Test. When a new rule is selected, the network first tries it internally to see if it yields the same error as on the previous trial. If it does, negative reward is predicted and triggers the elimination of this rule, as would a normal external reward.

Summary

Artificial intelligence models problem solving as the exploration of a vast tree of possibilities under the guidance of various heuristics and evaluation functions. Prefrontal cortex is heavily involved in this active exploration. Neuronal models of planning and problem solving suggest at least three functions for prefrontal cortex areas. Some prefrontal circuits are involved in the active maintenance of problem information in working memory. Other circuits may be involved in maintaining specific mental strategies and in rapidly switching from one strategy to another. Evaluation circuits, finally, may contribute to the internal monitoring of ongoing plans. Evaluation circuits may cause the release of reward signals, for instance, from catecholaminergic projection systems, which may modulate the activity of other units and thus influence decision making.

DISORDERS OF THOUGHT IN SCHIZOPHRENIA

Psychopathology is the science and the art of describing what is wrong with a mentally ill patient. Schizophrenia (see Box 59.1),[33] which is arguably the most disruptive and at the same time enigmatic of the psychiatric disorders, has a lifetime prevalence of 1 to 1.5% and about 1 person in 3000 is treated for schizophrenia in any one year in the United States. Most patients become ill in their twenties, although there are rare cases of childhood onset and onset past the age of 55.

Schizophrenia Is Associated with Impaired Access to the Mental Lexicon and Working Memory and by Disordered Thought Content

The concept of schizophrenia was formed 90 years ago. At that time, disrupted thought processes were described in terms of association psychology, the prevailing school of thought in psychology. Currently, these changed processes of thinking are considered in a framework linked to the basic cognitive functions of working memory and thought discussed previously.

Schizophrenic thought disorder can be characterized by a *decreased accuracy of lexical access* combined with *decreased working memory*. Both dysfunctions may be related, and they may be caused by dysfunctional maplike semantic networks that can be localized in frontal and temporal cortical areas. Disorders of the content of thought (i.e., delusions) are discussed within a framework of **neuromodulation** and **neuroplasticity** in cortical networks processing semantic and possibly other high-level information. The structures of semantic networks as revealed by experimental psychological, neuropsychological, and functional magnetic resonance imaging (fMRI) studies bear a close resemblance to self-organizing feature maps (i.e., a type of neural network). In these maps, the influence of noise on plasticity can be demonstrated, and these findings can be related to the neuromodulatory function of dopamine, which appears to regulate the signal-to-noise ratio in network information processing. In sum, different methods and strategies provided by cognitive neuroscience are combined to bridge the gap between psychopathology and underlying brain pathology and to provide a comprehensive and parsimonious explanation of a number of otherwise inexplicable or unrelated phenomena.

The Mental Lexicon Is a Maplike Network

The store of words in the mind is called the **mental lexicon.** Words have a number of features regarding meaning, grammar, sound, writing, and aspects of use. Word associations and various ways to study them have played a major role in research on how the mental lexicon is organized—that is, how semantic, grammatical, phonetic, graphemic, and pragmatic information is stored and accessed when language is produced or understood.[34–36]

One of the more important methods for studying word-related computations is the **lexical decision experiment.** The subject must decide whether or not a given string of characters is a word. To investigate specific types of associations, two words are presented one at a time, and the effect of the relation between the words on the reaction time is measured. Figure 59.6 shows the series of events in a typical lexical decision experiment. A robust phenomenon that has been discovered using this technique of lexical decision is **semantic priming.**[37,38] A word is recognized faster if a meaningfully related word is shown immediately before. For example, "black" is recognized faster as a word if it is presented shortly after "white" than if it is presented shortly after a nonrelated word such as "soft."

BOX 59.1

SCHIZOPHRENIA

The Swiss psychiatrist Eugen Bleuler coined the term "schizophrenia" in a famous book, first published in 1911.[33] In this book, which is full of clinical descriptions, Bleuler made the point that this disorder is characterized by four types of symptoms regarding:

(1) thinking (disordered associative processes),
(2) affect (inappropriate, depressed, or manic affect),
(3) will (ambivalence, i.e., indecisiveness), and
(4) social behavior (autism, social withdrawal).

The following writing of a patient examplifies the disturbed thought processes in schizophrenic patients, driven by associative links between concepts rather than goal-directed thinking.

> I am writing on paper. The pen which I am using is from a factory called 'Perry & Co.' This factory is in England. I assume this. Behind the name of Perry & Co. the city of London is inscribed; but not the city. The city of London is in England. I know this from my school-days. Then, I always liked geography. My last teacher in that subject was Professor August A. He was a man with black eyes. I also like black eyes. There are also blue and gray eyes and other sorts, too. I have heard it said that snakes have green eyes. All people have eyes. There are some, too, who are blind. These blind people are led about by a boy. It must be very terrible not to be able to see. There are people who can't see and, in addition, can't hear. I know some who hear too much. One can hear too much. There are many sick people in Burgholzli; they are called patients. One of them I like a great deal. His name is E. Sch. He taught me that in Burgholzli there are many kinds, patients, inmates, attendants. Then there are some who are not here at all. They are all peculiar people. . . . (Bleuler 1911/1950, p. 80).[33]

Schizophrenia poses a major public health problem worldwide, with a lifetime prevalence of about 1% in most cultures and geographic locations studied. Its clinical manifestations usually appear in adolescence and early adulthood. Prognosis varies widely, with some patients showing a stable illness course with minimal disability during maintenance treatment and others a more severe, deteriorating course which eventually stabilizes at a level of marked disability.

While the etiology of schizophrenia is unknown, several lines of converging evidence suggest that schizophrenia is a neurodevelopmental disorder (or group of disorders) resulting from a combination of genetic susceptibility and acquired neuropathology arising early in life. Family, twin, and adoption studies have established the importance of genetic factors in the etiology of schizophrenia, but have also suggested a role for acquired factors.[11,12,14] For example, the concordance rates among monozygotic and dizygotic twins are about 50 and 20%, respectively.

The most compelling evidence for a neurodevelopmental process in schizophrenia comes from *postmortem* studies that show cytoarchitectural disorganization of the cerebral cortex, which if confirmed by future studies would imply a defect in cortical development arising during the second trimester of gestation. Recent neurodevelopmental models proposed for approaching the pathogenesis of schizophrenia attempt to account for these histopathological findings and the several-year latency between the developmental period and the onset of clinical symptoms. For example, some groups have shown that lesions placed in specific frontal cortical or hippocampal areas during the perinatal period in experimental animals result in cognitive impairment and hyperdopaminergic behaviors (see below) which are not pronounced until after puberty. Another model capitalizes on the observation that NMDA–glutamatergic receptor antagonists (e.g., phencyclidine) can produce the spectrum of psychotic, behavioral/social, and cognitive symptoms seen in schizophrenia in nonschizophrenic, postpubertal humans. In contrast, such reactions rarely if ever occur in children who receive NMDA antagonists as part of anesthetic regimens, suggesting that the development of sensitivity to the psychotomimetic effects of NMDA hypofunction has an age dependency similar to that seen for the onset of psychosis in schizophrenia. Moreover, when introduced *in utero* to rats at critical developmental stages, NMDA antagonists produce neurodegenerative changes in the limbic cortex. These observations led to the hypothesis that schizophrenia is associated with NMDA receptor hypofunction, which, during cortical development, produces excitotoxic damage and consequent microscopic abnormalities, and during the postpubertal period, produces susceptibility to psychosis.

Wayne C. Drevets

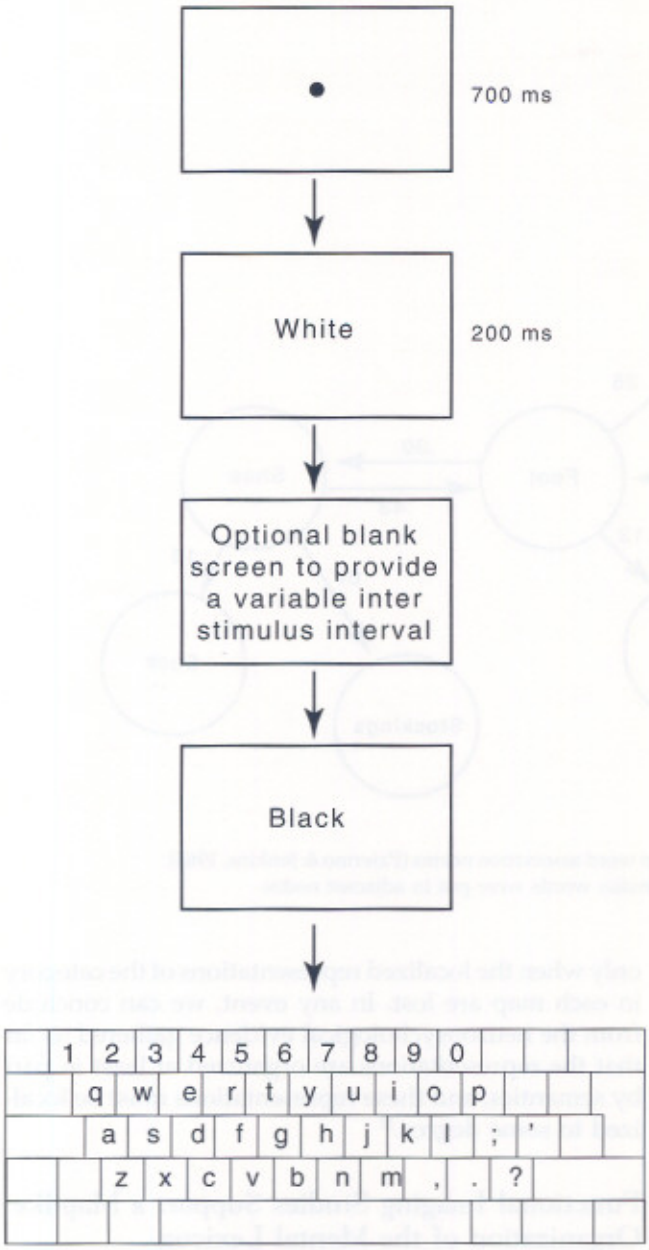

FIGURE 59.6 A typical sequence of events in semantic priming. Time runs vertically from top to bottom. The prime (e.g., white) stays on for 200 ms and the target (e.g., black) until the subjects responds.

Lexical Decision Experiments and Related Techniques Suggest That Word-Related Information Is Stored in the Form of an Associative Network

According to current network models of the mental lexicon, semantic (and possibly other) features of words are represented as "nodes" (which have been called "logogens") in a neuronal network (see Fig. 59.7). In the course of an utterance, these semantic units become activated for a short period of time and thereafter either decay rapidly or are actively inhibited.[38–40] This model of access to the mental lexicon further asserts that concepts activated in a semantic network by a prime serve as a source of activation that spreads to other related concepts. Such spreading of activation to nearby nodes in the semantic network lowers their thresholds of activation, that is, increases their probability of becoming used in the production or understanding of a subsequent utterance. If one of these concepts is denoted by a word that is a target in a lexical decision experiment, this target will be recognized faster because its processing is facilitated by its being already activated to some degree.

Ritter and Kohonen[41] have proposed a biologically and computationally plausible mechanism for the formation of semantic networks. They used a special type of neural network, a so-called self-organizing feature map, to simulate the organization of semantic input. This type of network is highly biologically plausible, since its basic features—lateral inhibition, a high degree of connectivity, and Hebbian learning—are features of the neuronal organization of the neocortex.[42,43] When the neural network is presented with any kind of coherent input, it will create an orderly, maplike representation of this input.[44,45] Ritter and Kohonen presented such a network with the names and characteristics of animals. This was done by using an arbitrary binary code for the animals' names and a binary representation of the presence or absence of the animals' crucial features. Upon presentation of the input, the network formed a map, on which 16 animals were represented in such a way that animals with similar features were close together and dissimilar animals were far apart. In their second experiment, the authors presented short sentences to a similar network. This time, the network organized this input not only according to semantic but also according to grammatical features of the words. Nouns, adverbs, and verbs were put on distinctive areas on the map, and within these areas, the words were organized according to semantic features.

The Brain Maps Semantic Concepts

Very different aspects of the outside world are represented in a maplike manner in the human cortex. Multiple somatosensory and motor maps were described even several decades ago.[46,47] Moreover, multiple retinotopic and tonotopic maps have been discovered in the primate cortex, and evidence indicates that

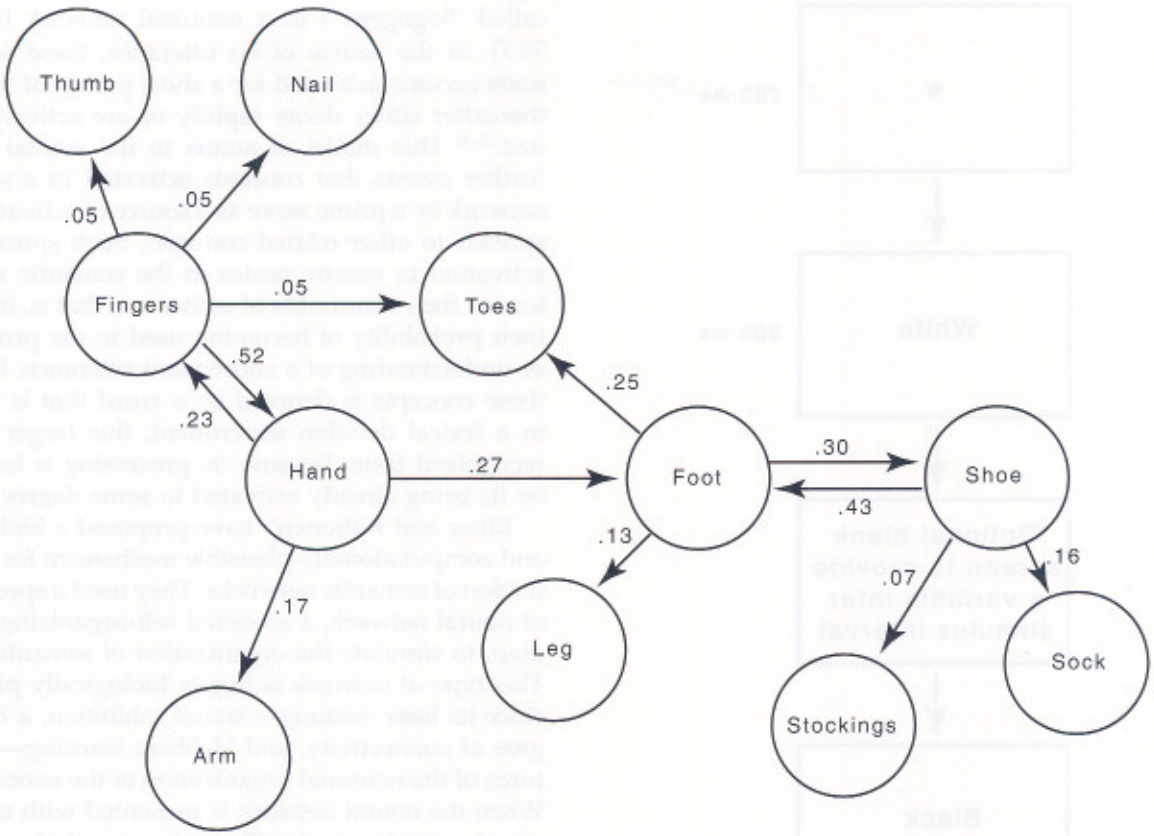

FIGURE 59.7 Semantic network (sketch) derived from word association norms (Palermo & Jenkins, 1963). Words that normal college students associated with stimulus words were put in adjacent nodes.

such maps exist in the human cortex[48–50] (see Chapter 50).

Data accumulated over the past decade from brain damaged-patients point to the existence of *semantic* maps in cortical areas. Some patients with aphasia related to brain damage display a loss of only a small fraction of their semantic memory. Some patients have no cognitive deficit except for the naming of, for example, living things. Others are unable to name vegetables or items inside the house.[18,51,52] These cases appear to be rare, but an increasing number of descriptions of patients with category-specific naming deficits suggests that a negative observation bias may have contributed, and may still be contributing, to the rarity of the phenomenon (see Chapter 57). In fact, as early as 1966, a high incidence of deficits in naming items of a specific category was found in a quantitative study of category-specific word comprehension deficits. This finding led to the conclusion that in aphasic patients such category-specific deficits may be the rule rather than the exception. The fact that the lesions in most of these cases were not small, but rather extended or even diffuse, suggests the existence of multiple semantic maps, so that selective category-specific deficits occur only when the localized representations of the category in each map are lost. In any event, we can conclude from the neuropsychological evidence gathered so far that the representations are organized at least in part by semantics, and these representations must be localized to some degree.[52]

Functional Imaging Studies Support a Maplike Organization of the Mental Lexicon

In one fMRI study, normal subjects were asked to covertly name pictures of items from four categories that had been chosen on the basis of previous neuropsychological evidence (i.e., animals, furniture, fruit, and tools). Color images of 20 items from each category were digitized and processed so that they were about equal in luminance and contrast and could be projected onto a screen mounted within the fMRI scanner. The stimuli were projected one at a time for 1.5 s each, resulting in 30-s stimulation epochs per category. One 6-min data acquisition run consisted of three presentations of each series of category-specific items, that is, of 12 epochs, in fixed succession (animals / furniture / fruit / tools / animals / furniture, etc.). Four such runs

were executed by each subject, with a pause of about 15 min between runs. Each run was preceded by a 30-s "warm-up" phase in which 20 items from the four categories (5 of each) had to be named. The purpose of this warm-up was to exclude activation caused by the mere beginning of the task, a phenomenon that was seen in a pilot experiment. The stimuli were different from those in the subsequent tasks. While the subjects performed the task, they were scanned with a 1.5-T (Tesla) MR scanner equipped with echo planar imaging and a surface coil placed over the left frontotemporal region of the head. During each 30-s epoch, 15 images were taken in 5 planes.

Data were analyzed on a unit volume basis (i.e., voxel by voxel). Figure 59.8 displays the time course of activation for the furniture and tool categories. Most subjects showed areas of increased cortical activation caused by naming items of one of the categories. Because of the way images were obtained, only the left frontal temporal areas could be assessed.[54] In another study involving subjects naming only animals or furniture items, activated areas were located not too far apart within the brain of each subject. Several such areas were found in each subject, in the *left* frontal and temporal lobes. Therefore, this study provides further evidence for distributed maplike semantic representations of semantic information. Two positron emission tomography studies of category-specific brain activation found further evidence of localized storage of high-level representations. Both studies involved the categories of animals and tools and provide converging evidence for the localized storage of aspects of these representations.

Associative Disorders Are Revealed by Abnormal Semantic Priming in Schizophrenic Patients

Schizophrenic patients suffering from thought disorders exhibit a number of language-related abnormalities.[55] Their utterances in normal conversations contain more repetitions, which may occur at the level of syllables, words, or phrases, as well as more **associative intrusions**—inappropriate intrusions of words that are often very remotely associated with some previous words of the utterance. The finding that schizophrenic patients produce more nonstandard associations in the word association test has been replicated a number of times.[56]

To investigate the associative processes of schizophrenic patients in more detail, researchers have used the above-mentioned semantic priming protocol.[57] When one study was run, a most unexpected result was obtained: an *increased* semantic priming effect in schizophrenic patients who suffered from formal thought disorder (TD) compared to the effect in non-thought-disordered (NTD) schizophrenic patients and normal control subjects.[57,58]

This experimental result is highly unexpected, because the very patients who by definition suffer from strange and unexpected associations benefit more from normal associations in the lexical decision experiment. The increased semantic priming effect (due to normal associations) can be reconciled with the presence of thought disorder (i.e., pathology of associations) by the spreading activation hypothesis of lexical access, if it is assumed that activation during lexical access

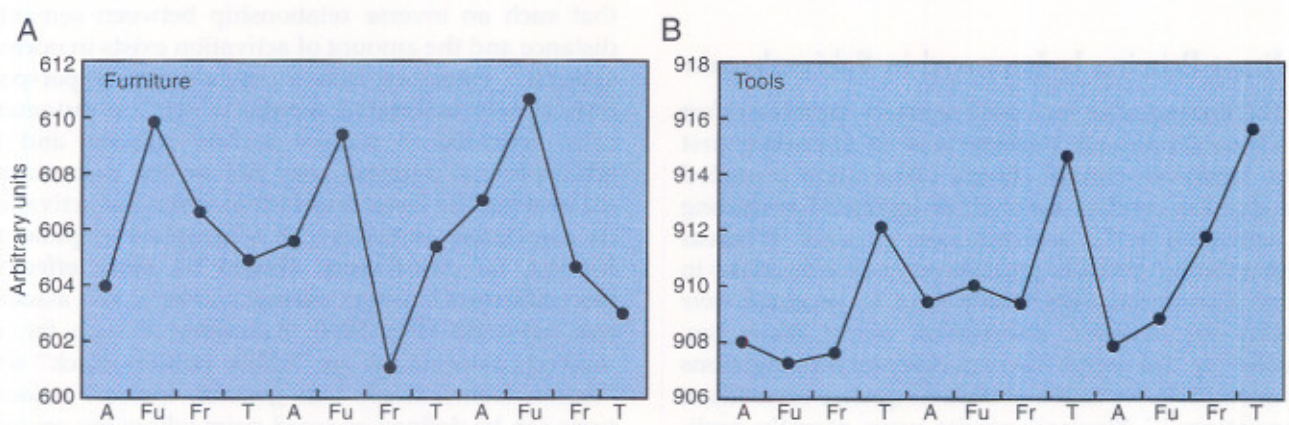

FIGURE 59.8 Results of MRI studies using picture naming as the primary task. Within a left frontal brain area related to semantic processing different sites become active depending on whether the subject sees pictures of furniture (Fu), animals (A), fruit (Fr), or tools (T). (A) Activity level of one set of sites during presentation of different types of pictures. These sites are most sensitive to furniture. (B) Activity level of another set of sites that most sensitive to tools.

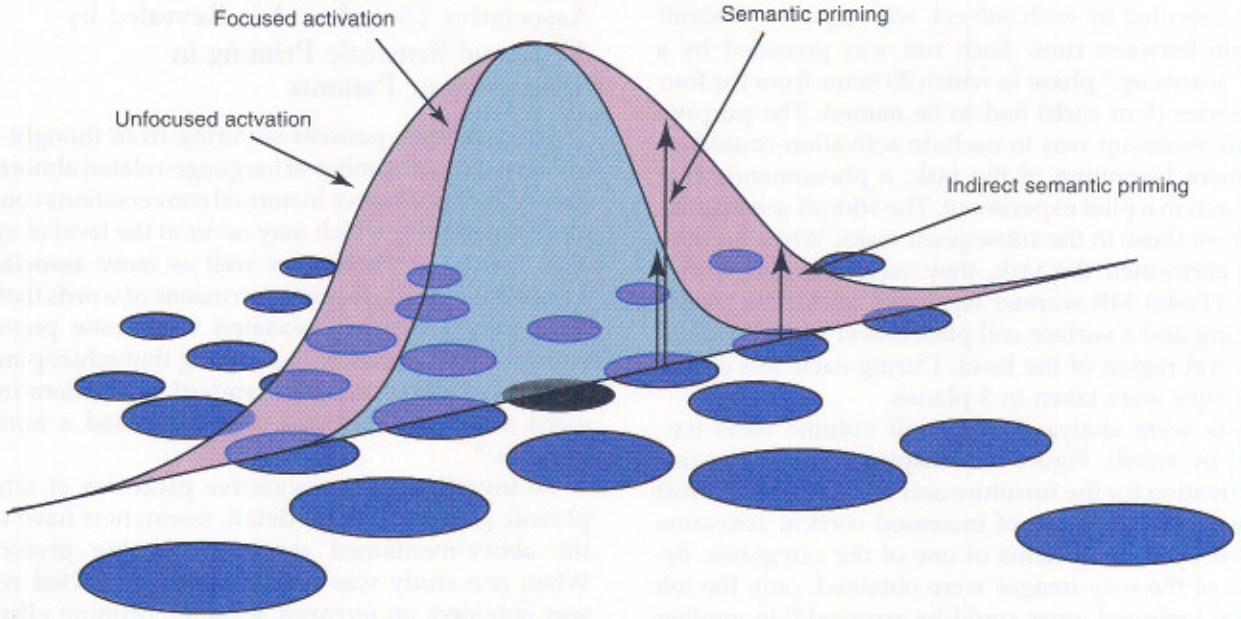

FIGURE 59.9 Illustration of how direct and indirect priming help us to study the spread of semantic activation. The elipses represent nodes in memory that code semantic information. Activation of the solid elipse is driven by an input item (e.g., the prime "white" in Fig. 59.6). The activation spreads to closely related nodes (which include "black" as one example), activated nodes within the focused gray curve. However, there may also be unfocused activation beyond the gray area, shown by the white area. Indirect priming (for example, of "night," which is associated to black but only indirectly to white) can be used to study activation that spreads beyond the focused area.

spreads faster and farther in the semantic network in TD schizophrenic patients than in normal subjects. This causes the increased activation of normal associations on the one hand (and hence an increased semantic priming effect) and the intrusion of oblique and unusual associations into utterances, because activity spreads out quickly to farther nodes, causing pathological intrusions.

Indirect Priming Is Improved in Schizophrenia

The finding of an increased semantic priming effect in TD schizophrenic patients was an important first step. However, careful clinical observation provided the clue for another measure of increased spreading of activation in TD schizophrenic patients. Whereas schizophrenic patients produce *fewer close associates* in word association tests,[59] they tend to produce *more indirect*, or *mediated, associations*. Many years ago, Bleuler[33] (p. 26) noted: "In experimental investigations of association, we find a notable frequency of 'mediate associations.'" Bleuler's results were recently replicated using a standard word association task given to 20 normal control subjects and 20 schizophrenic patients.[53] Compared to the normal control subjects, the TD schizophrenic patients showed fewer standard associations, fewer associations driven by meaning, and more indirect associations.

Finally, the following line of argument, derived from the associative network model of semantic priming, can be made to suggest a better measure for the spreading of activation in such networks. According to this model, activation is postulated to dissipate with distance. Empirical evidence indicates that such an inverse relationship between semantic distance and the amount of activation exists in normal subjects.[60] Therefore, from a psycholinguistic perspective, closely associated words, which are automatically activated in normal control subjects and in schizophrenic patients, may not be the best stimuli for proving the faster and farther spread of activation in associative networks of schizophrenic patients. Instead, far associations should be more effective discriminators between normal and activated associative networks (Fig. 59.9). Examples of such far, or indirect, associations are "chalk (white)–black" and "lemon (sour)–sweet." In general, indirect associations can be defined as word pairs where the connection between the words is obvious only via a mediating associated word. Applied to lexical decision experiments, this means that the target is an association to an association of the prime.

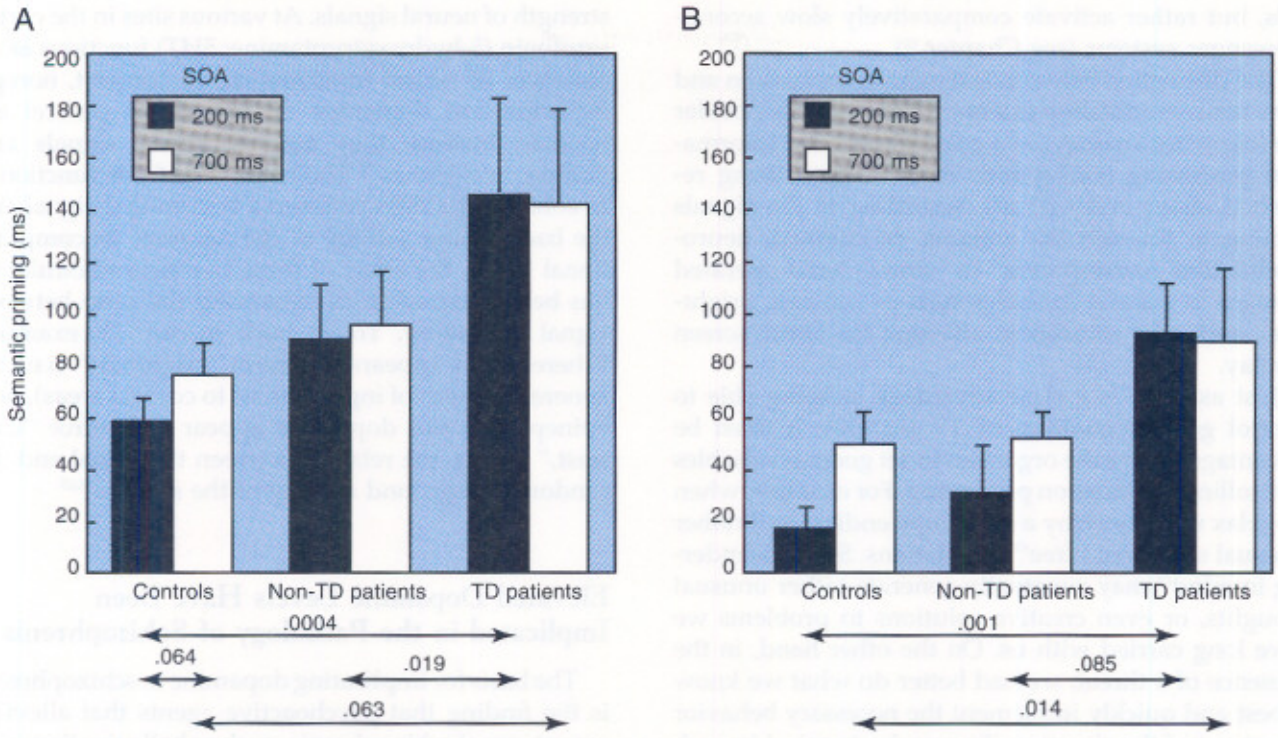

FIGURE 59.10 Direct (A) and indirect (B) priming score (difference between RT to related and unrelated words) for 50 normal controls, 21 non-thought-disordered schizophrenics, and 29 thought-disordered schizophrenics. Reproduced with permission from Spitzer.[62]

Indirect Semantic Priming Proved to Be a Better Measure of Spreading Activation Than Direct Semantic Priming

Two studies of indirect and direct semantic priming in normal control subjects and schizophrenic patients clearly confirmed the validity of these ideas. As can be seen in Fig. 59.10, the spreading of activation reaches more distant nodes only after several hundred milliseconds in normal subjects. In this group a significant indirect semantic priming effect was observed only when the target word appeared 700 ms after the prime word. This interval between the beginning of the prime and the beginning of the target is called stimulus onset asynchrony (SOA). In contrast, TD schizophrenic patients displayed a significant indirect semantic priming effect when the SOA was short, i.e., only 200 ms. This was interpreted as a sign of the fast and far spreading of activation in this group.

Differences in (direct) semantic priming between TD schizophrenic patients and normal subjects—particularly at the long SOA—can be attributed exclusively to the general slowness of the patients. When this slowness was taken into account, no difference in the priming effect was visible. This was not the case, however, for the indirect semantic priming effect. This study not only provided support for the spreading activation model of thought disorder but also suggested that indirect semantic priming in lexical decision tasks with a short SOA is a good measure of spatiotemporal characteristics of the spreading of activation in semantic networks.

SCHIZOPHRENIA AND DOPAMINE

The Modulatory Activity of Neurons Changes Relatively Slowly Compared to the Rate of Neurotransmission

Most neurons in the brain are engaged in excitatory or inhibitory point-to-point fast signal transmission (see Chapter 8). Only a small fraction of all brain neurons do not engage in fast information processing of signals. These cells have connections that are diffusely and widely spread across the cortex. They use substances such as monoamines (norepinephrine, dopamine), serotonin, and acetylcholine for transmission, and receptor sites for these substances can be found in many cortical areas and within the various layers of the cortex. The transmitters do not act on ion chan-

nels, but rather activate comparatively slow second-messenger systems (see Chapter 9).

The distinction between fast neurotransmission and slow neuromodulation is a matter of degree. Consider the following analogy: In a color TV set, fast information processing corresponds to the screen being refreshed about every 30 ms, according to the signals coming in through the antenna. In contrast, neuromodulation corresponds to slow, hand-operated changes in general variables such as contrast, brightness, and color saturation, affecting the entire screen display.

Just as there is a clear advantage in being able to control general qualities of TV pictures, it must be advantageous for the organism to set general variables controlling information processing. For example, when we relax we may enjoy a seemingly endless and rather unusual stream of "free" associations. Such "wandering in mind" may eventually generate rather unusual thoughts, or even creative solutions to problems we have long carried with us. On the other hand, in the presence of a threat, we had better do what we know is best and quickly implement the necessary behavior to get out of the situation. To wander in mind in such circumstances would clearly be disadvantageous. In other words, just as the adjustment of contrast or brightness of a TV improves the overall quality of the display in various circumstances (such as conditions of lighting or picture displayed), the adjustment of general processing features is advantageous for the organism because it allows better adaptation to the varying demand characteristics of an ever-changing environment.

Which General Processing Features Can Be Assumed to Be under the Control of Neuromodulators?

From a subjective point of view, we are all familiar with such general functional characteristics of mentation as vigilance and mood. We are more or less alert and do things in our mind with more or less ease; we find ourselves in various mood states and accordingly think differently about different aspects of the environment. Vigilance and affect are two concepts derived from subjective experience that can be reframed in neuromodulatory terms. Of course, it is unlikely that there is a clear-cut match between verbal abstractions from subjective accounts of mental processes and neuronal modulatory systems; however, it is important to realize that just as vigilance and affect are general aspects of mentation, neuromodulatory systems control general aspects of fast neural transmission.

One such general aspect of mentation concerns the strength of neural signals. At various sites in the cortex serotonin (5-hydroxytryptamine; 5HT) functions as an enhancer of neural transmission. In contrast, norepinephrine and dopamine do not act as general enhancers; instead, they amplify strong signals and dampen weak ones.[63] The net effect of this function is to enhance the ratio between a transmitted signal and the background activity of the neurons. In computational terms, the effect of these two neuromodulators has been referred to as enhancing the ratio between signal and noise. To go back to our TV example: Whereas 5HT appears to control "brightness" (i.e., the general strength of input signals to cortical areas), norepinephrine and dopamine appear to control "contrast," that is, the relation between the signal and the random background activity of the system.[64,65]

Elevated Dopamine Levels Have Been Implicated in the Pathology of Schizophrenia

The basis for implicating dopamine in schizophrenia is the finding that psychoactive agents that alleviate symptoms of schizophrenia, such as hallucinations and delusions, block dopamine receptors in the brain.[66] Moreover, a side effect of these agents is parkinsonism, a disorder chacterized by motor phenomena that are caused by a hypodopaminergic state.

These findings led to the so-called **dopamine hypothesis of schizophrenia.** The standard dopamine hypothesis of schizophrenic psychopathology attributes schizophrenic psychopathology to elevated dopamine levels in the brain. Although this hypothesis dominated the literature for decades, much evidence has accrued to challenge it. First, it never fit several clinical observations. For example, whereas neuroleptic drugs, as dopamine-receptor antagonists are called in the context of psychopharmacology, produce Parkinsonlike symptoms almost immediately, their action on psychopathological symptoms is delayed by days or sometimes even weeks. Second, some symptoms of schizophrenia do not respond to neuroleptic drugs and sometimes even become worse under therapy. In particular, so-called "positive" symptoms such as hallucinations and delusions may respond favorably to neuroleptic drugs, whereas this is not the case with most "negative" symptoms such as social withdrawal, anhedonia, and apathy. Third, if schizophrenia were necessarily associated with a hyperactive dopamine system, one would expect increases in the levels of dopamine and its metabolites in the brain of schizophrenic patients. Yet neither *in vivo* nor postmortem studies have produced consistent results in this regard. Finally, homeostatic mechanisms that are built into the dopamine

system would act to restore dopamine equilibrium following destabilization.

Negative symptoms in schizophrenia may be due to a *decrease* in dopaminergic activity rather than to an increase.[66–72] Various proposals have been based on this concept. For example, negative symptoms may be caused by low prefrontal dopamine activity, which leads to excessive dopamine activity in mesolimbic dopaminergic neurons, which in turn may eventually lead to positive symptoms.[68]

Dopamine May Modulate Signal-to-Noise Ratios in the Neuronal Network

Several lines of evidence suggest that dopamine modulates one general parameter of cortical information processing, namely, its signal-to-noise ratio.[65] According to this model, a decreased dopaminergic activation of cortical areas leads to a decrease in the functional focus of cortical neuronal network activity, thereby reducing the ability to produce appropriate output.

Whereas a high signal-to-noise ratio may at first appear to be desirable under any conditions, it can also at times be counterproductive. As discussed earlier, network models suggest that noise is an important factor driving neuroplasticity. In semantic networks, the relative absence of noise, for example, that produced by a state of moderate anxiety, may cause a more focused activation of ideas, concepts, and meanings. The upside of this effect is that under a given threat, human beings will engage in the one behavior they have learned to be appropriate. With regard to language, it has been demonstrated that stress and anxiety can lead to the production of an increased number of standard associations (such as black–white, doctor–nurse) in normal subjects.[73] In other words, anxiety may order the thoughts of normal subjects, turning them in the opposite direction of the type of thought disorders discussed above. The downside of such highly focused activation characteristics of semantic networks is that creativity is less likely to occur. We all know that anxious candidates will not do too well on an examination because they will not be able to produce creative solutions of problems that require more than the reproduction of rote-learned facts. To come up with "creative" solutions, we need to "unfocus" our mental activity to some extent and allow for the intrusion of unusual thoughts. In sum, the capacity to modulate the relative amount of signal and noise appears to be highly advantageous to an organism, and therefore neuromodulatory systems may have evolved.

L-Dopa Has Been Used to Study the Effects of Dopamine in the Central Nervous System

If dopamine modulates the signal-to-noise ratio in cortical networks, if such networks are involved in the storage of semantic information in the form of maps, and if these maps are accessed during semantic information processing more or less reliably (i.e., with more or less noise involved), the ingestion of L-dopa, a precursor of dopamine and norepinephrine (see Chapter 8), should cause an increase in the focus of activation in semantic networks, and hence, a decrease in the effects of spreading activation, which are represented in Fig. 59.9. (L-dopa is used because dopamine itself cannot pass the blood–brain barrier.)

Because L-dopa is a precursor of dopamine and norepinephrine and both substances have been implicated in modulating the signal-to-noise ratio in cortical networks,[64,69,74] the effects of the two catecholamines are difficult to discern. However, neuroanatomical considerations render dopamine the more likely candidate when it comes to the modulation of semantic processes. The noradrenergic system originates in the locus ceruleus and projects mainly to primary sensory cortical areas, modulating sensory input from the thalamus. In contrast, the dopaminergic projections arise from the ventral tegmental area and terminate predominantly in frontal areas, which are involved in language processing. Moreover, dopamine receptors (mainly D1 receptors) have been demonstrated on spines of pyramidal cells in cortical layer III, where corticocortical projection neurons are located, linking association cortices with each other.[75–77] Anatomical properties thus make the dopaminergic mesocortical neurons particulaarly suited to exerting a modulatory effect on the processing of semantic information.

The hypothesis of a dopaminergic neuromodulatory influence on semantic networks was directly tested in a study on indirect semantic priming in normal volunteers.[78] In a double-blind placebo-controlled design, a speeded lexical decision task and indirectly related word pairs as well as nonrelated word pairs were used to assess the effect of 100 mg of L-dopa (plus 25 mg benserazide, a peripheral decarboxylase inhibitor) on the time course of spreading activation in 31 normal subjects. If dopamine causes a sharper focus of lexical activation, the small indirect priming effect that normal subjects display at longer SOAs should decrease. The results of the study were in line with this hypothesis: When a long stimulus onset asynchrony was used to elicit indirect priming in normals, L-dopa produced a significant decrease of the indirect semantic priming effect from 28 to 14 ms. A small, nonsignificant reduction in semantic priming

indicated again that this measure may be less sensitive to changes in the spreading of activation in semantic networks. This study provided direct support for the hypothesis that dopamine increases the signal-to-noise-ratio in semantic networks, that is, causes a decreased spreading of activation during the process of lexical access. Indirectly, these data provide some support for the hypothesis that formal thought disorder is the result of a decreased dopaminergic tone.

Schizophrenics Also Show Working Memory Deficits

Since the time of the German psychiatrist Emil Kraepelin (1856–1926), the prefrontal cortex has been linked to schizophrenic psychopathology. However, until the advent of recent neuroimaging methods, little clear-cut evidence had been produced in support of this notion. This changed when single-photon emission tomography (SPECT) and positron emission tomography (PET) were used to image the brains of schizophrenic patients. The majority of the SPECT studies and quite a few of the PET studies suggest that schizophrenic patients suffer from a lower blood flow through the prefrontal cortex. In particular, more recent studies demonstrated a failure of activation of the prefrontal cortex in patients performing a task that is known to activate the prefrontal cortex in normal subjects, such as the Wisconsin Card Sorting Test, the Continuous Performance Test, the Porteus Mazes, or the Tower of London Test.[79,80] This observation led to the somewhat vague notion of "hypofrontality" in schizophrenia.[81]

The frontal lobes can be conceived as the site of the psychological function of working memory as discussed earlier in this chapter (also see Chapter 56). Working memory has been linked to dopamine activity in this brain area in human beings.[82] Because dopamine functioning is clearly involved in the pathogenesis of schizophrenia, because frontal cortex dysfunction in schizophrenia has been demonstrated by neuroimaging studies, and because delayed response tasks require subjects to guide their responses by memory of information newly stored for each trial, such tasks seem suitable for tapping working memory deficits in schizophrenic patients.[83] Moreover, working memory deficits have been demonstrated directly in schizophrenic patients using various delayed-response tasks.[63,84,85]

Features of schizophrenic thought may also be explained as a combined dysfunction of associative semantic and working memory. As we have already seen, the particular kind of schizophrenic concretism, that is, the tendency to make a remote aspect of a concept overly concrete, can easily be explained as the combined effect of a disinhibited (unfocused) associative memory and a reduced capacity of working memory. Furthermore, the clinically relevant aspect of schizophrenic thought and behavior, the patients' ubiquitous lack of sensitivity to context,[86,87] can easily be accounted for in terms of working and associative memory. The patients' failure to make appropriate use of contextual evidence in the production and understanding of language and in goal-directed behavior may be caused by the inability to keep relevant information "in mind" while pursuing a certain project. This relevant information must be represented in working memory, because it must be permanently used to guide behavior in the absence of immediate perceptual cues or even despite perceptual cues that suggest some alternative behavior.

Summary

Cognitive neuroscience is the most recent name for the human endeavor to understand the nature of mind and how it is related to the brain. Cognitive neuroscience provides a framework for the study of higher mental processes. In this chapter we reviewed the concepts of working memory, problem solving, and psychopathology. In doing so, we have drawn on concepts and knowledge of attention, language, memory, and emotion described in previous chapters of this section. Concepts such as neuroplasticity and neuromodulation have been used to understand the changes in such systems over long and short periods of time. We believe that there remains very great potential for the use of this framework in understanding higher mental functional and its pathologies.

References

1. Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63:** 81–97.
2. Peterson L. R., and Peterson, M. J. (1959). Short-term retention of individual verbal items. *J. Exp. Psychol.* **58:** 193–198.
3. Sternberg, S. (1966). High speed scaning in human memory. *Science* **153:** 652–654.
4. Wilson, F. A. W., O'Scalaidhe, S. P., and Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* **260:** 1955–1958.
5. Smith, E. E., and Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognit. Psychol.* **33:** 5–42.
6. Smith, E. E., Jonides, J., and Koeppe, R. A. (1996). Dissociating verbal and spatial memory using PET. *Cereb. Cortex* **6:** 11–20.
7. Baddeley, A. D., Thompson, N., and Buchanan, M. (1975). Word length and the structure of short term memory. *J. Verb. Learn. Verb. Behav.* **14:** 578–589.
8. Shallice, T. (1988). *From Neuropsychology to Mental Structure.* Cambridge University Press, Cambridge, UK.

9. Awh, E., Jonides, J., Smith, E. E., Schumacher, E., Koeppe, R., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from PET. *Psychol. Sci.* **7:** 25–31.
10. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature (London)* **331:** 585–589.
11. Paulesu, E., Frith, C. D., and Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature (London)* **362:** 342–345.
12. Gevins, A., and Cutillo, B. (1993). Spatiotemporal dynamics of component processing in human working memory. *Electroencephalogr. Clin. Neurophysiol.* **1:** 1–17.
13. Cohen, J., Forman, S. D., Braver, S., Casey, B. J., Servan-Schreiber, D., and Noll, D.C. (1995). Activation of prefrontal cotex in a nonspatial working memory task with functional MRI. *Hum. Brain Mapp.* **1:** 291–304.
14. Fuster, J. (1995). *Memory in the Cerebral Cortex.* MIT Press, Cambridge, MA.
15. Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* **19:** 450–466.
16. Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychol. Rev.* **99:** 122–149.
17. McCarthy, R., and Warrington, E. K. (1990). *Cognitive Neuropsychology: A Clinical Introduction.* Academic Press, San Diego. CA.
18. De Renzi, E., and Lucchelli, F. (1994). Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex* **30:** 3–25.
19. Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognit. Psychol.* **10:** 302–323.
20. Craik, F. I. M., and Jennings, J. M. (1992). Human memory. In *Handbook of Aging and Cognition* (F. I. M. Craik and T. A. Salthouse, eds.). Erlbaum, Hillsdale, NJ.
21. Campbell, J. I. D., and Charness, N. (1990). Age-related declines in working memory skills: Evidence from a complex calculation task. *Dev. Psychol.* **26:** 879–888.
22. Roland, P. C., and Friberg, L. (1985). Localization of cortical areasmactivated by thinking. *J. Neuropsychol.* **53:** 1219–1243.
23. Dehaene, S., and Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Math. Cognition* **1:** 83–120.
24. Schank, R. C., and Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures.* Erlbaum, Hillsdale, NJ.
25. Dehaene, S., and Changeux, J. P. (1989). A simple model of prefrontal cortex function in delayed-response tasks. *J. Cognit. Neurosci.* **1:** 244–261.
26. Dehaene, S., and Changeux, J. P. (1991). The Wisconsin Card Sorting Test: Theoretical analysis and modelling in a neuronal network. *Cereb. Cortex* **1:** 62–79.
27. Levine, D. S., and Prueitt, P. S. (1989). Modelling some effects of frontal lobe damage—novelty and perseveration. *Neural Networks* **2:** 103–116.
28. Mishkin, M., and Appenzeller, T. (1987). The anatomy of memory. *Sci. Am.* **256**(6): 80–89.
29. Cohen, J. D., Dunbar, K., and McClelland, J. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychol. Rev.* **97:** 332–361.
30. Dominey, P., Arbib, M., and Joseph, J. P. (1995). A model of cortico-striatal plasticity for learning occulomotor associations and sequences. *J. Cognit. Neurosci.* **7:** 311–337.
31. Guigon, E., Dorizzi, B., Burnod, Y., and Schultz, W. (1995). Neural correlates of learning in the prefrontal cortex of the monkey: A predictive model. *Cereb. Cortex* **2:** 135–147.
32. Barto, A. G., Sutton, R. S., and Anderson, C. W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Trans. Syst. Man. Cybernetic.* **13:** 834–846.
33. Bleuler, E. (1911/1950). *Dementia praecox or the Group of Schizophrenias* (J. Ziskin and N. D. Lewis, transl.). International Universities Press, New York.
34. Aitchison, J. (1987). *Words in the Mind.* Blackwell, Oxford, and Cambridge, MA.
35. Levelt, W. J. M. (1989), *Speaking: From Intention to Articulation.* MIT Press, Cambridge, MA and London.
36. Miller, G. A., and Glucksberg, S. (1988). Psycholinguistic aspects of pragmatics and semantics. In *Steven's Handbook of Experimental Psychology* (R. C. Atkinson, R. J. Herrnstein, G. Lindzey, and R. D. Luce, eds.), Vol. 2, pp. 417–472. Wiley, New York.
37. Meyer, D. E., and Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *J. Exp. Psychol.* **20:** 227–234.
38. Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In *Basic Progresses in Reading and Visual Word Recognition* (D. Besner and G. W. Humphreys, eds.), pp. 264–333. Erlbaum, Hillsdale, NJ.
39. Collins, A. M., and Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychol. Rev.* **82:** 407–428.
40. Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *J. Exp. Psychol. Gen.* **106:** 226–254.
41. Ritter, H., and Kohonen, T. (1989). Self-organizing semantic maps. *Biol. Cybernet.* **61:** 241–254.
42. Creutzfeld, D. O. (1995). *Cortex Cerebri.* Oxford University Press, Oxford.
43. Thomson, A. M., and Deuchars, J. (1994). Temporal and spatial properties of local circuits in neocortex. *Trends Neurosci.* **17**(3): 119–126.
44. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybernet.* **43:** 59–69.
45. Kohonen, T. (1989). *Self-Organization and Associative Memory,* 3rd ed. Springer, Berlin.
46. Penfield, W., and Rasmussen, T. (1950). *The Cerebral Cortex of Man: A Clinical Study of Localization and Function.* Macmillan, New York.
47. Merzenich, M. M., and Sameshima, K. (1993). Cortical plasticity and memory. *Curr. Opin. Neurol.* **3:** 187–196.
48. Zeki, S. (1993). *A Vision of the Brain.* Blackwell, Oxford.
49. Tootell, R., Kwong, K., Belliveau, J., Baker, J., Stern, C., Hockfield, S., Breiter, H., Born, R., Benson, R., Brady, T., and Rosen, B. (1993). Mapping human visual cortex: Evidence from functional MRI and histology. *Invest. Ophthalmol. Visual Sci.* 813.
50. Woolsey, T. A., and van der Loos, H. (1970). The structural organization of layer IV in the somato-sensory region of the mouse cerebral cortex. *Brain Res.* **17:** 204–242.
51. Caramazza, A., Hillis, A., Leek, E. C., and Miozzo, M. (1994). The organization of lexical knowledge in the brain: Evidence from category- and modality-specific deficits. In *Mapping the Mind* (L. A. Hirschfeld and S. A. Gelman, eds.), pp. 68–84. Cambridge University Press, Cambridge, UK.
52. Farah, M. J., and Wallace, M. A. (1992). Semantically-bounded anomia: Implications for the neural implementation of naming. *Neuropsychologia* **30:** 609–621.
53. Spitzer, M., Kwong, K. K., Kennedy, W., Rosen, B. R., and Belli-

veau, J. W. (1995). Category specific brain activation of fMRI during picture naming. *Neuroreport* **6:** 2109–2112.
54. Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., and Ungerleider, L. G. (1995). Discrete cortical regions associate with knowledge of color and knowledge of action. *Science* **270:** 102–105.
55. Maher, B. A., and Spitzer, M. (1993). Thought disorder and language behavior in schizophrenia. In *Linguistic Disorders and Pathologies. Handbücher der Sprach-und Kommunikationswissenschaft* (G. Blanken, J. Dittmann, H. Grimm, J. C. Marshal, and C. W. Wallesch, eds.), Vol. 9, pp. 522–533. de Gruyter, New York and Berlin.
56. Cramer, P. (1968). *Word Association.* Academic Press, New York and London.
57. Maher, B. A., Manschreck, T. C., Hoover, T. M., and Weisstein, C. C. (1987). Thought disorder and measured features of language production in schizophrenia. In *Positive and Negative Symptoms in Psychosis: Description, Research and Future Directions* (P. Harvey and E. Walker, eds.), pp. 195–215. Erlbaum, Hillsdale, NJ.
58. Manschreck, T. C., Maher, B. A., Milavetz, J. J., Ames, D., Weisstein, C. C., and Schneyer, M. L. (1988). Semantic priming in thought disordered schizophrenic patients. *Schizophr. Res.* **1:** 61–66.
59. Kent, G. H., and Rosenoff, A. J. (1910). A study of associations in insanity. *Am. J. Insanity* **66/67** (Part I): 7–47; (Part II): 317–390.
60. den Heyer, and Briand, K. (1986). Priming single digit numbers: Automatic spreading activation dissipates as a function of semantic distance. *Am. J. Psychol.* **99:** 315–339.
61. Spitzer, M. (1993). Assoziative Netzwerke, formale Denkstörungen und Schizophrenie. *Nervenarzt* **64:** 147–159.
62. Spitzer, M. (1993). The psychopathology, neuropsychology and neurobiology of associative and working memory in schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **243:** 57–70.
63. Morrison, J. H., and Hof, P. R. (1992). The organization of the cerebral cortex: From molecules to circuits. *Discuss. Neurosci.* **9**(2): 7–79.
64. Servan-Schreiber, D., Printz, H., and Cohen, J. D. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio and behavior. *Science* **249:** 892–895.
65. Cohen, J., and Servan-Schreiber, D. (1992). Context, cortex and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychol. Rev.* **12:** 45–77.
66. Carlsson, A. (1988). The current status of the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* **1:** 179–203.
67. Crow, T. J. (1980). Molecular pathology of schizophrenia: More than one disease process? *Br. Med. J.* **137:** 383–386.
68. Davis, K. L., Kahn, R. S., Ko, G., and Davidson, M. (1991). Dopamine in schizophrenia: A review and reconceptualization. *Am. J. Psychiatry* **148:** 1474–1486.
69. Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience* **41:** 1–24.
70. Heritch, A. J. (1990). Evidence for reduced and dysregulated turnover of dopamine in schizophrenia. *Schizophr. Bull.* **16:** 605–615.
71. Mackay, A. V. P. (1980). Positive and negative schizophrenic symptoms and the role of dopamine. *Br. J. Psychiatry* **137:** 379–386.
72. Weinberger, D. R., Berman, K. F., and Illowsky, B. P. (1988). Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia: A new cohort of evidence for a monoaminergic mechanism. *Arch. Gen. Psychiatry* **45:** 606–615.
73. Mintz, S. (1969). Effect of actual stress on word associations. *J. Abnorm. Psychol.* **74:** 293–295.
74. Chiodo, L. A., and Berger, T. W. (1986). Interactions between dopamine and amino acid-induced excitation and inhibition in the striatum. *Brain Res.* **375:** 198–203.
75. Goldman-Rakic, P. S., Leranth, C., Williams, S. M., Mons, N., and Geffard, M. (1989). Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 9015–9019.
76. Smiley, J. F., and Goldman-Rakic, P. S. (1993). Heterogeneous targets of dopamine synapses in monkey prefrontal cortex demonstrated by serial section electron microscopy. *Cereb. Cortex* **3:** 223–238.
77. Smiley, J. F., Levey, A. I., Ciliax, B. J., and Goldman-Rakic, P. S. (1994). D1 dopamine receptor immunoreactivity in human and monkey cerebral cortex: Predominant and extrasynaptic localization in dendritic spines. *Proc. Natl. Acad. Sci. U.S.A.* **91**(12): 5720–5724.
78. Kischka, U., Kammer, T., Weisbrod, M., Meier, S., Thimm, M., and Spitzer, M. (1996). Dopaminergic modulation of semantic network activation. *Neuropsychologia* **34:** 1107–1113.
79. Andrasen, N. C., Rezai, K. R., Alliger, R., Swayze, V. W., Flaum, M., Kirchner, P., Cohen, G., and O'Leary, D. S. (1992). Hypofrontality in neuroleptic-naive patients and in patients with chronic schizophrenia. Assessment with Xenon 133 single-photon emission computed tomography and the Tower of London. *Arch. Gen. Psychiatry,* **49:** 943–948.
80. Andreasen, N. C., Swayze, V. W., Flaum, M., O'Leary, D. S., and Alliger, R. (1994). The neural mechanisms of mental phenomena. In *Schizophrenia: From Mind to Molecule* (N. C. Andreasen, ed.), pp. 49–91. American Psychiatric Press, Washington, DC.
81. Winn, P. (1994). Schizophrenia research moves to the prefrontal cortex. *Trends Neurosci.* **17:** 265–268.
82. Luciana, M., Depue, R. A., Arbisi, P., and Leon, A. (1992). Facilitation of working memory in humans by a D2 dopamine receptor agonist. *J. Cognit. Neurosci.* **4:** 58–68.
83. Goldman-Rakic, P. S. (1991). Cortical dysfunction in schizophrenia: The relevance of working memory. In *Psychopathology and the Brain* (B. J. Carroll and J. E. Barrett, eds.). Raven Press, New York.
84. Park, S., and Holzman, P. S. (1992). Schizophrenic show spatial working memory deficits. *Archi. Gen. Psychiatry* **49:** 975–982.
85. Park, S. (1995). Spatial working memory function in schizophrenia. In *Experimental Psychopathology* (M. Spitzer and B. A. Maher, eds.). Cambridge University Press, New York.
86. Chapman, L. J., Chapman, J. P., and Miller, G. A. (1964). A theory of verbal behavior in schizophrenia. In *Progress in Experimental Personality Research* (B. A. Maher, ed.), Vol. 1, pp. 49–77. Academic Press, New York and London.
87. de Silva, W. P., and Hemsley, D. R. (1977). The influence of context on language perception in schizophrenia. *Br. J. Clin. Psychol.* **16:** 337–345.