62

Learning and Memory

Eric R. Kandel

Irving Kupfermann

Susan Iversen

BEHAVIOR IS THE RESULT OF the interaction between genes and the environment. In earlier chapters we saw how genes influence behavior. We now examine how the environment influences behavior. In humans the most important mechanisms by which the environment alters behavior are learning and memory. Learning is the process by which we acquire knowledge about the world, while memory is the process by which that knowledge is encoded, stored, and later retrieved.

Many important behaviors are learned. Indeed, we are who we are largely because of what we learn and what we remember. We learn the motor skills that allow us to master our environment, and we learn languages that enable us to communicate what we have learned, thereby transmitting cultures that can be maintained over generations. But not all learning is beneficial. Learning also produces dysfunctional behaviors, and these behaviors can, in the extreme, constitute psychological disorders. The study of learning therefore is central to understanding behavioral disorders as well as normal behavior, since what is learned can often be unlearned. When psychotherapy is successful in treating behavioral disorders, it often does so by creating an environment in which people can learn to change their patterns of behavior.

As we have emphasized throughout this book, neural science and cognitive psychology have now found a common ground, and we are beginning to benefit from the increased explanatory power that results from the convergence of two initially disparate disciplines. The rewards of the merger between neural science and cognitive psychology are particularly evident in the study of learning and memory.

In the study of learning and memory we are interested in several questions. What are the major forms of

learning? What types of information about the environment are learned most easily? Do different types of learning give rise to different memory processes? How is memory stored and retrieved?

In this chapter we review the major biological principles of learning and memory that have emerged from clinical and cognitive/psychological approaches. In the next chapter we shall examine learning and memory processes at the cellular and molecular level.

Memory Can Be Classified as Implicit or Explicit on the Basis of How Information Is Stored and Recalled

As early as 1861 Pierre Paul Broca had discovered that damage to the posterior portion of the left frontal lobe (Broca's area) produces a specific deficit in language (Chapter 1). Soon thereafter it became clear that other mental functions, such as perception and voluntary movement, can be related to the operation of discrete neural circuits in the brain. The successes of efforts to localize brain functions led to the question: Are there also discrete systems in the brain concerned with memory? If so, are all memory processes located in one region, or are they distributed throughout the brain?

In contrast to the prevalent view about the localized operation of other cognitive functions, many students of learning doubted that memory functions could be localized. In fact, until the middle of the twentieth century many psychologists doubted that memory was a discrete function, independent of perception, language, or movement. One reason for the persistent doubt is that memory storage does indeed involve many different regions of the brain. We now appreciate, however, that these regions are not equally important. There are several fundamentally different types of memory storage, and certain regions of the brain are much more important for some types of storage than for others.

The first person to obtain evidence that memory processes might be localized to specific regions of the human brain was the neurosurgeon Wilder Penfield. Penfield was a student of Charles Sherrington, the pioneering English neurophysiologist who, at the turn of the century, mapped the motor representation of anesthetized monkeys by systematically probing the cerebral cortex with electrodes and recording the activity of motor nerves. By the 1940s Penfield had begun to apply similar methods of electrical stimulation to map the motor, sensory, and language functions in the cerebral cortex of patients undergoing brain surgery for the relief of focal epilepsy. Since the brain itself does not have pain receptors, brain surgery is painless and can be carried out under local anesthesia in patients that are fully awake. Thus, patients undergoing brain surgery are able to describe what they experience in response to electrical stimuli applied to different cortical areas. On hearing about these experiments, Sherrington, who had always worked with monkeys and cats, told Penfield, "It must be great fun to put a question to the [experimental] preparation and have it answered!"

Penfield explored the cortical surface in more than a thousand patients. On rare occasions he found that electrical stimulation of the temporal lobes produced what he called an *experiential response* —a coherent recollection of an earlier experience. These studies were provocative, but they did not convince the scientific community that the temporal lobe is critical for memory because all of the patients Penfield studied had epileptic seizure foci in the temporal lobe, and the sites most effective in eliciting experiential responses were near those foci. Thus the responses might have been the result of localized seizure activity. Furthermore, the responses occurred in only 8% of all attempts at stimulating the temporal lobes. More convincing evidence that the temporal lobes are important in memory emerged in the mid 1950s from the study of patients who had undergone bilateral removal of the hippocampus and neighboring regions in the temporal lobe as treatment for epilepsy.

The first and best-studied case of the effects on memory of bilateral removal of portions of the temporal lobes was the patient called H.M., studied by Brenda Milner, a colleague of Penfield and the surgeon William Scoville. H.M., a 27-year-old man, had suffered for over 10 years from untreatable bilateral temporal lobe seizures as a consequence of brain damage sustained at age 9 when he was hit and knocked over by someone riding a bicycle. As an adult he was unable to work or lead a normal life. At surgery the hippocampal formation, the amygdala, and parts of the multimodal association area of the temporal cortex were removed bilaterally (Figure 62-1).

H.M.'s seizures were much better controlled after surgery, but the removal of the medial temporal lobes left him with a devastating memory deficit. This memory deficit (or *amnesia)* was quite specific. H.M. still had normal short-term memory, over seconds or minutes. Moreover, he had a perfectly good long-term memory for events that had occurred before the operation. He remembered his name and the job he held, and he vividly remembered childhood events, although he showed some evidence of a retrograde amnesia for information acquired in the years just before surgery. He retained a perfectly good command of language, including his normally varied vocabulary, and his IQ remained unchanged in the range of bright-normal.

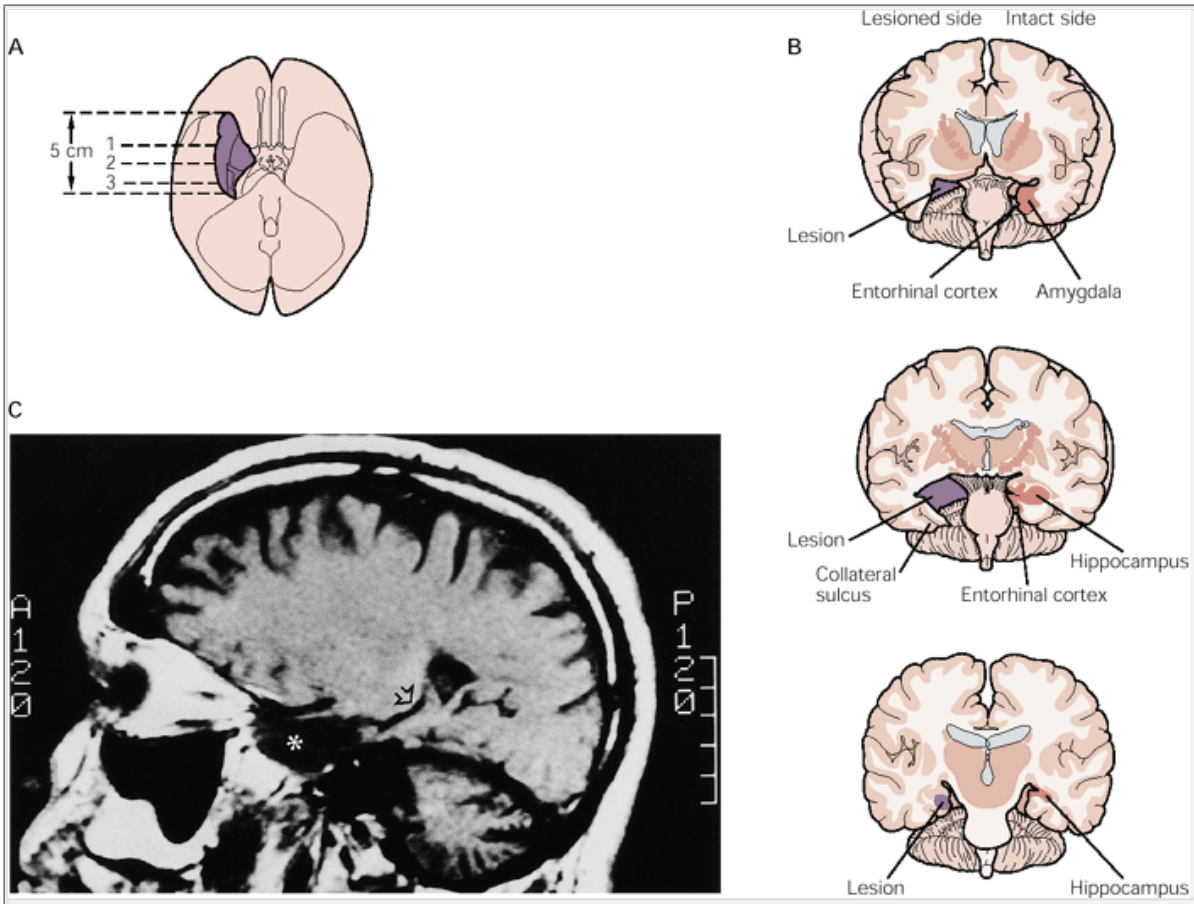


Figure 62-1 The medial temporal lobe and memory storage.

A. The longitudinal extent of the temporal lobe lesion in the patient known as H.M. in a ventral view of the brain.

B. Cross sections showing the estimated extent of surgical removal of areas of the brain in the patient H.M. Surgery was a bilateral, single-stage procedure. The right side is shown here intact to illustrate the structures that were removed. (Modified from Milner 1966.)

C. Magnetic resonance image (MRI) scan of a parasagittal section from the left side of H.M.'s brain. The calibration bar on the right side of the panel has 1 cm increments. The resected portion of the anterior temporal lobes is indicated with an **asterisk.** The remaining portion of the intraventricular portion of the hippocampal formation is indicated with an **open arrow.** Approximately 2 cm of preserved hippocampal formation is visible bilaterally. Note also the substantial cerebellar degeneration obvious as enlarged folial spaces. (From Corkin et al. 1997.)

P.1229

What H.M. now lacked, and lacked dramatically, was the ability to transfer new short-term memory into long-term memory. He was unable to retain for more than a minute information about people, places, or objects. Asked to remember a number such as 8414317, H.M. could repeat it immediately for many minutes, because of his good short-term memory. But when distracted, even briefly, he forgot the number. Thus, H.M. could not recognize people he met after surgery, even when he met them again and again. For example, for several years he saw Milner on an almost monthly basis, yet each time she entered the room H.M. reacted as though he had never seen her before. H.M. had a similarly profound difficulty with spatial orientation. It took him about a year to learn his way around a new house. H.M. is not unique. All patients with extensive bilateral lesions of the limbic association areas of the medial temporal lobe, from either surgery or disease, show similar memory deficits.

The Distinction Between Explicit and Implicit Memory Was First Revealed With Lesions of the Limbic Association Areas of the Temporal Lobe

Milner originally thought that the memory deficit after bilateral medial temporal lobe lesions affects all forms of memory equally. But this proved not to be so. Even though patients with lesions of the medial temporal lobe have profound memory deficits, they are able to learn certain types of tasks and retain this learning for as long as normal subjects. The spared component of

P.1230

memory was first revealed when Milner discovered that H.M. could learn new motor skills at a normal rate. For example, he learned to draw the outlines of a star while looking at his hand and the star in a mirror (Figure 62-2). Like normal subjects learning this task, H.M. initially made many mistakes, but after several days of training his performance was error-free and indistinguishable from that of normal subjects.

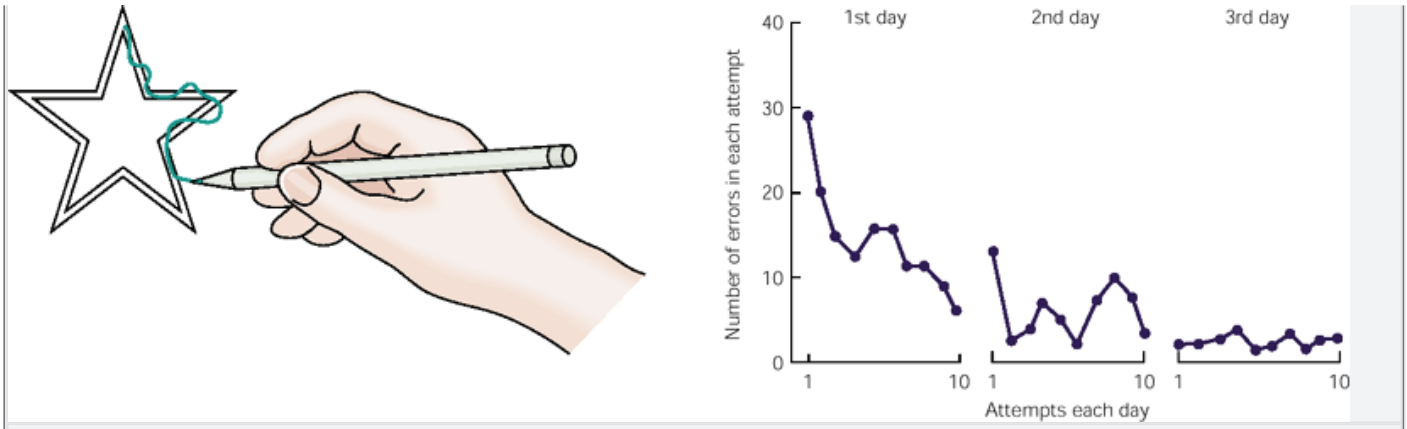


Figure 62-2 The patient H.M. showed definite improvement in any task involving learning skilled movements. He was taught to trace between two outlines of a star while viewing his hand in a mirror. He improved considerably with each fresh test, although he had no recollection that he had ever done the task before. The graph plots the number of times, in each trial, that he strayed outside the outlines as he drew the star. (From Blakemore 1977.)

Later work by Larry Squire and others has made it clear that the memory capacities of H.M. and other patients with bilateral medial temporal lobe lesions are not limited to motor skills. Rather, these patients are capable of various forms of simple reflexive learning, including habituation, sensitization, classical conditioning, and operant conditioning, which we discuss later in this chapter. Furthermore, they are able to improve their performance on certain perceptual tasks. For example, they do well with a form of memory called *priming*, in which the recall of words or objects is improved by prior exposure to the words or object. Thus, when shown the first few letters of previously studied words, a subject with amnesia correctly selects as many previously presented words as do normal subjects, even though the subject has *no* conscious memory of having seen the word before (Figure 62-3)

The memory capability that is spared in patients with bilateral lesions of the temporal lobe typically involves learned tasks that have two things in common. First, the tasks tend to be reflexive rather than reflective in nature and involve habits and motor or perceptual skills. Second, they do not require conscious awareness or complex cognitive processes, such as comparison and evaluation. The patient need only respond to a stimulus or cue, and need not try to remember anything. Thus, when given a highly complex mechanical puzzle to solve the patient may learn it as quickly and as well as a normal person, but will not consciously remember having worked on it previously. When asked why the performance of a task is much better after several days of practice than on the first day, the patient may respond, "What are you talking about? I've never done this task before."

Although these two fundamentally different forms of memory—for skills and for knowledge—have been demonstrated in detail in amnesia patients with lesions of the temporal lobe, they are not unique amnesiacs. Cognitive psychologists had previously distinguished these two types of memory in normal subjects. They refer to information about how to perform something as *implicit memory* (also referred to as *nondeclarative memory*), a memory that is recalled unconsciously. Implicit memory is typically involved in training reflexive motor or perceptual skills. Factual knowledge of people, places, and things, and what these facts mean, is referred to as *explicit memory* (or *declarative memory*). This is recalled by a deliberate, conscious effort (Figure 62-4). Explicit memory is highly flexible and involves the association of multiple bits and pieces of information. In contrast, implicit memory is more rigid and tightly connected to the original stimulus conditions under which the learning occurred

The psychologist Endel Tulving first developed the idea that explicit memory can be further classified as *episodic* (a memory for events and personal experience) or *semantic* (a memory for facts). We use episodic memory when we recall that we saw the first flowers of spring yesterday or that we heard Beethoven's *Moonlight Sonata*

several months ago. We use semantic memory to store and recall objective knowledge, the kind of knowledge we learn in school and from books. Nevertheless, all explicit memories can be concisely expressed in declarative statements, such as "Last summer I visited my grandmother at her country house" (episodic knowledge) or "Lead is heavier than water" (semantic knowledge).

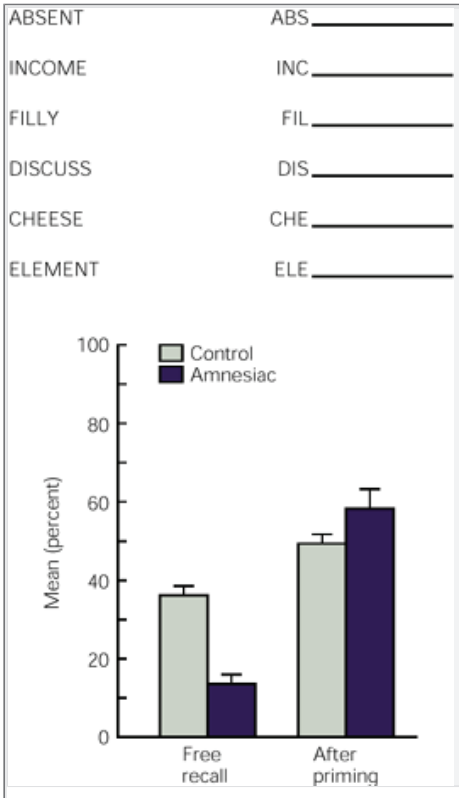


Figure 62-3 In a study of recall of words, amnesiacs and normal control subjects were tested under two conditions. First they were presented with common words and then asked to recall the words (free recall). Amnesiac patients were impaired in this condition. However, when subjects were given the first three letters of a word and instructed simply to form the first word that came to mind (completion), the amnesiacs performed as well as normal subjects. The baseline guessing rate in the word completion condition was 9%. (From Squire 1987.)

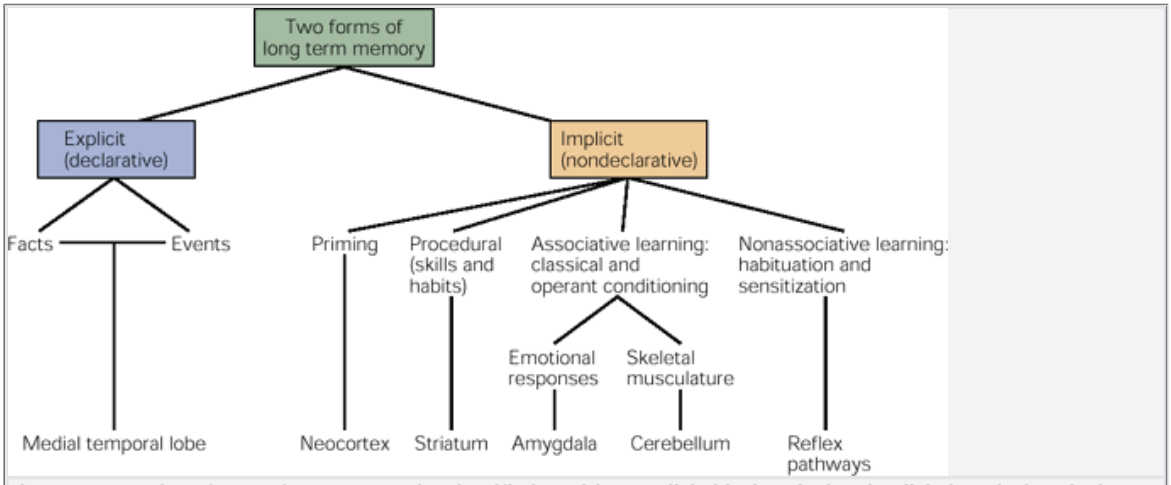


Figure 62-4 Various forms of memory can be classified as either explicit (declarative) or implicit (nondeclarative).

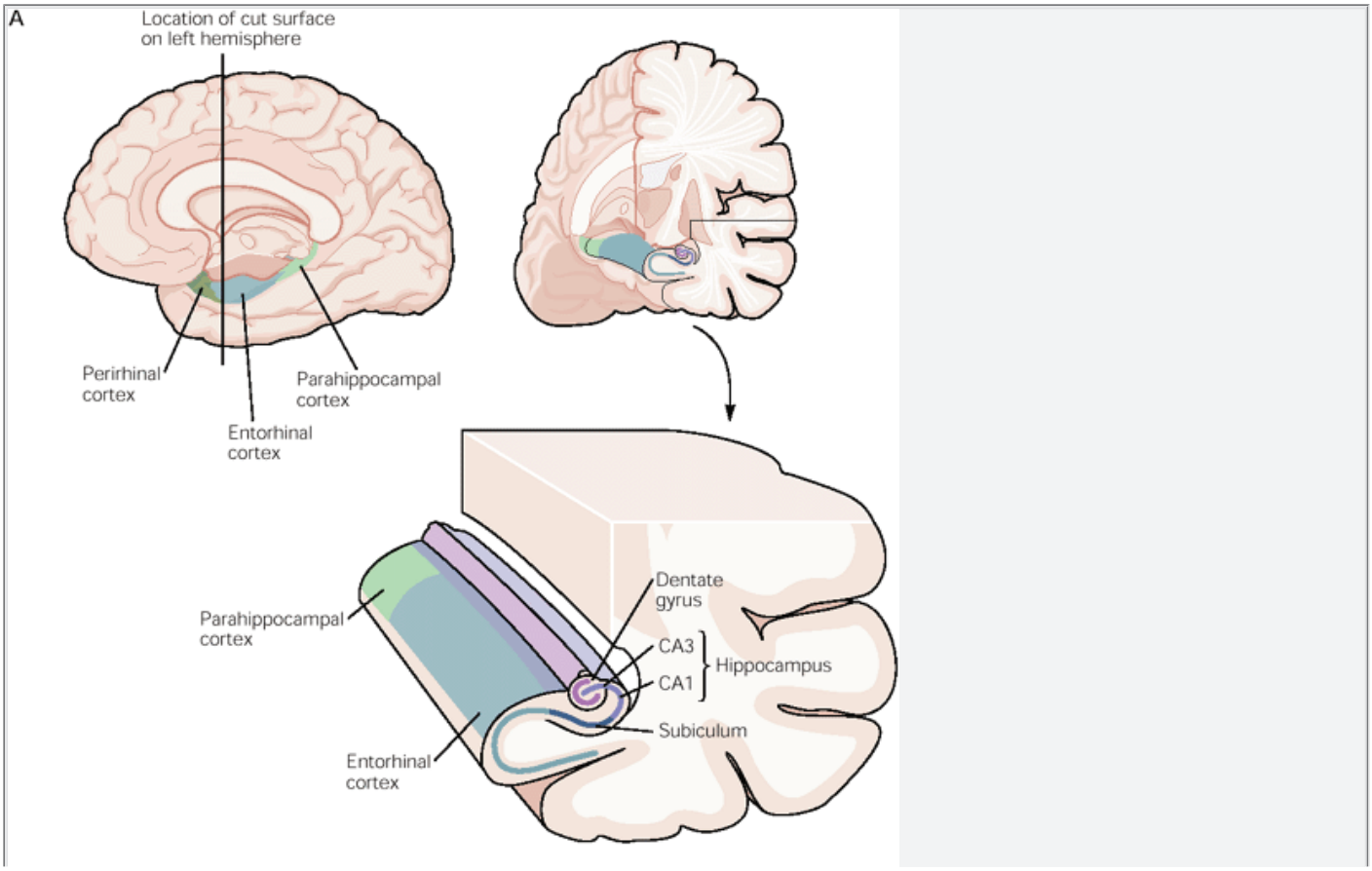
Animal Studies Help to Understand Memory

The surgical lesion of H.M.'s temporal lobe encompassed a number of regions, including the temporal pole, the ventral and medial temporal cortex, the amygdala, and the hippocampal formation (which includes the hippocampus proper, the subiculum, and the dentate gyrus) as well as the surrounding entorhinal, perirhinal, and parahippocampal cortices. Since lesions restricted to any one of these several sectors of the medial temporal lobe are rare in humans, experimental lesion studies in monkeys have helped define the contribution of the different parts of the temporal lobe to memory formation.

Mortimer Mishkin and Squire produced lesions in monkeys identical to those reported for H.M. and found defects in explicit memory for places and objects similar to those observed in H.M. Damage to the amygdala alone had no effect on explicit memory. Although the amygdala stores components of memory concerned with emotion (Chapter 50), it does not store factual information. In contrast, selective damage to the hippocampus or the polymodal association areas in the temporal cortex with which the hippocampus connects—the perirhinal and parahippocampal cortices—produces clear impairment of explicit memory.

Thus, studies with human patients and with experimental animals suggest that knowledge stored as explicit memory is first acquired through processing in one or more of the three polymodal association cortices (the prefrontal, limbic, and parieto-occipital-temporal cortices) that synthesize visual, auditory, and somatic information. From there the information is conveyed in series to the parahippocampal and perirhinal cortices,

then the entorhinal cortex, the dentate gyrus, the hippocampus, the subiculum, and finally back to the entorhinal cortex. From the entorhinal cortex the information is sent back to the parahippocampal and perirhinal cortices and finally back to the polymodal association areas of the neocortex (Figure 62-5).



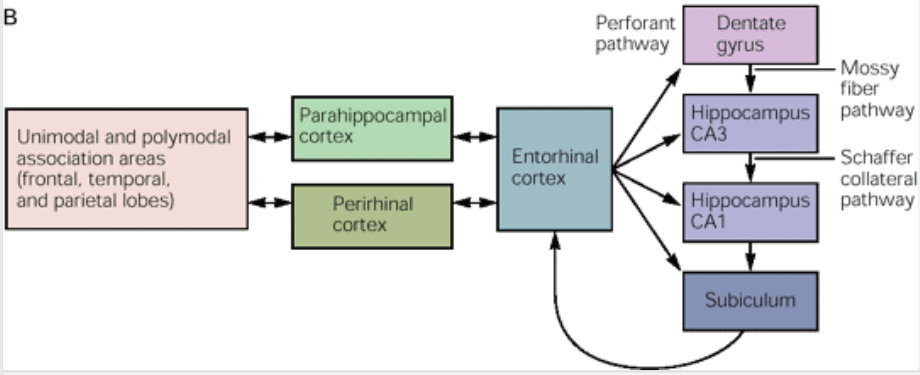


Figure 62-5 The anatomical organization of the hippocampal formation.

A. The key components of the medial temporal lobe important for memory storage can be seen in the medial (**left**) and ventral (**right**) surface of the cerebral hemisphere.

B. The input and output pathways of the hippocampal formation.

Thus, in processing information for explicit memory storage the entorhinal cortex has dual functions. First, it is the main input to the hippocampus. The entorhinal cortex projects to the dentate gyrus via the perforant pathway and by this means provides the critical input pathway through which the polymodal information from the association cortices reaches the hippocampus (Figure 62-5B). Second, the entorhinal cortex is also the major output of the hippocampus. The information coming to the hippocampus from the polymodal association cortices and that coming from the hippocampus to the association cortices converge in the entorhinal cortex. It is therefore understandable why the memory impairments associated with damage to the entorhinal cortex are particularly severe and why this damage affects not simply one but all sensory modalities. In fact, the earliest pathological changes in Alzheimer disease, the major degenerative disease that affects explicit memory storage, occurs in the entorhinal cortex.

Damage Restricted to Specific Subregions of the Hippocampus Is Sufficient to Impair Explicit Memory Storage

Given the large size of the hippocampus proper, how extensive does a bilateral lesion have to be to interfere with explicit memory storage? Clinical evidence from several patients, as well as studies in experimental animals, suggests that a lesion restricted to *any* of the major components of the system can have a significant effect on memory storage. For example Squire, David Amaral, and their collegues found that the patient R.B. had only one detectable lesion after a cardiac arrest—a destruction of the pyramidal cells in the CA1 region of the hippocampus. Nevertheless, R.B. had a defect in explicit memory that was qualitatively similar to that of H.M., although quantitatively it was much milder.

The different regions of the medial temporal lobe may, however, not have equivalent roles. Although the hippocampus is important for object recognition, for example, other areas in the medial temporal lobe may be even more important. Damage to the perirhinal, parahippocampal, and entorhinal cortices that spares the underlying hippocampus produces a greater deficit in memory storage, such as object recognition than do selective lesions of the hippocampus that spare the overlying cortex

On the other hand, the hippocampus may be relatively more important for spatial representation. In mice and rats lesions of the hippocampus interfere with memory for space and context, and single cells in the hippocampus encode specific spatial information (Chapter 63). Moreover, functional imaging of the brain of normal human subjects shows that spatial memories involve more intense hippocampal activity in the right hemisphere than do memories for words, objects, or people, while the latter involve greater activity in the hippocampus in the dominant left hemisphere. These physiological findings are consistent with the finding that lesions of the right hippocampus give rise to problems with spatial orientation, whereas lesions of the left hippocampus give rise to defects in verbal memory (Figure 62-6).

Explicit Memory Is Stored in Association Cortices

Lesions of the medial temporal lobe in patients such as H.M. and R.B. interfere only with the long-term storage of *new* memories. These patients retain a reasonably good memory of earlier events, although with severe lesions such as those of H.M. there appears to be some retrograde amnesia for the years just before the operation. How does this come about?

The fact that patients with amnesia are able to remember their childhood, the lives they have led, and the factual knowledge they acquired before damage to the hippocampus suggests that the hippocampus is only a temporary way station for long-term memory. If so, long-term storage of episodic and semantic knowledge would occur in the unimodal or multimodal association areas of the cerebral cortex that initially process the sensory information

For example, when you look at someone's face, the sensory information is processed in a series of areas of the cerebral cortex devoted to visual information, including the unimodal visual association area in the inferotemporal cortex specifically concerned with face recognition (see Box 28-1 and Chapter 28). At the same time, this visual information is also conveyed through the mesotemporal association cortex to the parahippocampal, perirhinal, and entorhinal cortices, and from there through the perforant pathway to the hippocampus. The hippocampus and the rest of the medial temporal lobe may then act, over a period of days or weeks, to facilitate storage of the information about the face initially processed by the visual association area of the inferotemporal lobe. The cells in the visual association cortex concerned with faces are interconnected with other regions that are thought to store additional knowledge about the person whose face is seen, and these connections could also be modulated by the hippocampus. Thus the hippocampus might also serve to bind together the various components of a richly processed memory of a person.

Viewed in this way the hippocampal system would mediate the initial steps of long-term storage. It would then slowly transfer information into the neocortical storage system. The relatively slow addition of information to the neocortex would permit new data to be stored in a way that does not disrupt existing information. If the association areas are the ultimate repositories for explicit memory, then damage to association cortex should destroy or impair recall of explicit knowledge that is acquired before the damage. This is in fact what happens. Patients with lesions in association areas have difficulty in recognizing faces, objects, and places in their familiar world. Indeed, lesions in different association areas give rise to specific defects in either semantic or episodic memory.

Semantic (Factual) Knowledge Is Stored in a Distributed Fashion in the Neocortex

As we have seen, semantic memory is that type of long-term memory that embraces knowledge of objects, facts, and concepts as well as words and their

meaning. It includes the naming of objects, the definitions of spoken words, and verbal fluency.

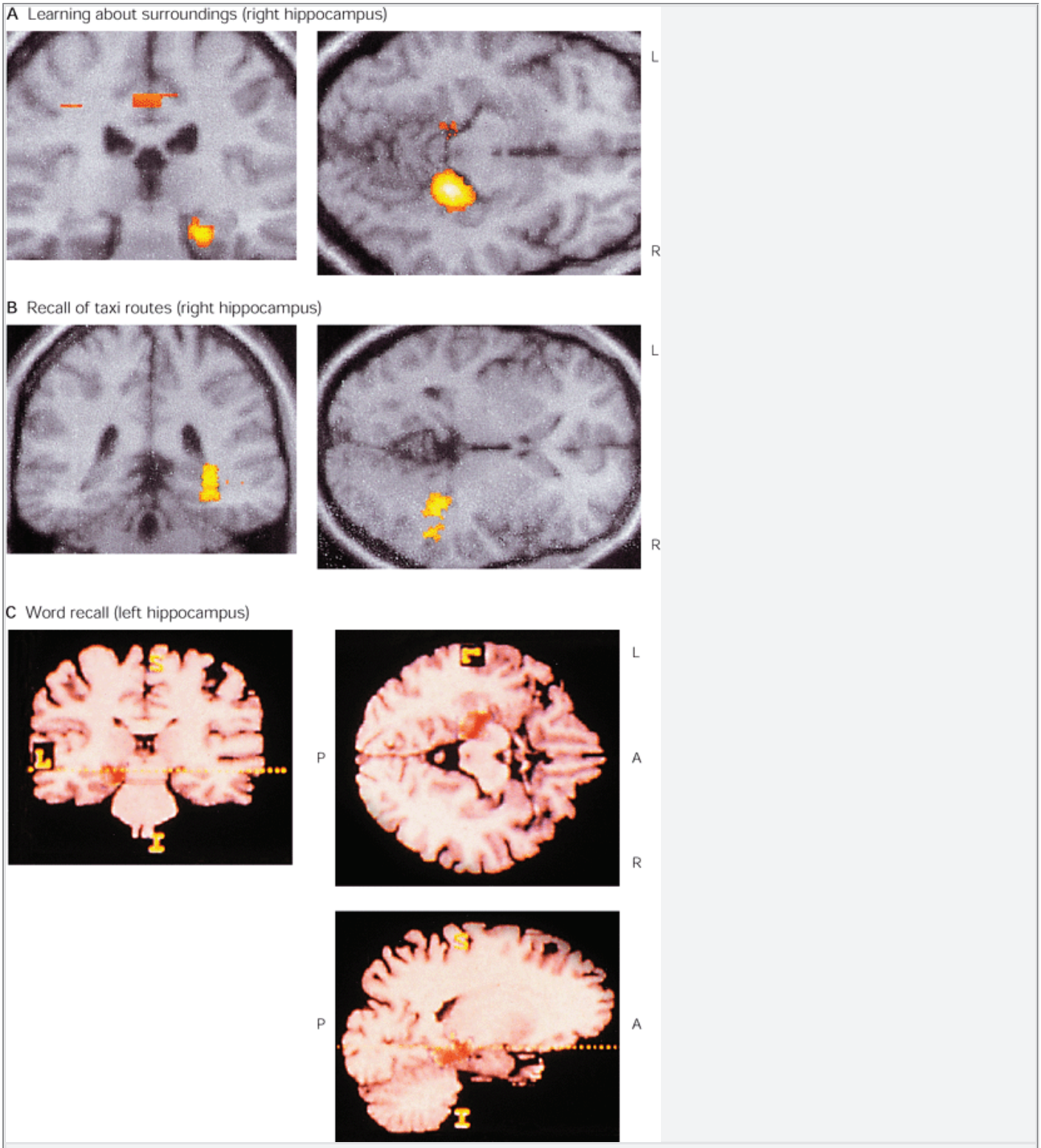


Figure 62-6 The role of the hippocampus in memory. We spend much of our time actively moving around our environment. This requires that we have a representation in our brain of the external environment, a representation that can be used to find our way around. The right hippocampus seems to be importantly involved in this representation, whereas the left hippocampus is concerned with verbal memory.

A. The right hippocampus is activated during learning about the environment. These scans were made while subjects watched a film that depicted navigation through the streets of an Irish town. The activity during this task was compared with that in the control task where the camera was static and people and cars came by it. In the latter case there was no learning of spatial relationships and the hippocampus was not activated. Areas with significant changes in activity, indexed by local perfusion change, are indicated in **yellow** and **orange.** The scan on the left is a coronal section and the scan on the right is a transaxial section; in each panel the front of the brain is on the right and the occipital lobe on the left. (From Maguire et al. 1996.)

B. The right hippocampus also is activated during the recall by licensed taxi drivers of routes around the city of London. These people spend a long time learning the intricacies of the road network in the city and are able to describe the shortest routes between landmarks as well as the names of the various streets. The right parahippocampal and hippocampal regions are significantly activated when they do this task. The scan on the left is a coronal section and the scan on the right is a transaxial section; in each panel the front of the brain is on the right and the occipital lobe on the left. Areas with significant changes in activity, indexed by local perfusion change, are depicted in **yellow** and **orange.** (From Maguire et al. 1996.)

C. Three anatomical slices in the coronal **(left upper**), transverse (**right upper**), and sagittal (**right lower**) planes show activation (**red**) in the left hippocampus associated with the successful retrieval of words from long lists that have to be memorized. **A** = anterior, **P** = posterior, **I** = inferior.

How is semantic knowledge built up? How is it stored in the cortex? The organization and flexibility of semantic knowledge is both remarkable and surprising. Consider a complex visual image such as a photograph of an elephant. Through experience this visual image becomes associated with other forms of knowledge about elephants, so that eventually when we close our eyes and conjure up the image of an elephant, the image is based on a rich representation of the concept of an elephant. The more associations we have made to the

P.1235

image of the elephant, the better we *encode* that image, and the better we can recall the features of an elephant at a future time. Furthermore, these associations fall into different categories. For example, we commonly know that an elephant is a living rather than a nonliving thing, that it is an animal rather than a plant, that it lives in a particular environment, and that it has unique physical features and behavior patterns and emits a distinctive set of sounds. Moreover, we know that elephants are used by humans to perform certain tasks and that they have a specific name. The word *elephant* is associated with all of these pieces of information, and any one bit of information can open access to all of our knowledge about elephants.

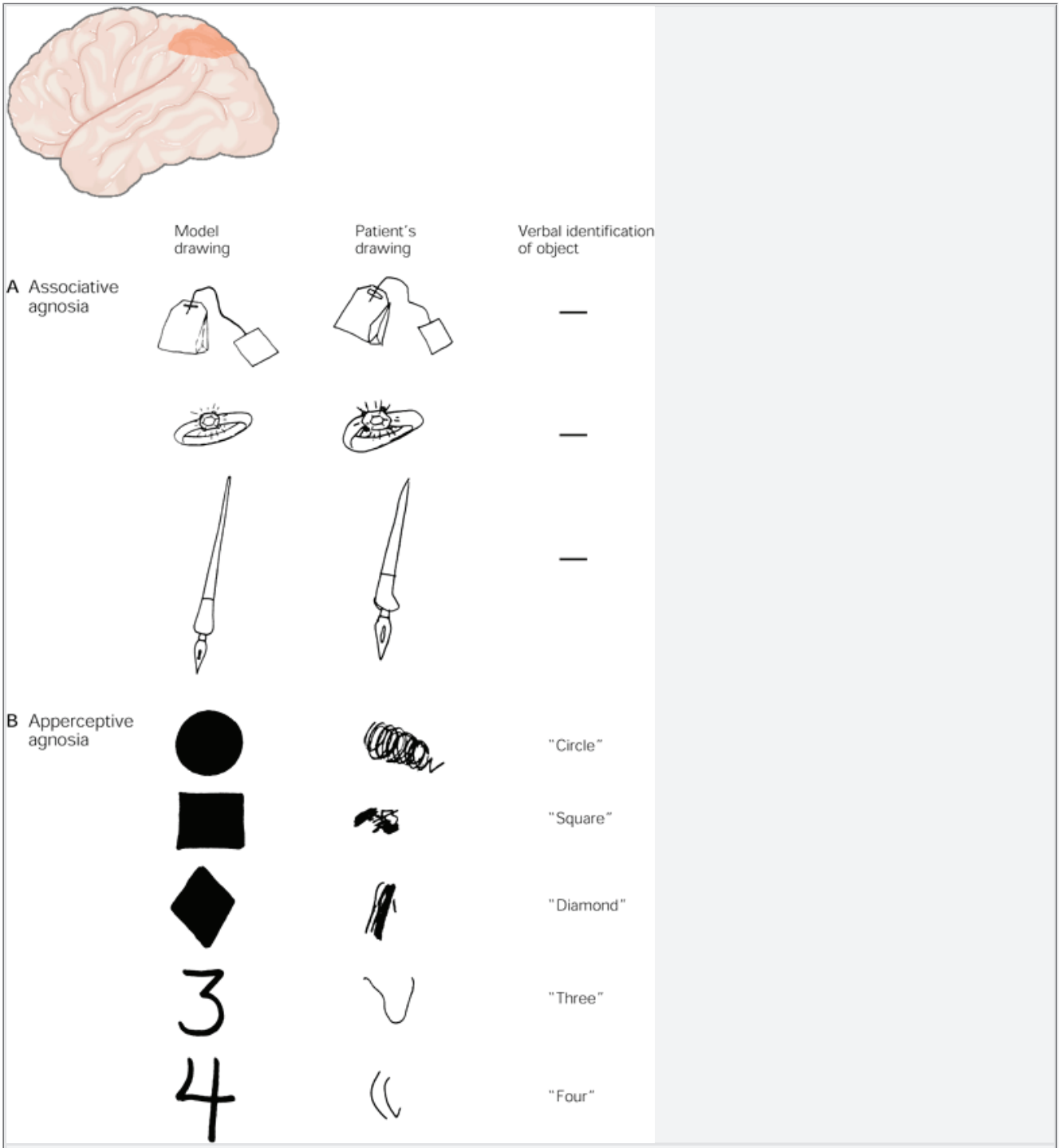


Figure 62-7 Selective lesions in the posterior parietal cortex produce selective defects in semantic knowledge. (From Farah 1994.)

A. A patient with associative agnosia is able to accurately copy drawings of a tea bag, ring, and pen but is unable to name the objects copied.

B. A patient with apperceptive agnosia is unable to reproduce even simple drawings but nevertheless can name the objects in the drawings.

As this example illustrates, we build up semantic knowledge through associations over time. The ability

P.1236

to recall and use knowledge—our *cognitive efficiency* —is thought to depend on how well these associations have organized the information we retain. As we first saw in Chapter 1, when we recall a concept it comes to mind in one smooth and continuous operation. However, studies of patients with damage to the association cortices have shown that different representations of an object—say, different aspects of elephants—are stored separately. These studies have made clear that our experience of knowledge as a seamless, orderly, and cross-referenced database is the product of integration of multiple representations in the brain at many distinct anatomical sites, each concerned with only one aspect of the concept that came to mind. Thus, there is no general semantic memory store; semantic knowledge is not stored in a single region. Rather, each time knowledge about anything is recalled, the recall is built up from distinct bits of

information, each of which is stored in specialized *(dedicated)* memory stores. As a result, damage to a specific cortical area can lead to loss of specific information and therefore a fragmentation of knowledge.

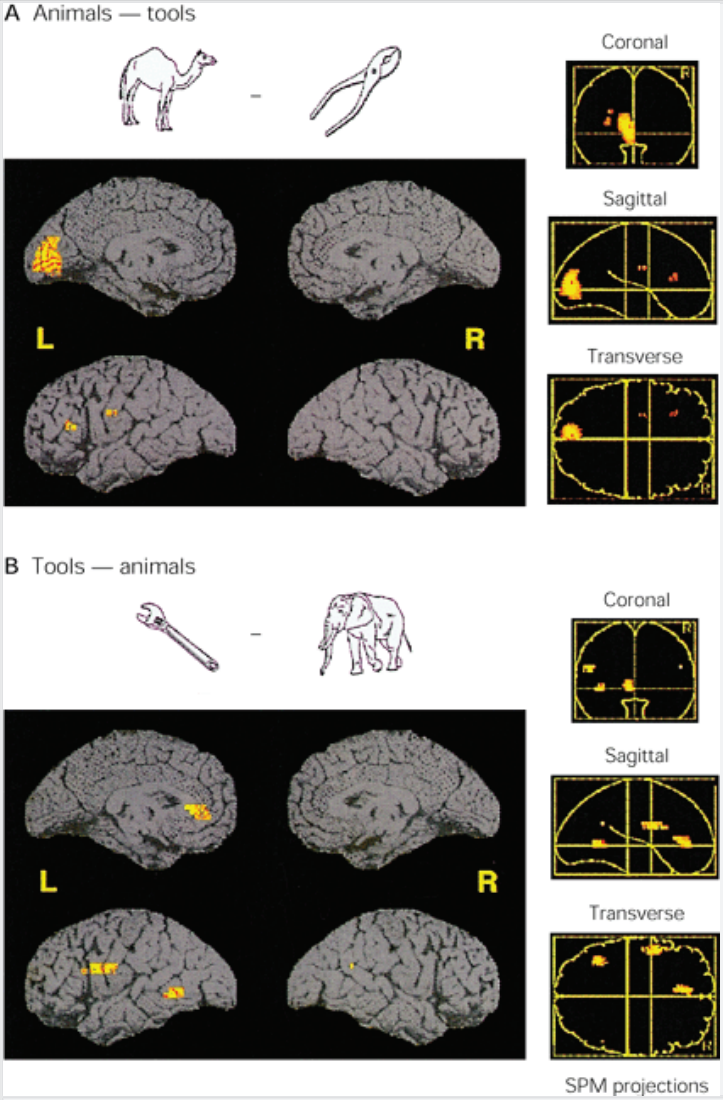


Figure 62-8 Neural correlates of category-specific knowledge are dependent on the intrinsic properties of the objects. (From Martin et al. 1996.)

A. Regions selectively activated when subjects silently named drawings of animals include the calcarine sulcus, the left and right fusiform gyri of the temporal lobe, left putamen, left thalamus, left insula and inferior frontal region (Broca's area), and right lateral and medial cerebellum. Active regions are identified by increased blood flow. In the statistical procedure regions that are active during the naming of both animals and tools cancel each other and do not show a signal. In this panel and in part B the top pair of brain images are medial views, the lower pair are lateral views. **SPM** = statistical parametric maps.

B. Regions selectively activated when subjects silently named tools are predominantly in areas of the left hemisphere associated with hand movements and the generation of action words.

For example, damage to the posterior parietal cortex can result in *associative visual agnosia;* patients cannot *name* objects but they can *identify* objects by selecting the correct drawing and can faithfully reproduce detailed drawings of the object (Figure 62-7). In contrast, damage to the occipital lobes and surrounding region can result in an *apperceptive visual agnosia;* patients are unable to draw objects but they can name them if appropriate perceptual cues are available (Figure 62-7).

While verbal and visual knowledge about objects involve different circuitry, visual knowledge involves even further specialization. For example, visual knowledge about faces and about inanimate objects is represented in different cortical areas. As we have seen in Chapter 25, lesions in the inferotemporal cortex can result in

prosopagnosia, the inability to recognize familiar faces or learn new faces, but these lesions can leave intact all other aspects of visual recognition. Conversely, positron emission tomography (PET) imaging studies show that object recognition activates the left occipitotemporal cortex and not the areas in the right hemisphere associated with face recognition. Thus, not all of our visual knowledge is represented in the same locus in the occipitotemporal cortex.

Category-specific defects in object recognition were first described by Rosaleen McCarthy and Elizabeth Warrington. They found that certain lesions interfere with the memory (knowledge) of living objects but not with memory of inanimate, manufactured objects. For example, one patient's verbal knowledge of living things was greatly impaired. When asked to define "rhinoceros" the patient responded by merely saying "animal." But when shown a picture of a rhinoceros he responded, "enormous, weighs over a ton, lives in Africa." The same patient's semantic knowledge of inanimate things was readily accessible through both verbal and visual cues. For example, when asked to define "wheelbarrow" he replied, "The thing we have here in the garden for carrying bits and pieces; wheelbarrows are used by people doing maintenance here on your buildings. They can put their cement and all sorts of things in it to carry it around."

To investigate further the neural correlates of categoryspecific knowledge for animate and inanimate objects, Leslie Ungerleider and her colleagues used PET scanning to map regions of the normal brain that are associated with naming animals and regions that are involved in naming of tools. They found that naming of animals and tools both involved bilateral activation of the ventral temporal lobes and Broca's area. In addition the naming animals selectively activated the left medial temporal lobe, a region involved in the earlier stages of visual processing. In contrast, the naming tools selectively activated a left premotor area, an area also activated with hand movements, as well as an area in the left middle temporal gyrus that is activated when action words are spoken. Thus, the brain regions active during object identification are dependent in part on the intrinsic properties of the objects presented (Figure 62-8).

Episodic (Autobiographical) Knowledge About Time and Place Seems to Involve the Prefrontal Cortex

Whereas some lesions to multimodal association areas interfere with semantic knowledge, others interfere with the capacity to recall any episodic event experienced more than a few minutes previously, including dramatic personal events such as accidents and deaths in the family that occurred before the trauma. Remarkably, patients with loss of episodic memory still have the ability to recall vast stores of factual (semantic) knowledge. One patient could remember all personal facts about his friends and famous people, such as their names and their characteristics, but could not remember any specific events involving these individuals.

The areas of the neocortex that seem to be specialized for long-term storage of episodic knowledge are the association areas of the frontal lobes. These prefrontal areas work with other areas of the neocortex to allow recollection of when and where a past event occurred (Chapter 19). A particularly striking symptom in patients with frontal lobe damage is their tendency to forget how information was acquired, a deficit called *source amnesia.* Since the ability to associate a piece of information with the time and place it was acquired is at the core of how accurately we remember the individual episodes of our lives, a deficit in source information interferes dramatically with the accuracy of recall of episodic knowledge.

Explicit Knowledge Involves at Least Four Distinct Processes

We have learned three important things about episodic and semantic knowledge. First, there is not a single, all-purpose memory store. Second, any item of knowledge has multiple representations in the brain, each of which corresponds to a different meaning and can be accessed independently (by visual, verbal, or other sensory clues). Third, both semantic and episodic knowledge are the result of at least four related but distinct types of processing: encoding, consolidation, storage, and retrieval (Figure 62-9).

Encoding refers to the processes by which newly learned information is attended to and processed when first encountered. The extent and nature of this encoding are critically important for determining how well the learned material will be remembered at later times. For a memory to persist and be well remembered, the incoming information must be encoded thoroughly and deeply. This is accomplished by attending to the information and associating it meaningfully and systematically with knowledge that is already well established in memory so as to allow one to integrate the new information with what one already knows. Memory storage is stronger when one is well motivated.

Consolidation refers to those processes that alter the newly stored and still labile information so as to make it more stable for long-term storage. As we shall learn in the next chapter, consolidation involves the expression of genes and the synthesis of new proteins, giving rise to

structural changes that store memory stably over time.

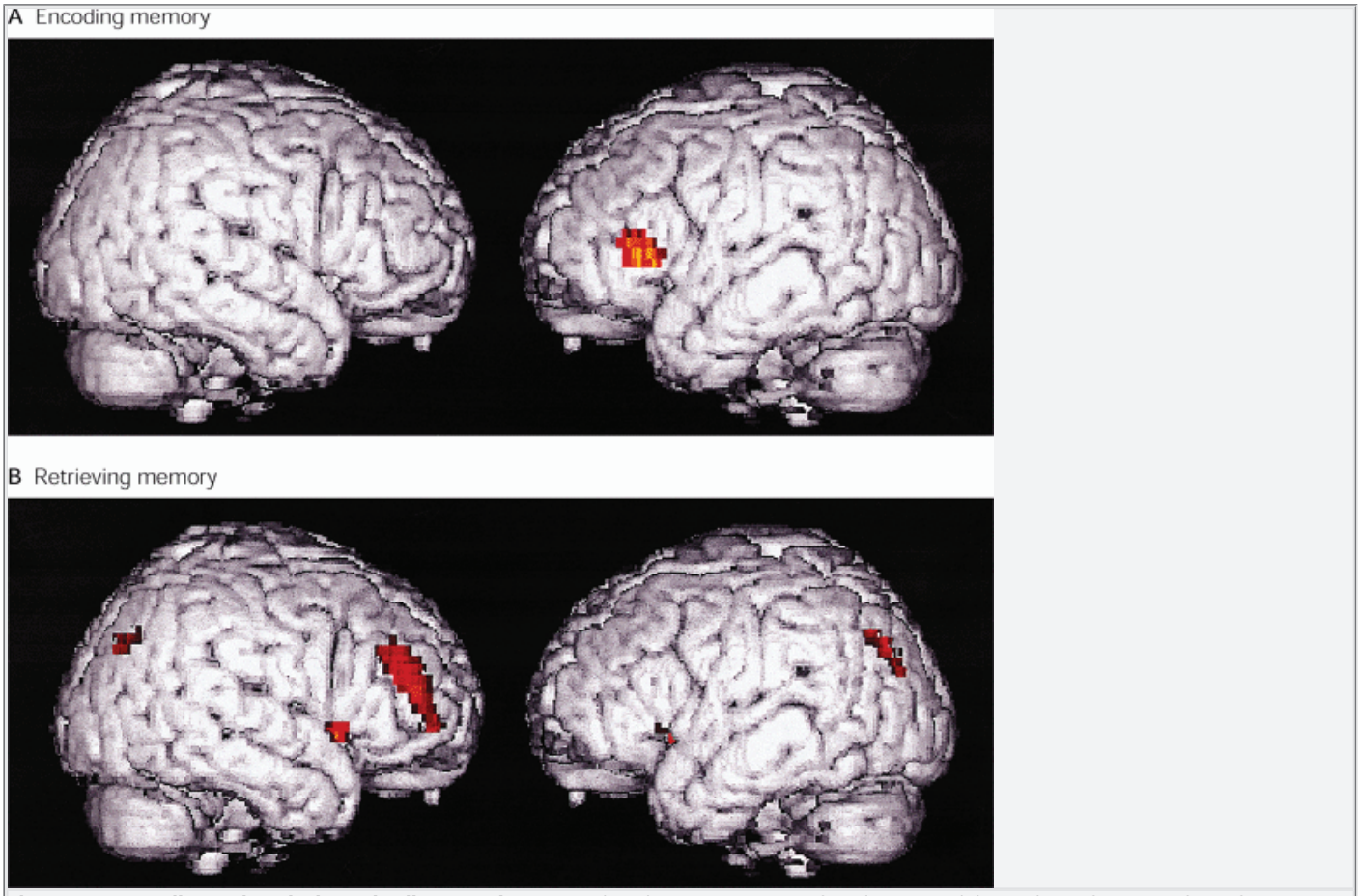


Figure 62-9 Encoding and retrieving episodic memories. Areas where brain activity is significantly increased during the performance of specific memory tasks are shown in **orange** and **red** on surface projections of the human brain (left hemisphere on the right, right hemisphere on the left).

A. Activity in the left prefrontal cortex is particularly associated with the encoding process. Subjects were scanned while attempting to memorize words paired with category labels: country—Denmark, metal—platinum, etc.

B. Activity in the right frontal cortex is associated with retrieval. Four subjects were presented with a list of category labels and examples that were not paired with the category. The subjects were then scanned when attempting to recall the examples. In addition to right frontal activation a second posterior region in the medial parietal lobe (the precuneus) is also activated.

Storage refers to the mechanism and sites by which memory is retained over time. One of the remarkable features about long-term storage is that it seems to have an almost unlimited capacity. In contrast, short-term working memory is very limited.

Finally, *retrieval* refers to those processes that permit the recall and use of the stored information. Retrieval involves bringing different kinds of information together that are stored separately in different storage sites. Retrieval of memory is much like perception; it is a constructive process and therefore subject to distortion, much as perception is subject to illusions (Box 62-1).

Retrieval of information is most effective when it occurs in the same context in which the information was acquired and in the presence of the same cues (retrieval cues) that were available to the subject during learning. Retrieval, particularly of explicit memories, is critically dependent on short-term working memory, a form of memory to which we now turn.

Working Memory Is a Short-Term Memory Required for Both the Encoding and Recall of Explicit Knowledge

How is explicit memory recalled and brought to consciousness? How do we put it to work? Both the initial encoding and the ultimate recall of explicit knowledge (and perhaps some forms of implicit knowledge as well) are thought to require recruitment of stored information

P.1239

into a special short-term memory store called *working memory*. As we learned in Chapter 19, working memory is thought to have three component systems.

Box 62-1 The Transformation of Explicit Memories

How accurate is explicit memory? This question was explored by the psychologist Frederic Bartlett in one series of studies in which the subjects were asked to read stories and then retell them. The recalled stories were shorter and more coherent than the original stories, reflecting reconstruction and condensation of the original. The subjects were unaware that they were editing the original stories and often felt more certain about the edited parts than about the unedited parts of the retold story. The subjects were not confabulating; they were merely interpreting the original material so that it made sense on recall.

Observations such as these lead us to believe that explicit memory, at least episodic (autobiographical) memory, is a constructive process like sensory perception. The information stored as explicit memory is the product of processing by our perceptual apparatus. As we saw in earlier chapters, sensory perception itself is not a faithful record of the external world but a constructive process in which incoming information is put together according to rules inherent in the brain's afferent pathways. It is also constructive in the sense that individuals interpret the external environment from the standpoint of a specific point in space as well as from the standpoint of a specific point in their own history. As discussed in Chapter 25, optical illusions nicely illustrate the difference between perception and the world as it is.

Moreover, once information is stored, later recall is not an exact copy of the information originally stored. Past experiences are used in the present as clues that help the brain reconstruct a past event. During recall we use a variety of cognitive strategies, including comparison, inferences, shrewd guesses, and suppositions, to generate a consistent and coherent memory.

An attentional control system (or *central executive*), thought to be located in the prefrontal cortex (Chapter 19), actively focuses perception on specific events in the environment. The attentional control system has a very limited capacity (less than a dozen items).

The attentional control system regulates the information flow to two rehearsal systems that are thought to maintain memory for temporary use: the articulatory loop for language and the visuospatial sketch pad for vision and action. The *articulatory loop* is a storage system with a rapidly decaying memory trace where memory for words and numbers can be maintained by subvocal speech. It is this system that allows one to hold in mind, through repetition, a new telephone number as one prepares to dial it. The *visuospatial sketch pad* represents both the visual properties and the spatial location of objects to be remembered. This system allows one to store the image of the face of a person one meets at a cocktail party.

The information processed in either one of these rehearsal, working memory systems has the possibility of entering long-term memory. The two rehearsal systems are thought to be located in different parts of the posterior association cortices. Thus, lesions of the extrastriate cortex impair rehearsal of visual imagery whereas lesions in the parietal cortex impair rehearsal of spatial imagery.

Implicit Memory Is Stored in Perceptual, Motor, and Emotional Circuits

Unlike explicit memory, implicit memory does not depend directly on conscious processes nor does recall require a conscious search of memory. This type of memory builds up slowly, through repetition over many trials, and is expressed primarily in performance, not in words. Examples of implicit memory include perceptual and motor skills and the learning of certain types of procedures and rules.

Different forms of implicit memory are acquired through different forms of learning and involve different brain regions. For example, memory acquired through fear conditioning, which has an emotional component, is thought to involve the amygdala. Memory acquired through operant conditioning requires the striatum and cerebellum. Memory acquired through classical conditioning, sensitization, and habituation—three simple forms of learning we shall consider later—involves charges in the sensory and motor systems involved in the learning.

P.1240

Implicit memory can be studied in a variety of perceptual or reflex systems in either vertebrates or invertebrates. Indeed, simple invertebrates provide useful models for studying the neural mechanisms of implicit learning.

Implicit Memory Can Be Nonassociative or Associative

Psychologists often study implicit forms of memory by exposing animals to controlled sensory experiences. Two major procedures (or paradigms) have emerged from such studies, and these have identified two major subclasses of implicit memory: nonassociative and associative. In nonassociative learning the subject learns about the properties of a single stimulus. In associative learning the subject learns about the relationship between two stimuli or between a stimulus and a behavior.

Nonassociative learning results when an animal or a person is exposed once or repeatedly to a single type of stimulus. Two forms of nonassociative learning are common in everyday life: habituation and sensitization. *Habituation* is a decrease in response to a benign stimulus when that stimulus is presented repeatedly. For example, most people are startled when they first hear the sound of a firecracker on the Fourth of July, Independence Day in the United States, but as the celebration progresses they gradually become accustomed to the noise. *Sensitization* (or *pseudoconditioning*) is an enhanced response to a wide variety of stimuli after the presentation of an intense or noxious stimulus. For example, an animal responds more vigorously to a mild tactile stimulus after it has received a painful pinch. Moreover, a sensitizing stimulus can override the effects of habituation, a process called *dishabituation.* For example, after the startle response to a noise has been reduced by habituation, one can restore the intensity of response to the noise by delivering a strong pinch.

Sensitization and dishabituation are not dependent on the relative timing of the intense and the weak stimulus; no association between the two stimuli is needed. Not all forms of nonassociative learning are as simple as habituation or sensitization. For example, imitation learning, a key factor in the acquisition of language, has no obvious associational element.

Two forms of associative learning have also been distinguished based on the experimental procedures used to establish the learning. Classical conditioning involves learning a relationship between two stimuli, whereas operant conditioning involves learning a relationship between the organism's behavior and the consequences of that behavior.

Classical Conditioning Involves Associating Two Stimuli

Since Aristotle, Western philosophers have traditionally thought that learning is achieved through the association of ideas. This concept was systematically developed by John Locke and the British empiricist school of philosophy, important forerunners of modern psychology. Classical conditioning was introduced into the study of learning at the turn of the century by the Russian physiologist Ivan Pavlov. Pavlov recognized that learning frequently consists of becoming responsive to a stimulus that originally was ineffective. By changing the appearance, timing, or number of stimuli in a tightly controlled stimulus environment and observing the changes in selected simple reflexes, Pavlov established a procedure from which reasonable inferences could be made about the relationship between changes in behavior (learning) and the environment (stimuli). According to Pavlov, what animals and humans learn when they associate *ideas* can be examined in its most elementary form by studying the association of *stimuli*.

The essence of classical conditioning is the pairing of two stimuli. The *conditioned stimulus* (CS), such as a light, tone, or tactile stimulus, is chosen because it produces either no overt response or a weak response usually unrelated to the response that eventually will be learned. The reinforcement, or *unconditioned stimulus* (US), such as food or a shock to the leg, is chosen because it normally produces a strong, consistent, overt response (the *unconditioned response*), such as salivation or withdrawal of the leg. Unconditioned responses are innate; they are produced without learning. When a CS is followed by a US, the CS will begin to elicit a new or different response called the *conditioned response.* If the US is rewarding (food or water), the conditioning is termed *appetitive;* if the US is noxious (an electrical shock), the conditioning is termed *defensive*.

One way of interpreting conditioning is that repeated pairing of the CS and US causes the CS to become an anticipatory signal for the US. With sufficient experience an animal will respond to the CS as if it were anticipating the US. For example, if a light is followed repeatedly by the presentation of meat, eventually the sight of the light itself will make the animal salivate. Thus, classical conditioning is a means by which an animal learns to predict events in the environment.

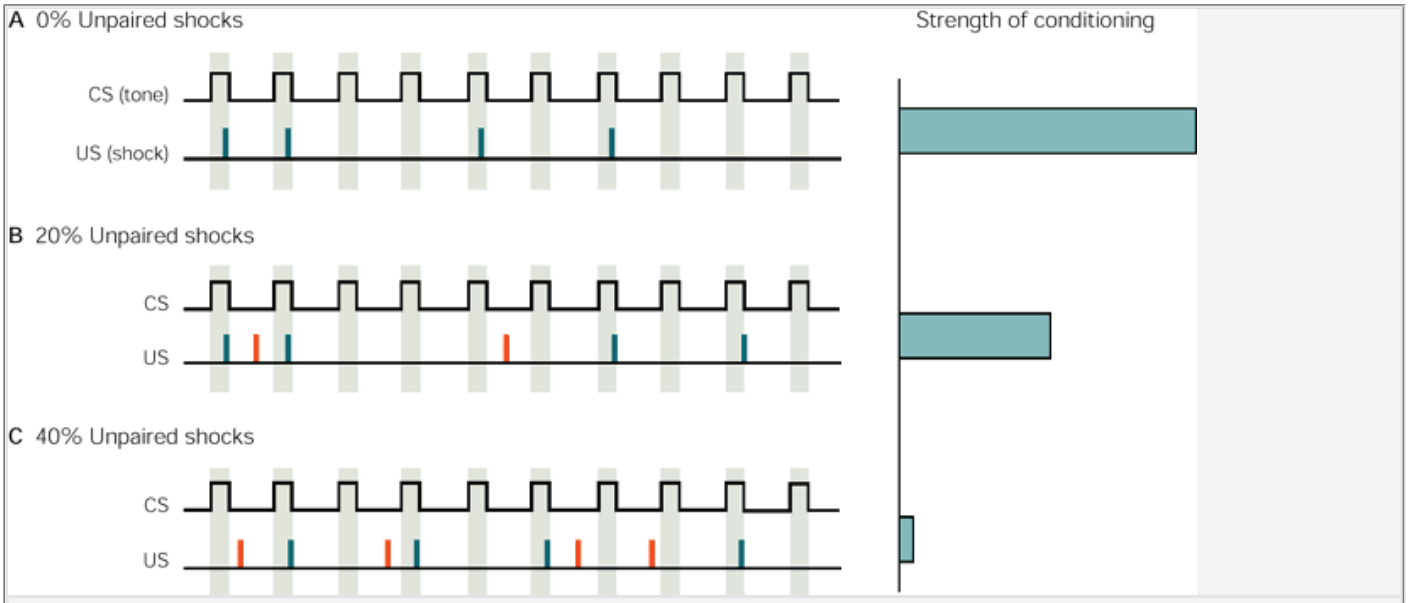


Figure 62-10 The capacity for a conditioned stimulus (CS) to produce a classically conditioned response is not a function of the number of times the CS is paired with an unconditioned stimulus (US) but rather the degree to which the CS and US are correlated. In this experiment on rats all animals were presented with a repeated tone (the CS) paired with an electric shock (the US) in 40% of the trials (**blue** vertical lines). Sometimes the shock was also presented when the tone was not present (**red** vertical lines). The percentage of these uncorrelated trials varied for different groups.

A. An experiment in which the US occurred only with the CS.

B-C. Examples in which the US was sometimes presented without the CS, either as often as the times the CS and US were paired (40%) or half as often (20%). After conditioning under the various circumstances, the degree of conditioning was evaluated by determining how effective the tone was in suppressing lever pressing to obtain food. Suppression of lever pressing is a sign of a conditioned emotional freezing response. The graph shows that in all three conditions the CS-US pairing was always 40%, but the percentage of US presentation with the absence of the CS pairing varied from 0% to 20% or 40%. When the shock occurred without the tone as often as with the tone (40%), little or no conditioning was evident. Some conditioning occurred when the shock occurred 20% of the time without the tone, and maximal conditioning occurred when the shock never occurred without the tone. (Adapted from Rescorla 1968.)

The intensity or probability of occurrence of a conditioned response decreases if the CS is repeatedly presented without the US (Figure 62-10). This process is known as *extinction.* If a light that has been paired with food is then repeatedly presented in the absence of food, it will gradually cease to evoke salivation. Extinction is an important adaptive mechanism; it would be maladaptive for an animal to continue to respond to cues in the environment that are no longer significant. The available evidence indicates that extinction is not the same as forgetting, but that instead something new is learned. Moreover, what is learned is not simply that the CS no longer precedes the US, but that the CS now signals that the US will *not* occur.

For many years psychologists thought that classical conditioning required only contiguity, that the CS precede the US by a critical minimum time interval. According to this view, each time a CS is followed by a reinforcing stimulus or US an internal connection is strengthened between the internal representation of the stimulus and the response or between one stimulus and another. The strength of the connection was thought to depend on the number of pairings of CS and US. This theory proved inadequate, however. A substantial body of empirical evidence now indicates that classical conditioning cannot be adequately explained simply by the temporal contiguity of events (Figure 62-10). Indeed, it would be maladaptive to depend solely on temporal contiguity. If animals learned to predict one type of event simply because it repeatedly occurred with another, they might often associate events in the environment that had no utility or advantage.

All animals capable of associative conditioning, from snails to humans, seem to associate events in their environment by detecting actual *contingencies* rather than simply responding to the *contiguity* of events. Why

is this faculty in humans similar to that in much simpler animals? One good reason is that all animals face common problems of adaptation and survival. Learning provides a successful solution to this problem, and once a successful biological solution has evolved it continues to be selected. Classical conditioning, and perhaps all forms of associative learning, may have evolved to enable animals to distinguish events that reliably and predictably occur together from those that are only randomly associated. In other words, the brain seems to have evolved mechanisms that can detect causal relationships in the environment, as indicated by positively correlated or associated events.

What environmental conditions might have shaped or maintained such a common learning mechanism in a wide variety of species? All animals must be able to recognize prey and avoid predators; they must search out food that is edible and nutritious and avoid food that is poisonous. Either the appropriate information can be genetically programmed into the animal's nervous system (as described in Chapter 3), or it can be acquired through learning. Genetic and developmental programming may provide the basis for the behaviors of simple organisms such as bacteria, but more complex organisms such as vertebrates must be capable of flexible learning to cope efficiently with varied or novel situations. Because of the complexity of the sensory information they process, higher-order animals must establish some degree of regularity in their interaction with the world. An effective means of doing this is to be able to detect causal or predictive relationships between stimuli, or between behavior and stimuli.

Operant Conditioning Involves Associating a Specific Behavior With a Reinforcing Event

A second major paradigm of associational learning, discovered by Edgar Thorndike and systematically studied by B. F. Skinner and others, is operant conditioning (also called *trial-and-error learning*). In a typical laboratory example of operant conditioning an investigator places a hungry rat or pigeon in a test chamber in which the animal is rewarded for a specific action. For example, the chamber may have a lever protruding from one wall. Because of previous learning as well as innate response tendencies and random activity, the animal will occasionally press the lever. If the animal promptly receives a positive reinforcer (eg, food) when it presses the level, it will subsequently press the lever more often than the spontaneous rate.

The animal can be described as having learned that among its many behaviors (for example, grooming, rearing, and walking) one behavior (lever-pressing) is followed by food. With this information the animal is likely to take the appropriate action whenever it is hungry.

If we think of classical conditioning as the formation of a predictive relationship between two stimuli (the CS and the US), operant conditioning can be considered as the formation of a predictive relationship between a stimulus (eg, food) and a behavior (eg, lever pressing). Unlike classical conditioning, which tests the responsiveness of specific reflex responses to selected stimuli, operant conditioning involves behaviors that occur either spontaneously or without an identifiable stimulus. Operant behaviors are said to be *emitted* rather than elicited; when a behavior produces favorable changes in the environment (when it is rewarded or leads to the removal of noxious stimuli) the animal tends to repeat the behavior. In general, behaviors that are rewarded tend to be repeated, whereas behaviors followed by aversive, though not necessarily painful, consequences (punishment or negative reinforcement) are usually not repeated. Many experimental psychologists feel that this simple idea, called the *law of effect*, governs much voluntary behavior.

Because operant and classical conditioning involve different kinds of association—classical conditioning involves learning an association between two stimuli whereas operant conditioning involves learning the association between a behavior and a reward—one might suppose the two forms of learning are mediated by different neural mechanisms. However, the laws of operant and classical conditioning are quite similar, suggesting that the two forms of learning may use the same neural mechanisms.

For example, timing is critical in both forms of conditioning. In operant conditioning the reinforcer usually must closely follow the operant behavior. If the reinforcer is delayed too long, only weak conditioning occurs. The optimal interval between behavior and reinforcement depends on the specific task and the species. Similarly, classical conditioning is generally poor if the interval between the conditioned and unconditioned stimuli is too long or if the unconditioned stimulus precedes the conditioned stimulus. In addition, predictive relationships are equally important in both types of learning. In classical conditioning the subject learns that a certain stimulus predicts a subsequent event; in operant conditioning the animal learns to predict the consequences of a behavior.

Associative Learning Is Not Random But Is Constrained by the Biology of the Organism

For many years it was thought that associative learning could be induced simply by pairing any two arbitrarily chosen stimuli or any response and reinforcer. More recent

studies have shown that associative learning is constrained by important biological factors.

As we have seen, animals generally learn to associate stimuli that are relevant to their survival. This feature of associative learning illustrates nicely a principle we encountered in the earlier chapters on perception. The brain is not a tabula rasa; it is capable of perceiving some stimuli and not others. As a result, it can discriminate some relations between things in the environment and not others. Thus, not all reinforcers are equally effective with all stimuli or all responses. For example, animals learn to avoid certain foods (called *bait shyness*, because animals in their natural environment learn to avoid bait foods that contain poisons). If a distinctive taste stimulus (eg, vanilla) is followed by a negative reinforcement (eg, nausea produced by a poison), an animal will quickly develop a strong aversion to the taste. Unlike most other forms of conditioning, food aversion develops even when the unconditioned response (poison-induced nausea) occurs after a long delay (up to hours) after the CS (specific taste). This makes biological sense, since the ill effects of naturally occurring toxins usually follow ingestion only after some delay.

For most species, including humans, food-aversion conditioning occurs only when taste stimuli are associated with subsequent illness, such as nausea and malaise. Food aversion develops poorly, or not at all, if the taste is followed by a nociceptive, or painful, stimulus that does not produce nausea. Conversely, an animal will not develop an aversion to a distinctive visual or auditory stimulus that has been paired with nausea. Evolutionary pressures have predisposed the brains of different species to associate certain stimuli, or a certain stimulus and a behavior, much more readily than others. Genetic and experiential factors can also modify the effectiveness of a reinforcer in one species. The results obtained with a particular class of reinforcer vary enormously among species and among individuals within a species, particularly in humans

Food aversion may be a means by which humans ordinarily learn to regulate their diets to avoid the unpleasant consequences of inappropriate food. It may also be induced in special circumstances, as in the malaise associated with certain forms of cancer chemotherapy. Aversive conditioning to foods in the ordinary diet of patients might account in part for the depressed appetite of many patients who have cancer. The nausea that follows chemotherapy for cancer can produce aversion to foods that were tasted shortly before the treatment.

Certain Forms of Implicit Memory Involve the Cerebellum and Amygdala

Lesions in several regions of the brain that are important for implicit types of learning affect simple classically conditioned responses. The best-studied case is classical conditioning of the protective eyeblink reflex in rabbits, a specific form of motor learning. A conditioned eyeblink can be established by pairing an auditory stimulus with a puff of air to the eye, which causes an eyeblink. Richard Thompson and his colleagues found that the conditioned response (eyeblink in response to a tone) can be abolished by a lesion at either of two sites. Damage to the vermis of the cerebellum, even a region as small as 2 mm^2 abolishes the conditioned response, but does not affect the unconditioned response (eyeblink in response to a puff of air). Interestingly, neurons in the same area of the cerebellum show learning-dependent increases in activity that closely parallel the development of the conditioned behavior. Second, a lesion in the interpositus nucleus, a deep cerebellar nucleus, also abolishes the conditioned eyeblink. Thus, both the vermis and the deep nuclei of the cerebellum play an important role in conditioning the eyeblink, and perhaps other simple forms of classical conditioning involving skeletal muscle movement.

Maseo Ito and his colleagues have shown that the cerebellum is involved in another form of implicit memory. The vestibulo-ocular reflex keeps the visual image fixed by moving the eyes when the head moves (Chapter 41). The speed of movement of the eyes in relation to that of the head (the gain of the reflex) is not fixed but can be modified by experience. For example, when one first wears magnifying spectacles, eye movements evoked by the vestibulo-ocular reflex are not large enough to prevent the image from moving across the retina. With experience, however, the gain of the reflex gradually increases and the eye can again track the image accurately. As with eyeblink conditioning, the learned changes in the vestibulo-ocular reflex require not only the cerebellum (the flocculus) but also one of the deep cerebellar nuclei (the vestibular) in the brain stem (see Chapters 41 and 42). Finally, as we have seen in Chapter 50, lesions of the amygdala impair conditioned fear.

Some Learned Behaviors Involve Both Implicit and Explicit Forms of Memory

Classical conditioning, we have seen, is effective in associating an unconscious reflexive response with a particular stimulus and thus typically involves implicit forms of memory. However, even this simple form of learning may also involve explicit memory, so that the learned response is mediated at least in part by cognitive processes. Consider the following experiment. A subject lays her hand, palm down, on an electrified grill;

a light (the CS) is turned on and at the same time she receives an electrical shock on one finger—she lifts her hand immediately (unconditioned response). After several light-shock conditioning trials she lifts her hand when the light alone is presented. The subject has been conditioned; but what exactly has been conditioned?

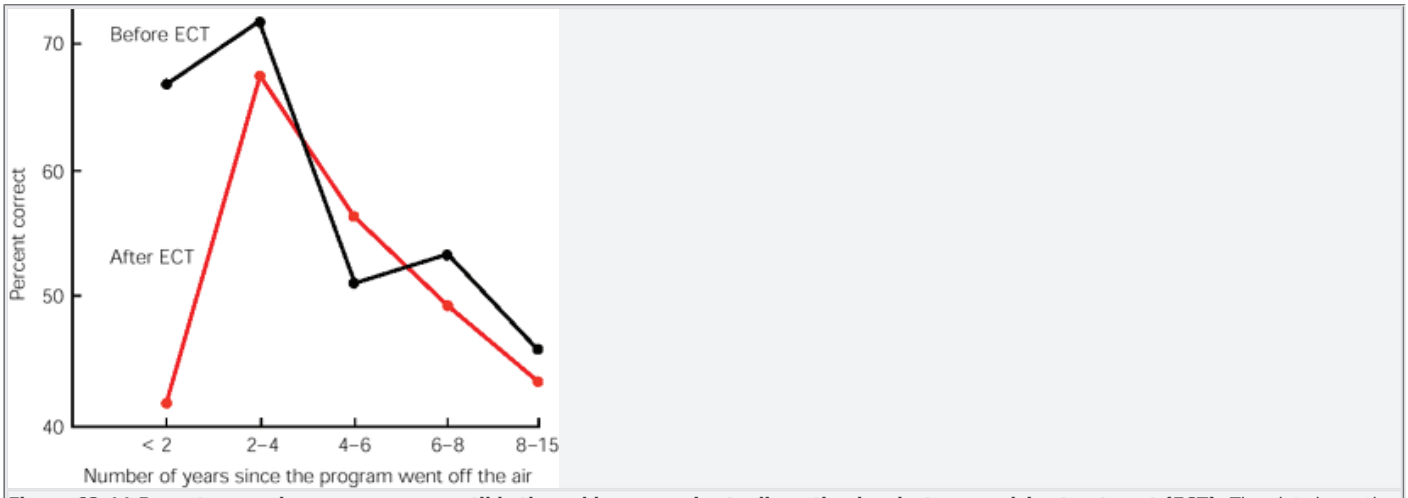


Figure 62-11 Recent memories are more susceptible than older memories to disruption by electroconvulsive treatment (ECT). The plot shows the responses of a group of patients who were tested on their ability to recall the names of television programs that were on the air during a single year between 1957 and 1972. Testing was done before and after the patients received ECT for treatment of depression. After ECT the patients showed a significant (but transitory) loss of memory for recent programs (1-2 years old) but not for older programs. (Adapted from Squire et al. 1975.)

It appears that the light is triggering a specific pattern of muscle activity (a *reflex*) that lifts the hand. However, what if the subject now places her hand on the grill, palm up, and the light is presented? If a specific reflex has been conditioned the light should produce a response that moves the hand *into* the grill. But if the subject has acquired information that the light means "grill shock," her response should be consistent with that information. In fact, the subject often will make an *adaptive response* and move her hand away from the grill. Therefore, the subject did not simply learn to apply a fixed response to a stimulus, but rather acquired information that the brain could use in shaping an appropriate response in a novel situation.

As this example makes clear, learning usually has elements of both implicit and explicit learning. For instance, learning to drive an automobile involves conscious execution of specific sequences of motor acts necessary to control the car; with experience, however, driving becomes an automatic and nonconscious motor activity. Similarly, with repeated exposure to a fact *(semantic learning*), recall of the fact with appropriate clues can eventually become virtually instantaneous—we no longer consciously and deliberately search our memory for it.

Both Explicit and Implicit Memory Are Stored in Stages

It has long been known that a person who has been knocked unconscious selectively loses memory for events that occurred before the blow (retrograde amnesia). This phenomenon has been documented thoroughly in animal studies using such traumatic agents as electroconvulsive shock, physical trauma to the brain, and drugs that depress neuronal activity or inhibit protein synthesis in the brain. Brain trauma in humans can produce particularly profound amnesia for events that occur within a few hours or, at most, days before the trauma. In such cases older memories remain relatively undisturbed. The extent of retrograde amnesia, however, varies among patients, from several seconds to several years, depending on the nature and strength of the learning and the nature and severity of the disrupting event.

Studies of memory retention and disruption of memory have supported a commonly used model of memory storage by stages. Input to the brain is processed into short-term working memory before it is transformed through one or more stages into a more permanent long-term store. A search-and-retrieval system surveys the memory store and makes information available for specific tasks.

According to this model, memory can be impaired at several points. For example, there can be a loss of the contents of a memory store; as we have seen (Chapter 58), in Alzheimer's disease there actually is a loss of nerve cells in the entorhinal cortex. Alternatively, the search-and-retrieval mechanism may be disrupted by head trauma. This latter conclusion is supported by the observation that trauma sometimes only temporarily disrupts memory, since considerable memory for past events gradually returns. If stored memory were completely lost, it obviously could not be recovered.

Studies of memory loss in patients undergoing electroconvulsive therapy (ECT) for depression have confirmed and extended the findings from animal experiments. Patients were examined using a memory test that could reliably quantify the degree of memory for relatively recent events (1-2 years old), old events (3-9 years old), and very old events (9-16 years old). The patients

were asked to identify, by voluntary recall, the names of television programs broadcast during a single year between 1957 and 1972. The patients were tested before ECT and then again afterward (with a different set of television programs). Both before and after ECT recall of the programs was more correct for more recent years. After ECT, however, the patients showed a significant but transitory loss of memory for more recent programs, while their recall of older programs remained essentially the same as it was before ECT (Figure 62-11).

One interpretation of these findings is that until memories have been converted to a long-term form, retrieval (recall) of recent memories is easily disrupted.

Once converted to a long-term form, however, the memories are relatively stable. With time, however, both the long-term memory and the capacity to retrieve it gradually diminish, even in the absence of external trauma. Because of this susceptibility to disruption, the total set of retrievable memories changes continually with time.

Several experiments studying the effects of drugs on learning support the idea that memory is time-dependent and subject to modification when it is first formed. For example, subconvulsant doses of excitant drugs, such as strychnine, can improve the retention of learning of animals even when the drug is administered after the training trials. If the drug is given to the animal soon after training, retention of learning on the following day is greater. The drug has no effect, however, when given after a long delay (several hours) after training. In contrast, inhibitors of protein synthesis selectively block the formation of long-term memory but not short-term memory when given during the training procedure.

An Overall View

The neurobiological study of memory has yielded three generalizations: memory has stages, long-term memory is represented in multiple regions throughout the nervous system, and explicit and implicit memories involve different neuronal circuits.

Different types of memory processes involve different regions and combinations of regions in the brain. Explicit memory underlies the learning of facts and experiences—knowledge that is flexible can be recalled by conscious effort and can be reported verbally. Implicit memory processes include forms of perceptual and motor memory—knowledge that is stimulus-bound, is expressed in the performance of tasks without conscious effort, and is not easily expressed verbally. Implicit memory flows automatically in the doing of things, while explicit memory must be retrieved deliberately.

Long-term storage of explicit memory requires the temporal lobe system. Implicit memory involves the cerebellum and amygdala and the specific sensory and motor systems recruited for the task being learned. Moreover, the memory processes for many types of learning involve several brain structures. For example, learned changes of the vestibulo-ocular reflex appear to involve at least two different sites in the brain, and explicit learning involves neocortical structures as well as the hippocampal formation. Furthermore, there are reasons to believe that information is represented at multiple sites even within one brain structure.

This parallel processing may explain in part why a limited lesion often does not eliminate a specific memory, even a simple implicit memory. Another important factor that may account for the failure of small lesions to adversely affect a specific memory may reside in the very nature of learning. As we shall see in the next chapter, memory involves both functional and structural changes at synapses in the circuits participating in a learning task. Although such changes are likely to occur only in particular types of neurons, the complex nature of many tasks makes it likely that these neurons are widely distributed within the pathways that mediate the response. Therefore some components of the stored information (ie, some of the synaptic changes) could remain undisturbed by a small lesion. Furthermore, the brain can take even the limited store of remaining information and construct a good representation of the original, just as the brain normally constructs conscious memory.

Selected Readings

Corkin S, Amaral DG, González RG, Johnson KA, Hyman BT. 1997. H.M.'s medial temporal lobe lesion: findings from magnetic resonance imaging. J Neurosci 17:3964–3979.

Kamin LJ. 1969. Predictability, surprise, attention, and conditioning. In: BA Campbell and RM Church (eds). *Punishment and Aversive Behavior*, pp. 279-296. New York: Appleton-Century-Crofts.

Maguire EA, Frackowiak RS, Frith CD. 1996. Learning to find your way: a role for the human hippocampal formation. Proc R Soc London B 263:1745–1750.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psych Rev 3:419–457.

Milner B. 1966. Amnesia following operation on the temporal lobes. In: CWM Whitty and OL Zangwill (eds). *Amnesia*, pp. 109-133. London: Butterworths.

Milner B, Squire LR, Kandel ER. 1998. Cognitive neuroscience and the study of memory. Neuron 20:445–468.

Muller R. 1996. A quarter of a century of place cells. Neuron 17:813–822.

Schwartz B, Robbins SJ. 1994. *Psychology of Learning and Behavior.* 4th ed. New York: Norton.

Schacter D. 1996. *Searching For Memory. The Brain, the Mind and the Past.* New York: Harper Collins/Basic Books.

Squire LR, Kandel ER. 1999. *Memory: From Mind to Molecules.* New York: Freeman.

Squire LR, Zola-Morgan S. 1991. The medial temporal lobe memory system. Science 253:1380–1386.

Steinmetz JE, Lavond DG, Ivkovich D, Logan CG, Thompson RF. 1992. Disruption of classical eyelid conditioning after cerebellar lesions: damage to a memory trace system or a simple performance deficit? J Neurosci 12:4403–4426.

References

Bartlett FC. 1932. *Remembering: a Study in Experimental and Social Psychology.* Cambridge, England: The University Press.

Blakemore C. 1977. *Mechanics of the Mind.* Cambridge, MA: Cambridge Univ. Press.

Domjan M, Burkhard B. 1986. *The Principles of Learning and Behavior*, 2nd ed. Monterey, CA: Brooks/Cole.

Drachman DA, Arbit J. 1966. Memory and the hippocampal complex II. Is memory a multiple process? Arch Neurol 15:52–61.

du Lac S, Raymond JL, Sejnowski TJ, Lisberger SG. 1995. Learning and memory in the vestibulo-ocular reflex. Annu Rev Neurosci 18:409–441.

Farah M. 1990. *Visual Agnosia.* Cambridge, MA: MIT Press.

Frackowiak RS. 1994. Functional mapping of verbal memory and language. Trends Neurosci 17:109–115.

Hebb DO. 1966. *A Textbook of Psychology.* Philadelphia: Saunders.

Martin A, Wiggs CL, Ungerleider LG, Haxby JV. 1996. Neural correlates of category-specific knowledge. Nature 379:649–652.

McCarthy, RA, Warrington EK. 1990. *Cognitive Neuropsychology: A clinical Introduction.* San Diego: Academic Press.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419–457.

McGaugh JL. 1990. Significance and remembrance: the role of neuromodulatory systems. Psychol Sci 1:15–25.

Pavlov IP. 1927. *Conditioned Reflexes: Investigation of the Physiological Activity of the Cerebral Cortex.* Anrep GV, trans. London: Oxford University Press.

Penfield W. 1958. Functional localization in temporal and deep sylvian areas. Res Publ Assoc Res Ment Dis 36:210–226.

Rescorla RA. 1968. Probability of shock in the presence and absence of CS in fear conditioning. J Comp Physiol Psychol 66:1–5.

Rescorla RA. 1988. Behavioral studies of Pavlovian conditioning. Annu Rev Neurosci 11:329–352.

Skinner BF. 1938. *The Behavior of Organisms: An Experimental Analysis.* New York: Appleton-Century-Crofts.

Squire LR. 1987. *Memory and Brain.* New York: Oxford University Press.

Squire LR, Slater PC, Chace PM. 1975. Retrograde amnesia: temporal gradient in very long term memory following electroconvulsive therapy. Science 187:77–79.

Thorndike EL. 1911. *Animal Intelligence: Experimental Studies.* New York: Macmillan.

Tulving E, Schacter DL. 1990. Priming and human memory systems. Science 247:301–306.

Warrington EK, Weiskrantz L. 1982. Amnesia: a disconnection syndrome? Neuropsychologia 20:233–248.

63

Cellular Mechanisms of Learning and the Biological Basis of Individuality

Eric R. Kandel

THROUGHOUT THIS BOOK we have emphasized that all behavior is a function of the brain and that malfunctions of the brain give rise to characteristic disturbances of behavior. Behavior, in turn, is shaped by learning. How does learning act on the brain to change behavior? How is new information acquired and, once acquired, how is it retained? In the preceding chapter we saw that memory—the outcome of learning—is not a single process but has at least two forms. *Implicit* (declarative) *memory* is unconscious memory for perceptual and motor skills, whereas *explicit* (nondeclarative) *memory* is a memory for people, places, and objects that requires conscious recall.

In this chapter we examine the cellular and molecular mechanisms that contribute to these two forms of memory by exploring the mechanisms that underlie simple implicit forms of memory storage in invertebrates and the more complex explicit forms in vertebrates. We shall see that the molecular mechanisms of memory storage are highly conserved throughout evolution, and that the more complex forms of learning and memory depend on many of the same molecular mechanisms used in the simplest forms. Finally, we shall consider the idea that these mechanisms contribute to individuality by changing the connectivity of neurons in our brains.

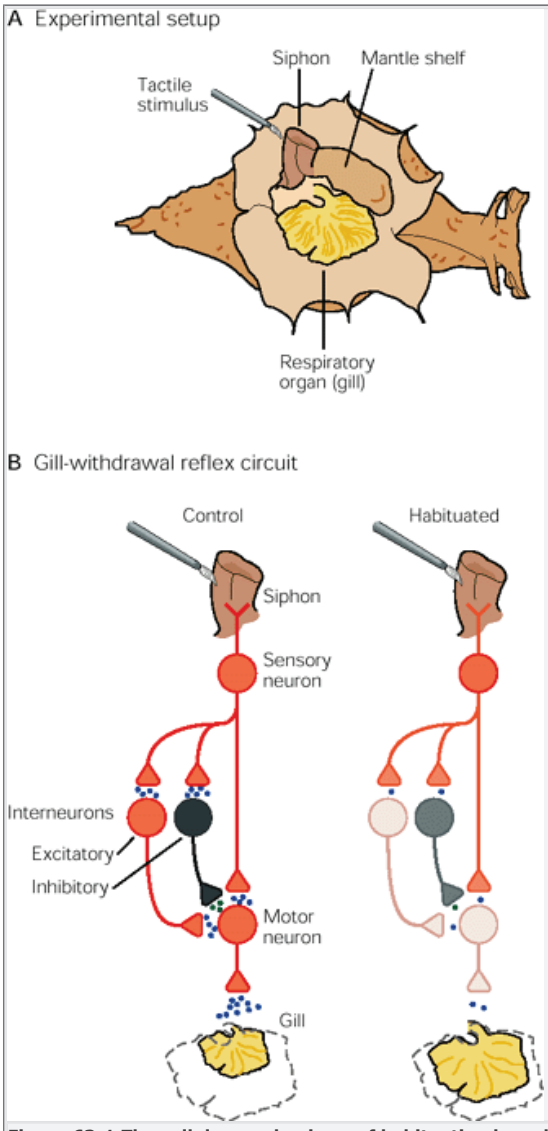


Figure 63-1 The cellular mechanisms of habituation have been investigated in the gill-withdrawal reflex of the marine snail*Aplysia*.

A. A dorsal view of *Aplysia* illustrates the respiratory organ (gill), which is normally covered by the mantle shelf. The mantle shelf ends in the siphon, a fleshy spout used to expel seawater and waste. A tactile stimulus to the siphon elicits the gill-withdrawal reflex. Repeated stimuli lead to habituation.

B. This simplified circuit shows key elements involved in the gill-withdrawal reflex as well as sites involved in habituation. In this circuit about 24 mechanoreceptors in the abdominal ganglion innervate the siphon skin. These glutaminergic sensory cells form synapses with a cluster of six motor neurons that innervate the gill and with several groups of excitatory and inhibitory interneurons that synapse on the motor neurons. (For simplicity, only one of each type of neuron is illustrated here.) Repeated stimulation of the siphon leads to a depression of synaptic transmission between the sensory and motor neurons as well as between certain interneurons and the motor cells.

Short-Term Storage of Implicit Memory for Simple Forms of Learning Results From Changes in the Effectiveness of Synaptic Transmission

Much progress in the cellular study of memory storage has come from examining elementary forms of learning: habituation, sensitization, and classical conditioning. These cellular modifications have been analyzed in the behavior of simple invertebrates and in a variety of vertebrate reflexes, such as flexion reflexes, fear responses, and the eyeblink. Most simple forms of implicit learning change the effectiveness of the synaptic connections that make up the pathway mediating the behavior.

Habituation Involves an Activity-Dependent Presynaptic Depression of Synaptic Transmission

In *habituation*, the simplest form of implicit learning, an animal learns about the properties of a novel stimulus that is harmless. An animal first responds to a new stimulus by attending to it with a series of orienting responses. If the stimulus is neither beneficial nor harmful, the animal learns, after repeated exposure, to ignore it.

Habituation was first investigated by Ivan Pavlov and Charles Sherrington. While studying posture and locomotion, Sherrington observed a decrease in the intensity of certain reflexes, such as the withdrawal of a limb, in response to repeated stimulation. The reflex response returned only after many seconds of rest. He suggested that this decrease, which he called habituation, results from diminished synaptic effectiveness within the pathways to the motor neurons that had been repeatedly activated.

This problem was later investigated at the cellular level by Alden Spencer and Richard Thompson. They found close cellular and behavioral parallels between habituation of the spinal flexion reflex in the cat and habituation of more complex behavioral responses in humans. They showed, through intracellular recordings from spinal motor neurons in cats, that habituation leads to a decrease in the strength of the synaptic connections between excitatory interneurons and motor neurons. The connections between the sensory neurons innervating the skin and the interneurons were unaffected.

Since the organization of interneurons in the spinal cord of vertebrates is quite complex, further analysis of the cellular mechanisms of habituation in the flexion reflex proved difficult. Progress in this effort required a simpler system. The marine sea slug *Aplysia californica*, which has a simple nervous system containing only about 20,000 central nerve cells, is an excellent simple system for studying implicit forms of memory. *Aplysia*

has a repertory of defensive reflexes for withdrawing its gill and its siphon, a small fleshy spout above the gill used to expel seawater and waste (Figure 63-1A). These reflexes are similar to the leg withdrawal reflex studied by Spencer and Thompson. For example, a mild tactile stimulus delivered to the siphon elicits reflex withdrawal of both siphon and gill. With repeated stimulation these reflexes habituate. They can also be sensitized and classically conditioned, as we shall see later.

Gill withdrawal in *Aplysia* has been studied in detail. In response to a novel tactile stimulus to the siphon, firing in the sensory neurons innervating the siphon generates excitatory synaptic potentials in interneurons and motor cells (Figure 63-1B). The synaptic potentials from the sensory neurons and interneurons summate both temporally and spatially to cause the motor cells to discharge repeatedly, leading to strong reflexive withdrawal of the gill. If the stimulus is repeated, the direct monosynaptic excitatory synaptic potentials produced by sensory neurons in both the interneurons and the motor cells become progressively smaller. Thus, with repeated stimulation, several of the excitatory interneurons also produce weaker synaptic potentials in the motor neurons, with the net result that the motor neuron fires much less briskly and consequently the reflex response diminishes.

What reduces the effectiveness of synaptic transmission by the sensory neurons? Quantal analysis revealed that the decrease in synaptic strength results from a decrease in the number of transmitter vesicles released from presynaptic terminals of sensory neurons. These sensory neurons use glutamate as their transmitter. Glutamate interacts with two types of receptors in motor cells: one similar to the *N* -methyl-d-aspartate (NMDA) type of glutamate receptors of vertebrates and the other to a non-NMDA-type (Chapter 12). There is no change in the sensitivity of these receptors with habituation. How this decrease in transmitter release occurs is not yet understood; it is thought to be due in part to a reduced mobilization of transmitter vesicles to the active zone (see Chapter 14). This reduction lasts many minutes.

These enduring *plastic* changes in the functional strength of synaptic connections thus constitute the cellular mechanisms mediating the short-term memory for habituation. Since these changes occur at several sites in the reflex circuit, memory in this instance is distributed and stored throughout the circuit, not at one specialized site. Synaptic depression of the connections made by sensory neurons, interneurons, or both is a common mechanism for habituation and explains habituation of the several well-studied escape responses of crayfish and cockroaches as well as startle reflexes of vertebrates.

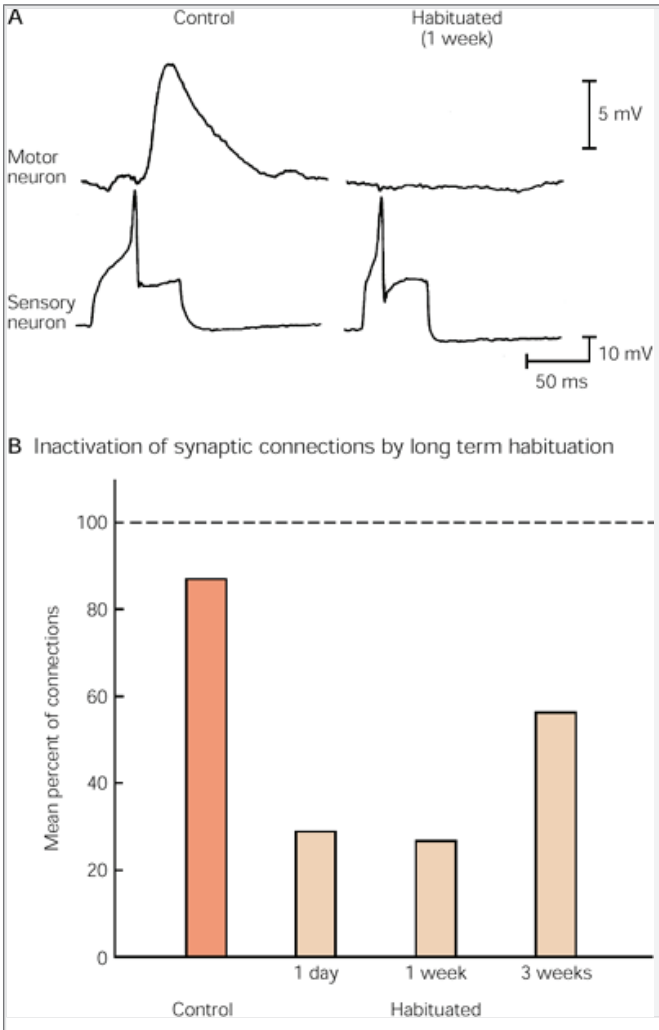


Figure 63-2 Long-term habituation of the gill-withdrawal reflex in *Aplysia* is represented on the cellular level by a dramatic depression of synaptic effectiveness between the sensory and motor neurons. (Adapted from Castellucci et al. 1978.)

A. Comparison of the synaptic potentials in a sensory and motor neuron in a control (untrained) animal and an animal that has been subjected to long-term habituation. In the habituated animal no synaptic potential is evident in the motor neuron one week after training.

B. The mean percentage of physiologically detectable connections in habituated animals at three points in time after long-term habituation training.

The synaptic mechanisms underlying habituation can vary in two ways. First, the locus of the depression can be situated at any of several synaptic sites. For example, in the flexion reflex of the spinal cord there is no depression of synaptic transmission at the direct connections made by the sensory neurons on interneurons.

P.1250

Rather, depression is thought to occur at downstream sites: at the synapses made by certain classes of interneurons on the motor neurons of the reflex. Second, mechanisms other than homosynaptic depression, such as enhancement of synaptic inhibition, can contribute to habituation.

These studies illustrate that learning can lead to changes in synaptic strength and that the duration of the short-term memory storage is determined by the duration of the synaptic change. In turn, theses studies raise the question: How much can the effectiveness of a synapse change and how long can the change last? Whereas a single session of 10 tactile stimuli to the siphon leads to a short-term memory for habituation of the *Aplysia* gill-withdrawal reflex lasting minutes, four such sessions, separated by periods ranging from several hours to one day, produce a long-term memory that lasts for as long as three weeks! In naive *Aplysia* 90% of the sensory neurons make physiologically detectable connections onto identified gill motor neurons. In contrast, in animals trained for long-term habituation the number of such connections is reduced to 30%; this low incidence persists for a week and does not completely recover for three weeks after the training (Figure 63-2). As we shall see later, during this long-term inactivation of synaptic transmission the actual structure of the sensory neurons changes.

Not all synapses are equally adaptable. The strength of some synapses in *Aplysia* rarely changes, even with repeated activation. However, at synapses specifically involved in learning and memory storage, such as the connections between sensory and motor neurons and some interneurons in the withdrawal-reflex circuit, a relatively small amount of training, especially if it is appropriately spaced over many minutes or hours, can produce large and enduring changes in synaptic strength. Massed training, whereby the habituating stimuli are given one after the other without rest between training sessions, produces a robust short-term memory but long-term memory is seriously compromised. This illustrates a general principle of learning psychology: Spaced training is usually much more effective than massed training in producing long-term memory.

Sensitization Involves Presynaptic Facilitation of Synaptic Transmission

When an animal repeatedly encounters a harmless stimulus it learns to habituate to it. In contrast, with a *harmful* stimulus the animal typically learns to respond more vigorously not only to that stimulus but also to other stimuli, even harmless ones. Defensive reflexes for withdrawal and escape become heightened. This enhancement of reflex responses, called *sensitization*, is more complex than habituation: A stimulus applied to one pathway produces a change in the reflex strength in another pathway. Like habituation, sensitization has both a short-term and a long-term form. Thus, whereas a single shock to an animal's tail produces short-term sensitization that lasts minutes, five or more shocks to the tail produce sensitization lasting days to weeks.

A noxious stimulus to the tail enhances synaptic transmission at several connections in the neural circuit of the gill-withdrawal reflex, including the connections made by sensory neurons with motor neurons and interneurons—the same synapses depressed by habituation. Thus a synapse can participate in more than one type of learning and store more than one type of memory. However, habituation and sensitization use different cellular mechanisms to produce synaptic change. Short-term habituation in *Aplysia* is a *homosynaptic process;* the decrease in synaptic strength is a direct result of activity in the sensory neurons and their central connections in the reflex pathway. In contrast, sensitization is a *heterosynaptic* process; the enhancement of

P.1251

P.1252

synaptic strength is induced by modulatory interneurons activated by stimulation of the tail.



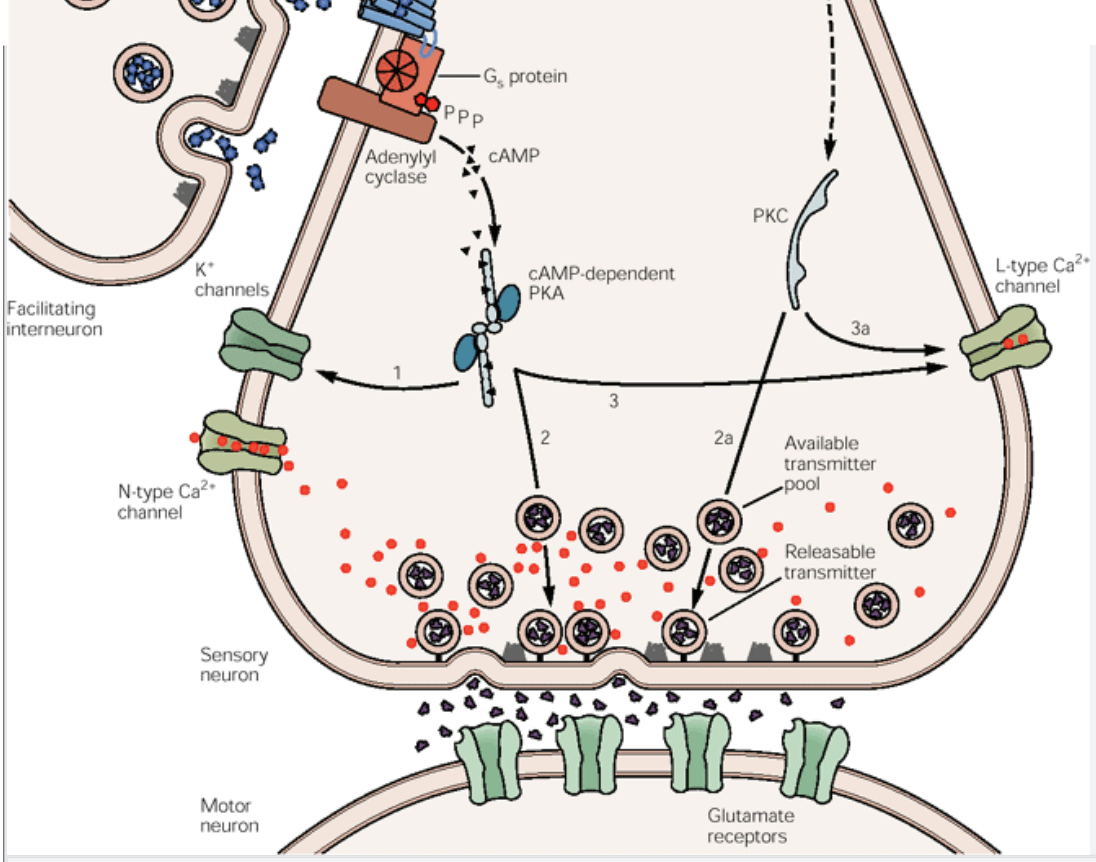


Figure 63-3 (Opposite) Short-term sensitization of the gill-withdrawal reflex in*Aplysia* involves presynaptic facilitation.

A. Sensitization of the gill is produced by applying a noxious stimulus to another part of the body, such as the tail. Stimuli to the tail activate sensory neurons in the tail that excite facilitating interneurons, which form synapses on the terminals of the sensory neurons innervating the siphon. At these axoaxonic synapses the facilitating interneurons enhance transmitter release from the sensory neurons (*presynaptic facilitation*).

B. Presynaptic facilitation in the sensory neuron is thought to occur by means of three biochemical pathways. The transmitter released by the presynaptic interneuron, here serotonin (**5-HT**, hydroxytryptamine), binds to two receptors. One engages a G protein ($\mathbf{G_s}$), which increases the activity of adenylyl cyclase. The adenylyl cyclase converts adenosine triphosphate (ATP) to cyclic adenosine monophosphate (**cAMP**), thereby increasing the level of cAMP in the terminal of the sensory neuron. The cAMP activates the cAMP-dependent protein kinase A (**PKA**) by attaching to its inhibitory regulatory subunit, thus releasing its active catalytic subunit.

The catalytic subunit of PKA acts along three pathways. In **pathway 1** the catalytic subunit phosphorylates K^+ channels, thereby decreasing the K^+ current. This prolongs the action potential and increases the influx of Ca^{2+}, thus augmenting transmitter release. In **pathway 2** vesicles containing transmitter are mobilized to the releasable transmitter pool at the active zone, and the efficiency of the exocytotic release machinery is also enhanced. In **pathway 3** L-type Ca^{2+} channels are opened. Serotonin, acting through a second receptor, engages the G protein ($\mathbf{G_o}$) that activates a phospholipase C (**PLC**), which in turn stimulates intramembranous diacylglycerol to activate protein kinase C (**PKC**). **Pathways 2-2a** and **3-3a** involve the joint action of PKA and PKC.

There are at least three groups of modulatory inter-neurons, the best studied of which release serotonin. The serotonergic and other modulatory interneurons form synapses on the sensory neurons, including axo-axonic synapses on their presynaptic terminals (Figure 63-3A). The serotonin and other modulatory transmitters released from the interneurons after a single shock to the tail bind to specific membrane-spanning receptors that activate the heterotrimeric GTP binding protein Gαs. The Gαs protein activates an adenylyl cyclase to produce the second messenger cyclic adenosine mono-phosphate (cAMP), which activates the cAMP-dependent protein kinase (PKA) (see the discussion of PKA in Chapter 13). PKA, together with protein kinase C, enhances release of transmitter from the sensory neurons' terminals for a period of minutes through the phosphorylation of several substrate proteins (Figure 63-3B). As we shall learn later, repeated sensitizing stimuli produce a strengthening of connections that lasts days.

Classical Conditioning Involves Presynaptic Facilitation of Synaptic Transmission That Is Dependent on Activity in Both the Presynaptic and the Postsynaptic Cell

Classical conditioning is a more complex form of learning than sensitization. Rather than learning only about one stimulus, the organism learns to associate one type of stimulus with another. As we have learned in Chapter 62, an initially weak conditioned stimulus can become highly effective in producing a response when paired with a strong unconditioned stimulus. In reflexes that can be enhanced by both classical conditioning and sensitization, classical conditioning results in a greater and longer-lasting enhancement.

The siphon- and gill-withdrawal reflexes of *Aplysia* are examples of reflexes that can be enhanced by both classical conditioning and sensitization. The gill-withdrawal reflex can be elicited in one of two ways: by stimulating either the siphon or a nearby structure called the *mantle shelf.* The siphon and the mantle shelf are separately innervated by distinct populations of sensory neurons. Thus, each reflex pathway can be conditioned independently by pairing a conditioned stimulus to the appropriate area (either the siphon or the mantle shelf) with an unconditioned stimulus (a strong shock to the tail). After such paired or associative training, the response of the conditioned (or paired) pathway to stimulation is significantly enhanced compared to that of the unpaired pathway (Figure 63-4).

In classical conditioning the timing of the conditioned and unconditioned stimuli is critical. The conditioned stimulus must *precede* the unconditioned stimulus, often within an interval of about 0.5 s. What cellular mechanisms are responsible for this requirement for temporal pairing of stimuli? In classical conditioning of the gill-withdrawal reflex of *Aplysia* one important feature is the timing of the convergence in individual sensory neurons of the conditioned stimulus (siphon touch) and the unconditioned stimulus (tail shock).

As we have seen, an unconditioned stimulus to the tail activates facilitating interneurons that make axo-axonic connections with the presynaptic terminals of the sensory neurons that carry information from the siphon and the mantle shelf (Figure 63-4A). The resulting presynaptic facilitation ordinarily gives rise to behavioral sensitization. However, if the unconditioned stimulus (to the tail) and the conditioned stimulus (to the siphon or mantle shelf) are timed so that the conditioned stimulus just precedes the unconditioned stimulus, then the modulatory interneurons engaged by the unconditional stimulus will activate the sensory neurons immediately *after* the conditioned stimulus has activated the sensory neurons. This sequential activation of the sensory neurons during a critical interval by the CS and the US leads to greater presynaptic facilitation than when the two stimuli are not appropriately paired (Figure 63-4B). This novel feature unique to classical conditioning is called activity dependence.

There are presynaptic and postsynaptic components to activity-dependent facilitation. Activity in the conditioned stimulus pathway leads to Ca^{2+} influx into the presynaptic sensory neuron with each action potential, and this influx activates the Ca^{2+}-binding protein calmodulin. The activated Ca^{2+}/calmodulin binds to adenylyl cyclase, potentiating its response to serotonin and enhancing its production of cAMP. Thus, the presynaptic cellular mechanism of classical conditioning in the monosynaptic pathway of the withdrawal reflex in *Aplysia* is in part an elaboration of the mechanism of sensitization in this same pathway. This is because adenylyl cyclase acts as a *coincidence detector.* That is, it recognizes the molecular representation of both the conditioned stimulus (spike activity in the sensory neuron and the consequent Ca^{2+} influx) and the unconditioned stimulus (serotonin released by tail stimuli), and it responds both to the conditioned stimulus (binding to the Ca^{2+}/calmodulin activated by the Ca^{2+}/influx following action potentials) and the unconditioned stimulus (binding to the Gαs activated by the binding of serotonin to a receptor).

The postsynaptic component of classical conditioning is a retrograde signal to the sensory neuron. In the withdrawal reflex pathway in *Aplysia* the postsynaptic motor cell has two types of receptors to glutamate: nonNMDA and NMDA-type receptors. As we have learned in Chapter 11, the extracellular mouth of the NMDA-type

receptor-channel is plugged by Mg^{2+} at the resting membrane potential. Thus, under normal circumstances and during habituation and sensitization only the non-NMDA receptor is activated because the NMDA receptor is blocked by Mg^{2+}. However, when the conditioned stimulus and unconditioned stimulus are paired appropriately during classical conditioning, the motor neuron fires a whole train of action potentials. The depolarization of the postsynaptic cell expels Mg^{2+} from the NMDA-type receptor-channel and Ca^{2+} flows into the cell. The Ca^{2+} influx is thought to activate signaling pathways in the motor cell that give rise to a retrograde messenger that is taken up in the presynaptic terminals of the sensory cell, where it acts to enhance transmitter release even further.

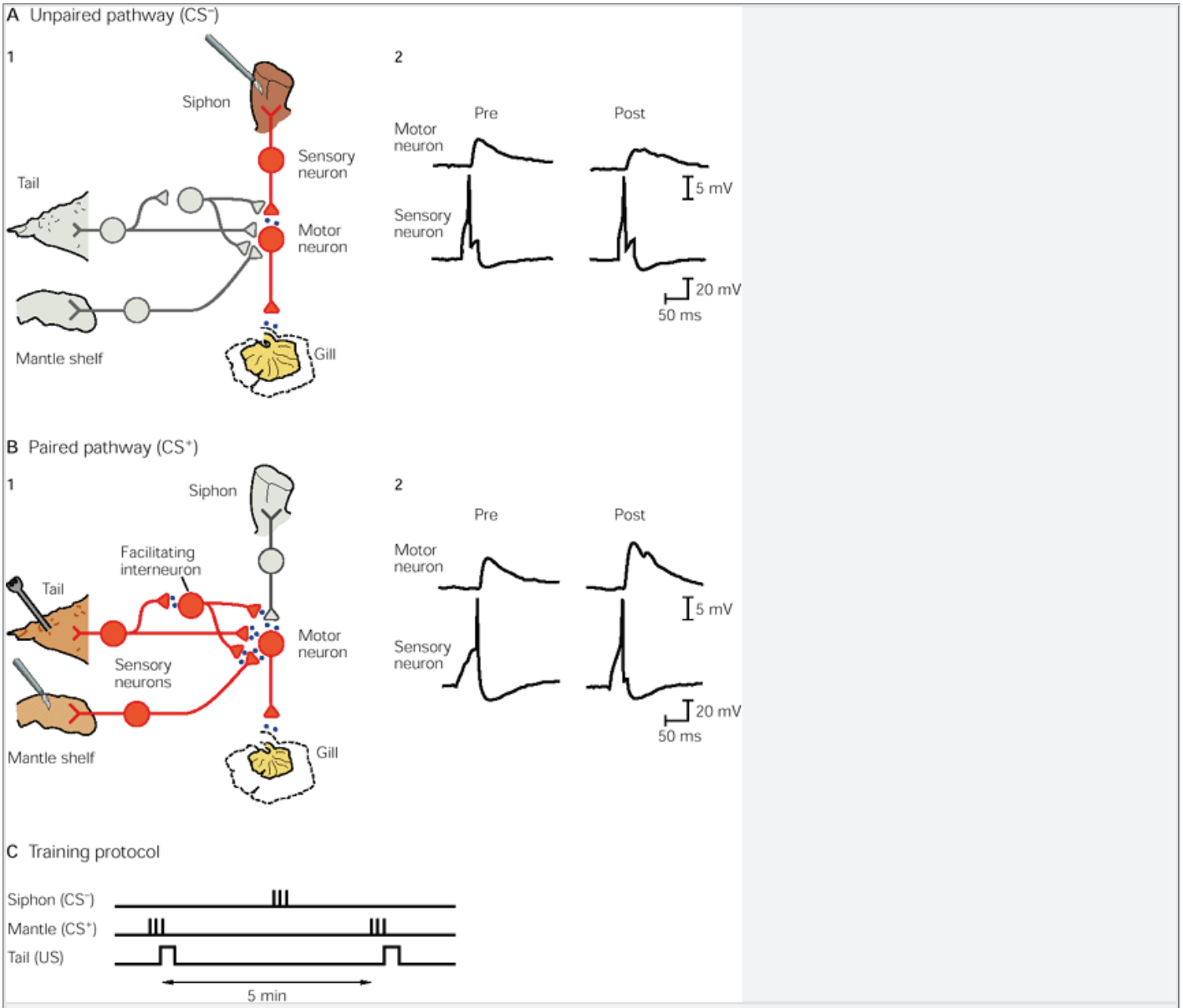


Figure 63-4 Classical conditioning of the gill-withdrawal reflex in *Aplysia*. A conditioned stimulus (**CS**) applied to the mantle shelf is paired with an unconditioned stimulus (**US**) applied to the tail. As a control, a CS applied to the siphon is not paired with the US. (Adapted from Hawkins et al. 1983.)

A. A shock to the tail (US) excites facilitating interneurons that form synapses on the presynaptic terminals of sensory neurons innervating the mantle shelf and siphon. This is the mechanism of sensitization (A1).

B. When the mantle pathway is activated by a CS just before the US, the action potentials in the mantle sensory neurons prime them so that they are more responsive to subsequent stimulation from the (serotonergic) facilitating interneurons in the US pathway. This is the mechanism of classical conditioning; it both amplifies the response of the CS pathway and restricts the amplification to that pathway (B1).

Recordings of the excitatory postsynaptic potentials produced in an identified motor neuron by the mantle and siphon sensory neurons were made before training (**Pre**) and 1 hour after training (**Post**). After training the excitatory postsynaptic potential due to input from the mantle (paired) sensory neuron (B2) is considerably greater than the one due to the siphon (unpaired) neuron (A2).

C. The experimental protocol for classical conditioning compares the responses of paired and unpaired stimuli mediated by siphon and mantle sensory neurons. In the mantle sensory neurons the action potentials produced by the CS are paired with those produced by the US (tail stimulus). In a siphon sensory neuron the action potentials produced by the CS are not paired with the same US.

Thus, three signals in a siphon sensory neuron must converge to produce the large increase in transmitter release that occurs with classical conditioning: (1) activation

of adenylyl cyclase by Ca^{2+} influx, representing the conditioned stimulus; (2) activation of serotonergic

P.1254

receptors coupled to adenylyl cyclase, representing the unconditioned stimulus; and (3) a retrograde signal indicating that the postsynaptic cell has been adequately activated by the uncondtioned stimulus.

Long-Term Storage of Implicit Memory for Sensitization and Classical Conditioning Involves the cAMP-PKA-MAPK-CREB Pathway

Molecular Biological Analysis of Long-Term Sensitization Reveals a Role for cAMP Signaling in Long-Term Memory

As with habituation and most other forms of learning, practice makes perfect. Repeated experience consolidates memory by converting the short-term form into a long-term form. These physiological consequences of repeated training have been best studied for sensitization. In *Aplysia* a single training session (or a single application of serotonin to the sensory neurons) gives rise to short-term sensitization, lasting only minutes, that does not require new protein synthesis. However, five training sessions produce long-term sensitization, lasting several days, that requires new protein synthesis. Further spaced training produces sensitization that persists for weeks. These behavioral studies of *Aplysia* (and similar ones in vertebrates) suggest that short-term and long-term memory are two independent but overlapping processes that blend into one another. Several findings point to this interpretation.

First, both short- and long-term memory for sensitization of the gill-withdrawal reflexes involve changes in the strength of connections at several synaptic sites, including the synaptic connections between sensory and motor neurons (Figure 63-5A). Second, in both the long-term and short-term processes the increase in the synaptic strength of the connections between the sensory and motor neurons is due to the enhanced release of transmitter. Third, the same transmitter (serotonin) released by stimulation of the tail produces short-term facilitation after a single exposure and long-term facilitation after five or more repeated exposures. Finally, cAMP and PKA, intracellular second-messenger pathways that are critically involved in short-term memory, are also recruited for long-term memory (Figure 63-5B).

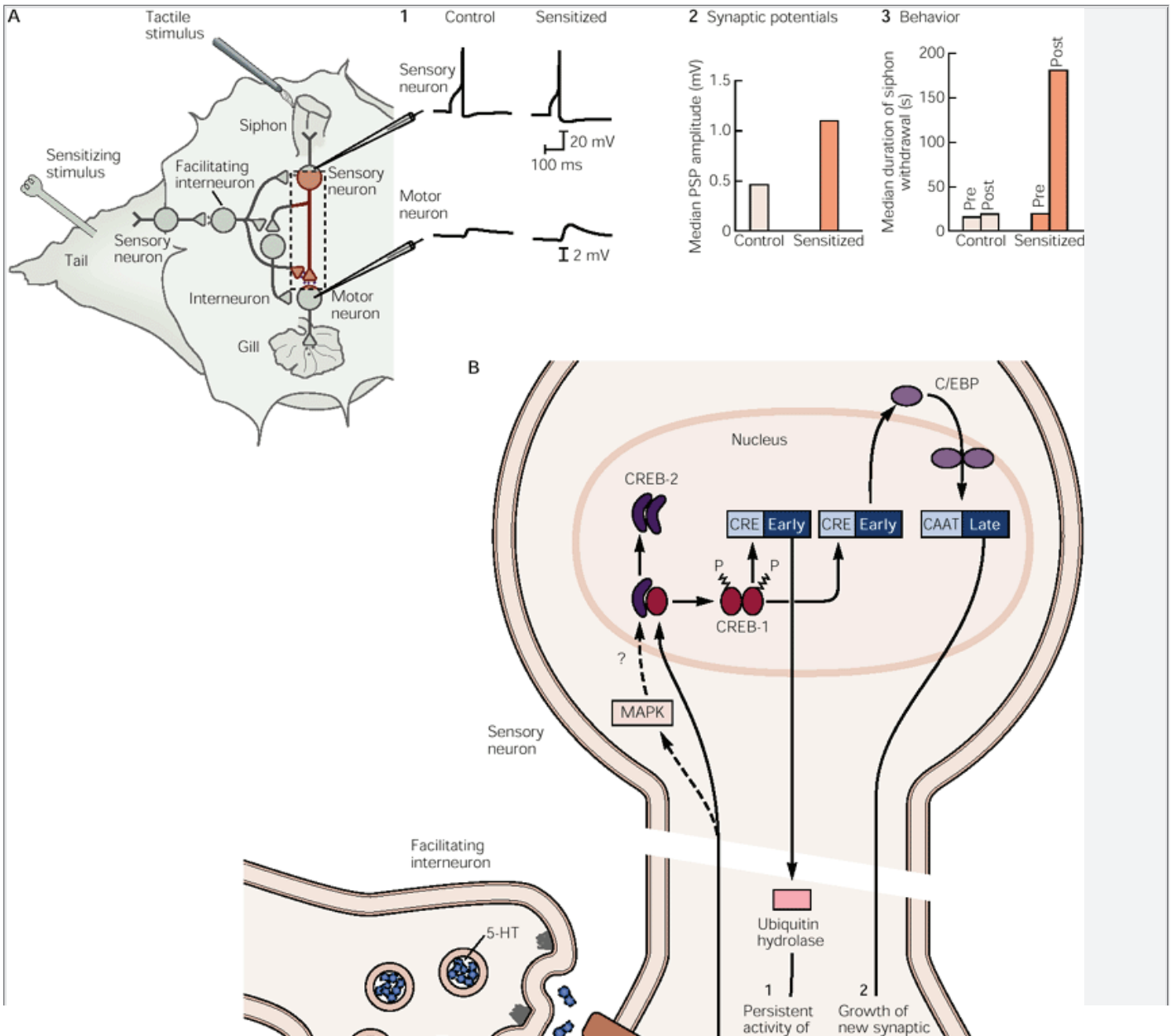
Despite these similarities, short- and long-term memory are distinct processes that can be distinguished by several criteria. In humans, epileptic seizure or head trauma affects long-term memory but not short-term memory. A similar dissociation between short- and long-term memory can be demonstrated in experimental animals using inhibitors of protein or mRNA synthesis to block long-term memory selectively.

As we saw in the preceding chapter, the process by which transient short-term memory is converted into a stable long-term memory is called *consolidation.*
Consolidation of long-term implicit memory for simple forms

P.1255

P.1256

of learning involves three processes: gene expression, new protein synthesis, and the growth (or pruning) of synaptic connections.



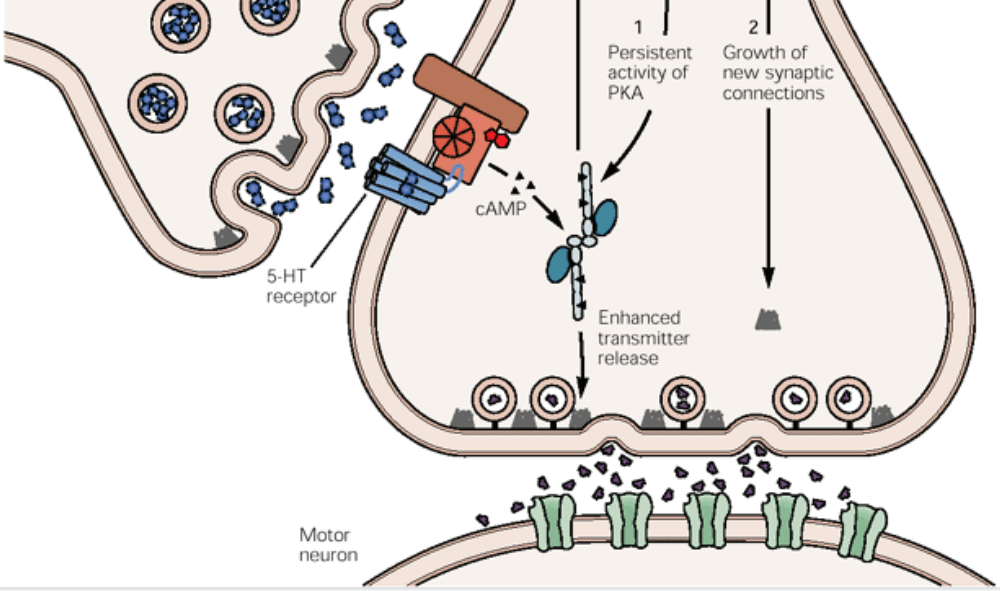


Figure 63-5 (Facing page) Persistent synaptic enhancement with long-term sensitization.

A. Long-term sensitization of the gill-withdrawal reflex of *Aplysia* involves facilitation of transmitter release at the connections between sensory and motor neurons. **1.** The recordings show representative synaptic potentials in a siphon sensory neuron and a gill motor neuron in a control animal and in an animal that received long-term sensitization training by repeated stimulation of its tail. The record was obtained one day after the end of training. **2.** The median amplitude of the post-synaptic potentials (**PSP**) in an identified gill motor neuron is greater in sensitized animals one day after training than in control animals. **3.** The effect of sensitization on the neural circuit of the gill- and siphon-withdrawal reflex is measured here by the median duration of withdrawal of the siphon (see Figure 63-1). (**Pre** = score before training; **post** = score after training.) The experimental group was tested one day after the end of training. (Adapted from Frost et al. 1985.)

B. Long-term sensitization of the gill-withdrawal reflex of *Aplysia* leads to two major changes in the sensory neurons of the reflex: persistent activity of protein kinase A and structural changes in the form of the growth of new synaptic connections.

Both the short-term and long-term facilitation are initiated by a serotonergic interneuron. Short-term facilitation (lasting minutes to hours), resulting from a single tail shock or a single pulse of serotonin, leads to covalent modification of preexisting proteins (short-term pathway). As shown in Figure 63-3, serotonin acts on a postsynaptic receptor to activate the enzyme adenylyl cyclase, which converts ATP to the second messenger cAMP. In turn, cAMP activates the cAMP-dependent protein kinase A, which phosphorylates and covalently modifies a number of target proteins, leading to enhanced transmitter availability and release. The duration of these modifications is a measure of the short-term memory.

Long-term facilitation (lasting one or more days) involves the synthesis of new proteins. The switch for this inductive mechanism is initiated by protein kinase A (**PKA**), which recruits the mitogen-activated kinase (**MAPK**) and together they translocate to the nucleus (long-term pathway), where PKA phosphorylates the cAMP-response element binding (**CREB**) protein. The transcriptional activators bind to cAMP response elements (**CRE**) located in the upstream region of two types of cAMP-inducible genes. To activate CREB-1, PKA needs also to remove the repressive action of CREB-2, which is capable of inhibiting the activation capability of CREB-1. PKA is thought to mediate the derepression of CREB-2 by means of another protein, MAPK. One gene activated by CREB encodes a ubiquitin hydrolase, a component of a specific ubiquitin protease that leads to the regulated proteolysis of the regulatory subunit of PKA. This cleavage of the (inhibitory) regulatory subunit results in persistent activity of PKA, leading to persistent phosphorylation of the substrate proteins of PKA, including both CREB-1 and the protein involved in the short-term process. The second gene activated by CREB encodes another transcription factor **C/EBP**. This binds to the DNA response element **CAAT**, which activates genes that encode proteins important for the growth of new synaptic connections.

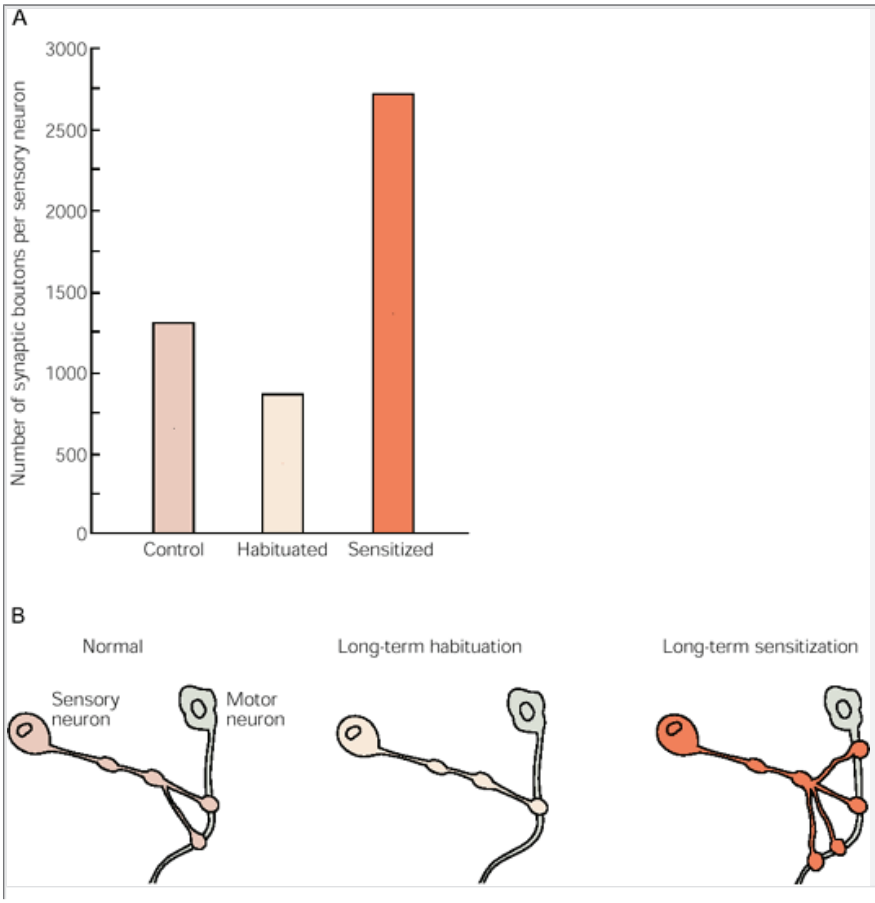


Figure 63-6 Long-term habituation and sensitization in *Aplysia* involve structural changes in the presynaptic terminals of sensory neurons. (Adapted from Bailey and Chen 1983.)

A. When measured 1 day or 1 week after training, the number of presynaptic terminals is highest in sensitized animals (about 2800) compared with control (1300) and habituated animals (800).

B. Long-term habituation leads to a loss of synapses and long-term sensitization leads to an increase in synapses.

How do genes and proteins operate in the consolidation of long-term functional changes? Studies of long-term sensitization of the gill-withdrawal reflex indicate that with repeated application of serotonin the catalytic subunit of PKA recruits another second messenger kinase, the mitogen-activated protein (MAP) kinase, a kinase commonly associated with cellular growth. Together the two kinases translocate to the nucleus of the sensory neurons, where they activate a genetic switch (see the discussion of transcriptional regulation in Chapter 13). Specifically, the catalytic subunit phosphorylates and thereby activates a transcription factor called *CREB-1* (*c* AMP *r* esponse *e* lement *b* inding protein). This transcriptional activator, when phosphorylated, binds to a promoter element called *CRE* (the *c* AMP *r* esponse *e* lement). By means of the MAP kinase the catalytic subunit of PKA also acts indirectly to relieve the inhibitory actions of *CREB-2*, a repressor of transcription.

The presence of both a repressor (CREB-2) and an activator (CREB-1) of transcription at the very first step in long-term facilitation suggests that the threshold for putting information into long-term memory is highly *regulated.* Indeed, we can see in everyday life that the ease with which short-term memory is transferred into long-term memory varies greatly depending on attention, mood, and social context. In fact, when the repressive action of CREB-2 is relieved (by injecting, for example, a specific antibody to CREB-2), a single pulse of serotonin, which normally produces only short-term facilitation lasting minutes, is able to produce long-term facilitation, the cellular homolog of long-term memory.

Under normal circumstances the physiological relief of the repressive action of CREB-2 and the activation of CREB-1 induce expression of downstream target genes, two of which are particularly important: (1) the enzyme ubiquitin carboxyterminal hydrolase, which activates proteasomes to make PKA persistently active, and (2) the transcription factor C/EBP, one of the components of a gene cascade necessary for the growth of new synaptic connections. The induction of the hydrolase is a key step in the recruitment of a regulated proteolytic

complex: the ubiquitin-dependent proteosome. As in other cellular contexts, ubiquitin-mediated proteolysis also produces a cellular change of state, here by removing inhibitory constraints on memory. One of the substrates of this proteolytic process is the regulatory subunit of PKA.

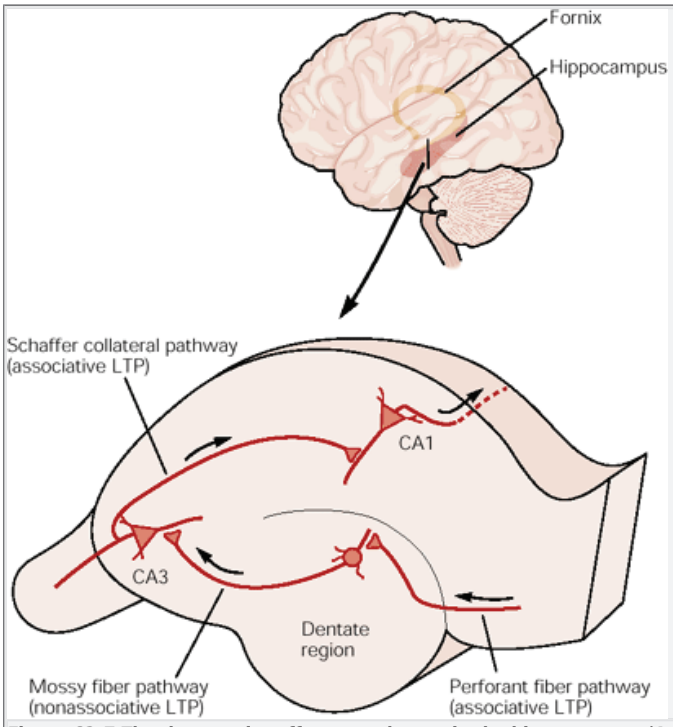


Figure 63-7 The three major afferent pathways in the hippocampus. (**Arrows** denote the direction of impulse flow.) The perforant fiber pathway from the entorhinal cortex forms excitatory connections with the granule cells of the dentate gyrus. The granule cells give rise to axons that form the mossy fiber pathway, which connects with the pyramidal cells in area CA3 of the hippocampus. The pyramidal cells of the CA3 region project to the pyramidal cells in CA1 by means of the Schaffer collateral pathway. Long-term potentiation (**LTP**) is nonassociative in the mossy fiber pathway and associative in the other two pathways.

PKA is made up of four subunits: two regulatory submits inhibit two catalytic subunits (Chapter 13). Long-term training and the induction of the hydrolase degrades about 25% of the regulatory (inhibitory) subunits in the sensory neurons. As a result, the catalytic subunits continue phosphorylating proteins important for enhancing transmitter release and strengthening the synaptic connections, including CREB-1, long after the second messenger, cAMP, has returned to its basal level (Figure 63-5B). This is the simplest mechanism for long-term memory: a second-messenger kinase critical for the short-term process is made persistently active for up to 24 hours by repeated training, without requiring a continuous signal of any sort. The kinase becomes autonomous and does not require either serotonin, cAMP, or PKA.

The second and more enduring consequence of the activation of CREB-1 is a cascade of gene activation that leads to the growth of new synaptic connections. It is this growth process that provides the stable, self- maintained state of long-term memory. In *Aplysia* the number of presynaptic terminals in the sensory neurons of the gill-withdrawal pathway increases and becomes twice as great in the long term in sensitized animals as in untrained animals (Figure 63-6). This structural change is not limited to the sensory neurons. In animals that have been sensitized for the long term, the dendrites of the motor neurons grow to accommodate the additional synaptic input. Such morphological changes do not occur with short-term sensitization. Long-term habituation, in contrast, leads to *pruning* of synaptic connections. The long-term inactivation of the functional connections between sensory and motor neurons reduces the number of terminals for each neuron by one-third (Figure 63-6), and the proportion of terminals with active zones from 40% to 10%.

Genetic Analyses of Implicit Memory Storage for Classical Conditioning Also Implicate the cAMP-PKA-CREB Pathway

How general is the role of the cAMP-PKA-CREB pathway in long-term memory storage? Does it apply to other species and other types of learning? The fruit fly *Drosophila* is particularly amenable to genetic manipulation. As first shown by Seymour Benzer and his students, *Drosophila* can be classically conditioned, and four interesting mutations in single genes that lead to a learning deficit have been isolated: *dunce, rutabaga, amnesiac*, and *PKA-R1.* Studies of these mutants have given

rise to two general conclusions. First, all of the mutants that fail to show classical conditioning also fail to show sensitization. Second, all four mutants have a defect in the cAMP cascade. *Dunce* lacks phosphodiesterase, an enzyme that degrades cAMP. As a result, this mutant has abnormally high levels of cAMP that are thought to be beyond the range of normal modulation. *Rutabaga* is defective in the Ca^{2+}/calmodulin-dependent adenylyl cyclase and therefore has a low basal level of cAMP. *Amnesiac* lacks a peptide transmitter that acts on adenylyl cyclase, and *PKA-R1* is defective in PKA.

More recently a reverse genetic approach has been used to explore memory storage in *Drosophila*. Various transgenes (see Chapter 3) are placed under the control

P.1258

of an inducible promoter that is heat-sensitive, so that by heating and cooling the fly a particular gene can be turned on or off. This inducible control over gene expression, which we shall return to again later in the chapter, is useful for studying synaptic or behavioral plasticity in adult animals. It minimizes any potential effect that a transgene might produce on the development of the brain and therefore allows one to read out the selective effect of the gene on adult behavior.

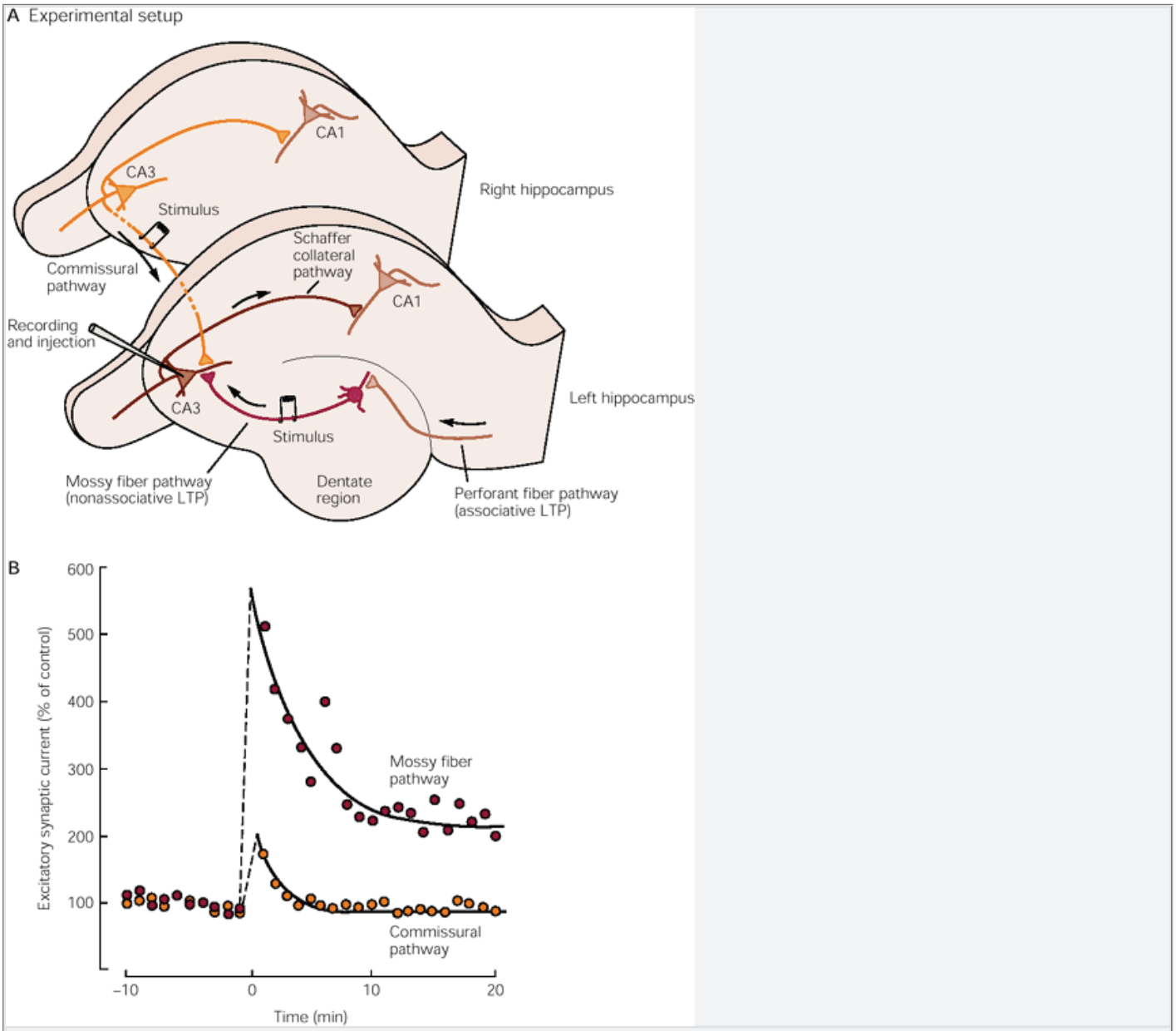


Figure 63-8 Long-term potentiation (LTP) of the mossy fiber pathway to the CA3 region of the hippocampus.

A. Experimental arrangement for studying LTP in the CA3 region of the hippocampus. Stimulating electrodes are placed so as to activate two independent pathways to the CA3 pyramidal cells: The commissural pathway from the CA3 region of the contralateral hippocampus and the ipsilateral mossy fiber pathway.

B. Whole-cell voltage-clamp recording allows injection of both fluoride and the Ca^{2+} chelator BAPTA into the cell body of the CA3 neuron. Together these two drugs are thought to block *all* second-messenger pathways in the postsynaptic cell. Despite this drastic biochemical blockade of the postsynaptic cell, LTP in the mossy fiber pathway is unaffected and is therefore thought to be presynaptically induced. In contrast, these injections do block LTP in the commissural pathway. This pathway requires activation of the *N* -methyl-D- aspartate (NMDA) receptor, and here induction of LTP is postsynaptic. (Adapted from Zalutsky and Nicoll 1990).

The first such experiment involved inducing the expression of transgenes that blocked the catalytic subunit of PKA. William Quinn and his colleagues found that blocking the action of PKA, even transiently, interferes with the fly's ability to learn and to form short-term memory. A similar disruption of learning and memory was observed in a mutant of a *Drosophila* homolog of the PKA catalytic subunit. These experiments indicate the importance of the cAMP signal transduction pathway is critical for associative learning and short-term memory in *Drosophila*.

Long-term memory after repeated training in *Droso-phila* also requires new protein synthesis. *Drosophila* expresses

P.1259

both a CREB activator and a CREB-2 repressor. Jerry Yin, Tim Tully, and their colleagues found that overexpression of the repressor (CREB-2), which presumably prevents the expression of cAMP-activated genes, selectively blocks long-term memory without interfering with learning or short-term memory. Conversely, overexpression of the CREB activator results in immediate long-term memory, even with a training procedure that produces only short-term memory in wild-type flies.

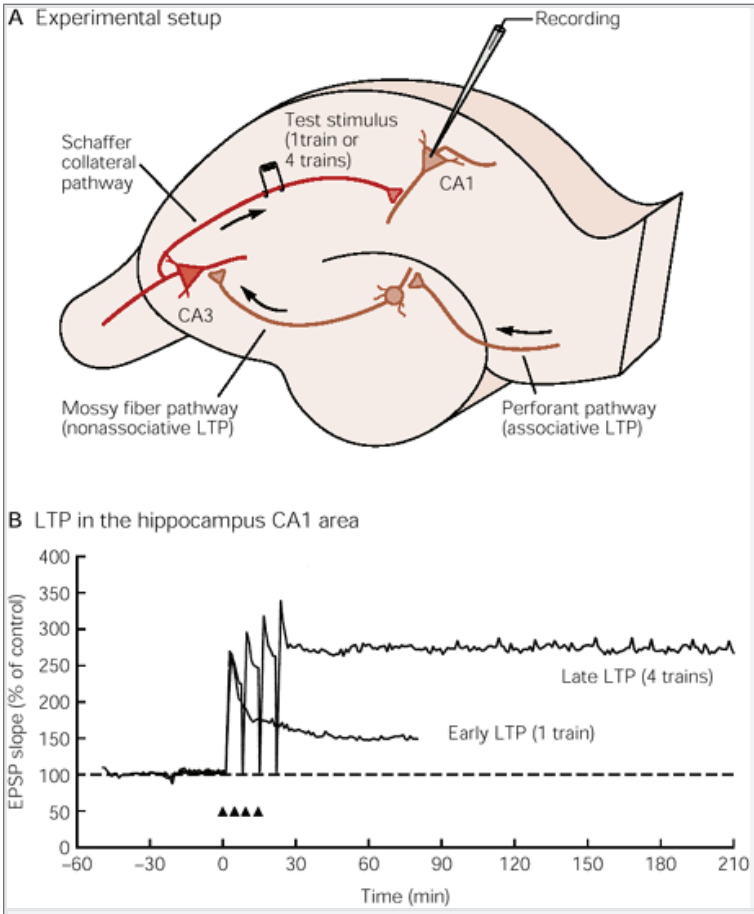


Figure 63-9 Long-term potentiation (LTP) in the Schaffer collateral pathway to the CA1 region of the hippocampus.

A. Experimental setup for studying LTP in the CA1 region of the hippocampus. The Schaffer collateral pathway is stimulated electrically and the response of the population of pyramidal neurons is recorded.

B. Comparison of early and late LTP in a cell in the CA1 region of the hippocampus. The graph is a plot of the slope (rate of rise) of the excitatory postsynaptic potentials (**EPSP**) in the cell as a function of time. The slope is a measure of synaptic efficacy. Excitatory postsynaptic potentials were recorded from outside the cell. A test stimulus was given every 60 s to the Schaffer collaterals. To elicit early LTP a single train of stimuli is given for 1 s at 100 Hz. To elicit the late phase of LTP four trains are given separated by 10 min. The resulting early LTP lasts 2-3 hours, whereas the late LTP lasts 24 or more hours.

Explicit Memory in Mammals Involves Long-Term Potentiation in the Hippocampus

What mechanisms are used to store explicit memory—information about people, places, and objects? One important component of the medial temporal system of higher vertebrates involved in the storage of explicit memory is the hippocampus (Chapter 62). As first shown by Per Andersen, the hippocampus has three major pathways: (1) the *perforant pathway*, which projects from the entorhinal cortex to the granule cells of the dentate gyrus; (2) the *mossy fiber pathway*, which contains the axons of the granule cells and runs to the pyramidal cells in the CA3 region of the hippocampus; and (3) the *Schaffer collateral pathway*, which consists of the excitatory collaterals of the pyramidal cells in the CA3 region and ends on the pyramidal cells in the CA1 region (Figure 63-7).

In 1973 Timothy Bliss and Terje Lom•' discovered that each of these pathways is remarkably sensitive to the history of previous activity. A brief high-frequency train of stimuli (a tetanus) to any of the three major synaptic pathways increases the amplitude of the excitatory postsynaptic potentials in the target hippocampal neurons. This facilitation is called *long-term potentiation* (LTP). The mechanisms underlying LTP are not the same in all three pathways. LTP can be studied in the intact animal, where it can last for days and even weeks. It can also be examined in slices of hippocampus and in cell culture for several hours. We shall first consider the mossy fiber pathway.

Long-Term Potentiation in the Mossy Fiber Pathway Is Nonassociative

The mossy fiber pathway consists of the axons of the granule cells of the dentate gyrus. The mossy fiber terminals release glutamate as a transmitter, which binds to both NMDA and non-NMDA receptors on the target pyramidal cells. However, in this pathway the NMDA receptors have only a minor role in synaptic plasticity under most conditions; blocking the NMDA receptors has no effect on LTP. Similarly, blocking Ca^{2+} influx into the postsynaptic pyramidal cells in the CA3 region does not affect LTP (Figure 63-8).

Instead, LTP in the mossy fiber pathway region has been found to depend on Ca^{2+} influx into the *presynaptic* cell after the tetanus. The Ca^{2+} influx appears to activate Ca^{2+}/calmodulin-dependent adenylyl cyclase thereby increasing the level of cAMP and activating PKA in the presynaptic neuron, just as in the sensory neurons of *Aplysia* during associative learning. Moreover, mossy fiber LTP can be regulated by a modulatory input. This input is noradrenergic and engages β-adrenergic receptors, which activate adenylyl cyclase, as does the serotonergic input in *Aplysia*.

Long-Term Potentiation in the Schaffer Collateral and Perforant Pathways Is Associative

The Schaffer collateral pathway connects the pyramidal cells of the CA3 region of the hippocampus with those of the CA1 region (Chapter 5 and Figures 63-7 and 63-9A). Like the mossy fiber terminals, the terminals of the Schaffer collaterals also use glutamate as transmitter, but LTP in the Schaffer collateral pathway requires activation of the NMDA-type of glutamate receptor (Figures 63-9B and 63-10). Therefore, LTP in CA1 cells has two characteristic features that distinguish it from LTP in the mossy fiber pathway, both of which derive from the known properties of the NMDA receptor.

First, LTP in the Schaffer collateral pathway typically requires activation of several afferent axons together, a feature called *cooperativity.* This feature derives from the fact that the NMDA receptor-channel becomes functional and conducts Ca^{2+} only when two conditions are met: Glutamate must bind to the postsynaptic NMDA receptor *and* the membrane potential of the postsynaptic cell must be sufficiently depolarized by the cooperative firing of several afferent axons to expel Mg^{2+} from the

mouth of the channel (Figure 63-10). Only when Mg^{2+} is expelled can Ca^{2+} influx into the postsynaptic cell occur. Calcium influx initiates the persistent enhancement of synaptic transmission by activating two calcium-dependent serine-threonine protein kinases—the Ca^{2+}/calmodulin-dependent protein kinase and protein kinase C—as well as PKA and the tyrosine protein kinase fyn.

Second, LTP in the Schaffer collateral pathway requires concomitant activity in both the presynaptic and postsynaptic cells to adequately depolarize the post-synaptic cell, a feature called *associativity*. As we have seen, to initiate the Ca^{2+} influx into the postsynaptic cell, a strong presynaptic input sufficient to fire the postsynaptic cell is required.

The finding that LTP in the Schaffer collateral pathway requires simultaneous firing in both the postsynaptic and presynaptic neurons provides direct evidence for *Hebb's rule*, proposed in 1949 by the psychologist Donald Hebb: "When an axon of cell A… excites cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells so that A's efficiency as one of the cells firing B is increased." As discussed in Chapter 56, a similar principle is involved in fine-tuning synaptic connections during the late stages of development.

The induction of LTP in the CA1 region of the hippocampus depends on four postsynaptic factors: postsynaptic depolarization, activation of NMDA receptors, influx of Ca^{2+}, and activation by Ca^{2+} of several second-messenger systems in the postsynaptic cell. The mechanisms for the expression of this LTP, on the other hand, is still uncertain. It is thought to involve not only

an increase in the sensitivity and number of the postsynaptic non-NMDA (AMPA) receptors to glutamate as a result of being phosphorylated by the Ca^{2+}/calmodulin-dependent protein kinase, but also an increase in transmitter release from the *presynaptic* terminals of the CA3 neuron (Figure 63-11). Evidence for enhanced presynaptic function is based on two observations. First, biochemical studies suggest that the release of glutamate is enhanced during LTP. Second, as we shall see later, quantal analysis indicates that the probability of transmitter release increases greatly during LTP.

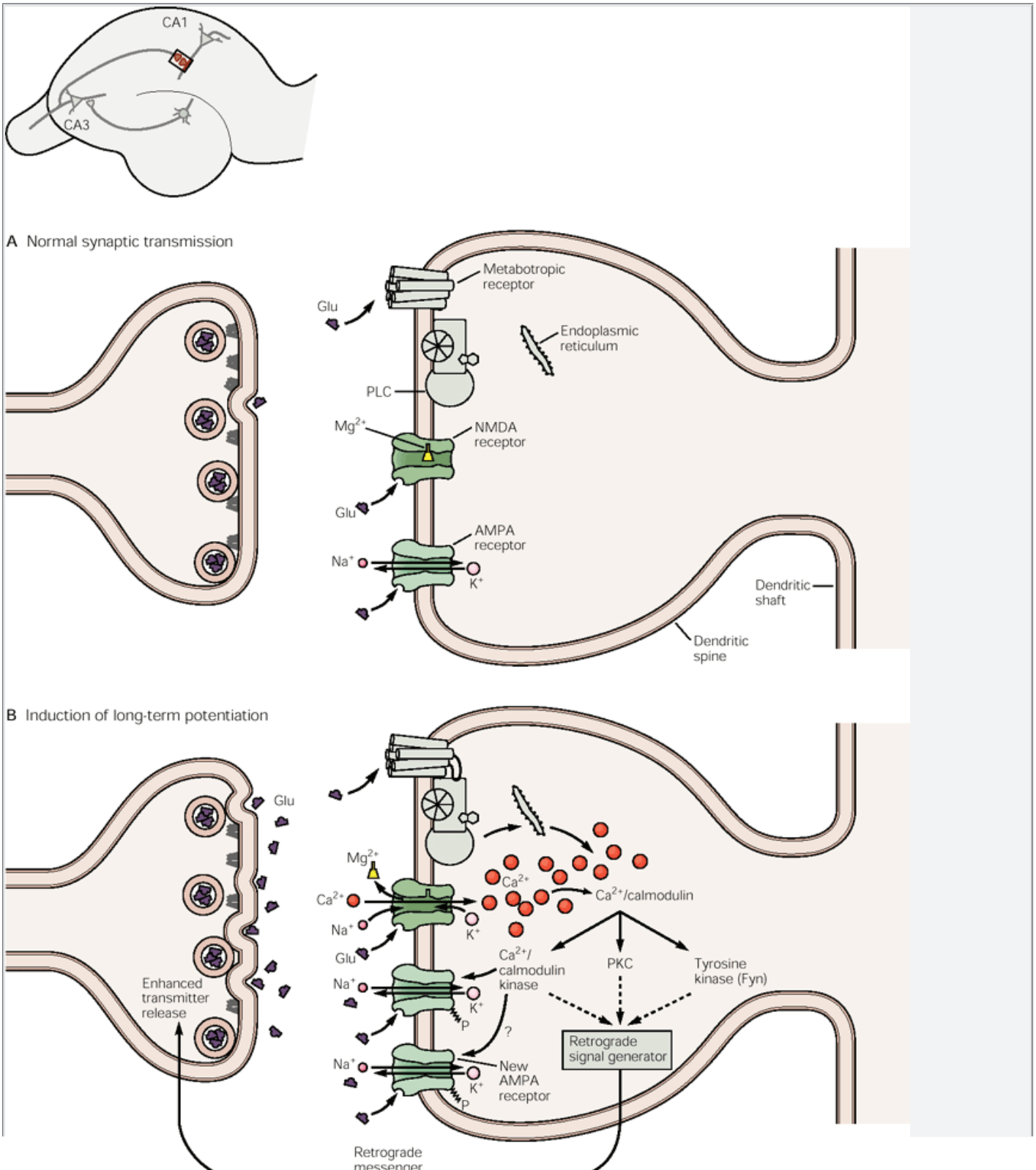




Figure 63-10 (Opposite) A model for the induction of the early phase of long-term potentiation. According to this model NMDA and non-NMDA receptor-channels are located near each other in dendritic spines.

A. During normal, low-frequency synaptic transmission glutamate (**Glu**) is released from the presynaptic terminal and acts on both the NMDA and non-NMDA receptors. The non-NMDA receptors here are the AMPA type. Na^+ and K^+ flow through the non-NMDA channels but not through the NMDA channels, owing to Mg^{2+} blockage of this channel at the resting level of membrane potential.

B. When the postsynaptic membrane is depolarized by the actions of the non-NMDA receptor-channels, as occurs during a high-frequency tetanus that induces LTP, the depolarization relieves the Mg^{2+} blockage of the NMDA channel. This allows Ca^{2+} to flow through the NMDA channel. The resulting rise in Ca^{2+} in the dendritic spine triggers calcium-dependent kinases (Ca^{2+}/calmodulin kinase and protein kinase C) and the tyrosine kinase Fyn that together induce LTP. The Ca^{2+}/calmodulin kinase phosphorylates non-NMDA receptor-channels and increases their sensitivity to glumate thereby also activating some otherwise silent receptor channels. These changes give rise to a postsynaptic contribution for the maintenance of LTP. In addition, once LTP is induced, the postsynaptic cell is thought to release (in ways that are still not understood) a set of retrograde messengers, one of which is thought to be nitric oxide, that act on protein kinases in the presynaptic terminal to initiate an enhancement of transmitter release that contributes to LTP.

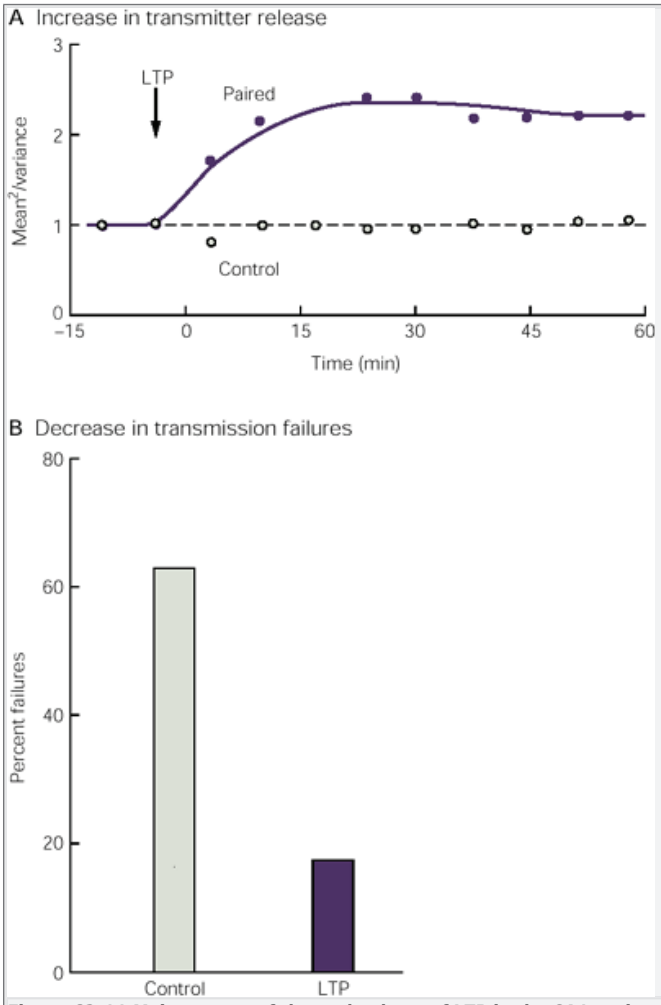


Figure 63-11 Maintenance of the early phase of LTP in the CA1 region of the hippocampus depends on an increase in presynaptic transmitter release. Quantal analysis of LTP in area CA1 is based on a coefficient of variation of evoked responses. This analysis assumes that the number of quanta of transmitter released follows a binomial distribution, where the coefficient of variation (mean squared/variance) provides an index of transmitter release from the presynaptic terminal that is independent of quantal size. (From Malinow and Tsien 1990.)

A. With LTP the ratio of mean squared to variance increases, indicating an increase in transmitter release. This increase occurs only in the pathway that is paired with depolarization of the postsynaptic cell. It does not occur in a control pathway that is not paired.

B. At normal rates of stimulation the number of failures in transmission is significant (60%). After LTP the percentage of failures decreases to 20%, another indication that LTP is presynaptic.

Since induction of LTP requires events only in the postsynaptic cell (Ca^{2+} influx through NMDA channels), whereas expression of LTP is due in part to a subsequent event in the presynaptic cells (increase in transmitter release), the presynaptic cells must somehow receive information that LTP has been induced. There is now evidence that calcium-activated second messengers, or perhaps Ca^{2+} itself, causes the postsynaptic cell to release one or more retrograde messengers from its active dendritic spines. Recent pharmacological and genetic experiments have identified nitric oxide (NO), a gas that diffuses readily from cell to cell, as one of the possible candidate retrograde messengers involved in LTP.

These studies of the Schaffer collateral pathway indicate that LTP in CA1 uses two associative mechanisms in series: a Hebbian mechanism (simultaneous firing in both the pre- and postsynaptic cells) and activity-dependent presynaptic facilitation. A similar set of mechanisms is responsible for LTP in the perforant pathway. As we saw earlier, two associative mechanisms in series also contribute to classical conditioning in *Aplysia*.

Long-Term Potentiation Has a Transient Early and a Consolidated Late Phase

As with memory storage (Chapter 62), LTP has phases. One stimulus train produces an early, short-term phase of LTP (called *early LTP*) lasting 1-3 hours; this component does not require new protein synthesis. Four or more trains induce a more persistent phase of LTP (called *late LTP*) that lasts for at least 24 hours and requires new protein and RNA synthesis. As we have seen, the mechanisms for the early, short-term phase are quite different in the Schaffer collateral and mossy fiber

pathways. However, the mechanisms for the late, long-term phase in the two pathways appear similar. In both pathways late-phase LTP requires the synthesis of new mRNA and protein and recruits the cAMP-PKA-MAPK-CREB signaling pathway.

What are the properties of this late phase of LTP? Does long-term explicit memory storage, like implicit

P.1263

P.1264

memory storage, also require the growth of new synaptic connections? In fact, cellular-physiological studies are beginning to suggest that the late phase of LTP involves the activation, perhaps the growth, of additional presynaptic machinery for transmitter release and the insertion of new clusters of postsynaptic receptors.

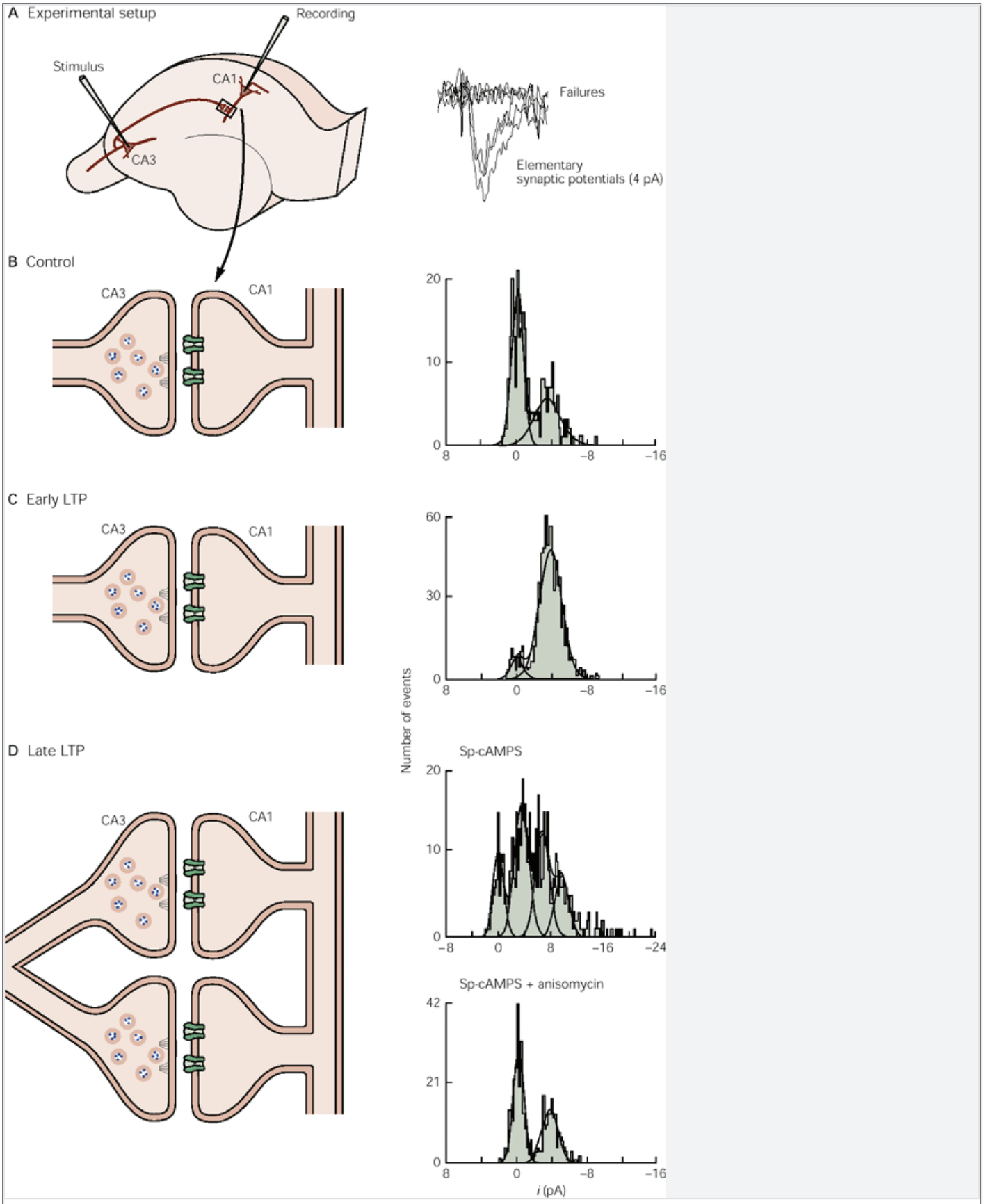


Figure 63-12 The early and late phases of LTP are evident in the synaptic transmission between a single CA3 cell and a single CA1 cell. (From Bolshakov et al. 1997.)

A. A single CA3 cell can be stimulated selectively to produce a single elementary synaptic potential in a CA1 cell. When the CA3 cell is stimulated repeatedly at low frequency, it gives rise to either an elementary response of the size of a miniature synaptic potential or a failure.

B. In control cells there are many failures; the synapse has a low probability of releasing vesicles. The distribution of the amplitudes of many responses can be approximated by two Gaussian curves, one centered on zero (the failures) and the other centered on 4 pA (the successful responses). These histograms are consistent with the type of synapse illustrated here, in which a single CA3 cell makes a single connection on a CA1 cell. This connection has a single active zone from which it releases a single vesicle in an all-or-none manner (failures or successes).

C. With the early phase of LTP the probability of release rises significantly, but the two Gaussian curves in the distribution of responses is consistent with the view that a single release site still releases only a vesicle but now with a high probability of release.

D. When the late phase of LTP is induced by a cAMP analog (**Sp-cAMPS**), the distribution of responses no longer fits two Gaussian curves but instead requires three or four Gaussian curves, suggesting the possibility that new presynaptic active zones and postsynaptic receptors have grown. These effects are blocked by anisomycin, an inhibitor of protein synthesis.

Charles Stevens and his colleagues and Steven Siegelbaum and Vadim Bolshakov have now examined LTP by stimulating a *single* presynaptic CA3 neuron and recording from a *single* CA1 postsynaptic cell. When the CA3 neuron is stimulated repeatedly at a low rate, most of the time the CA1 cell fails to respond with a synaptic potential. Only on occasion does activity in the presynaptic neuron lead to a small response, about 4 pA in amplitude, in the postsynaptic cell (Figure 63-12A). This response is approximately the size of a single spontaneous miniature synaptic potential (Chapter 14). When many failures and unitary responses are collected and measured, the failures of release and the unitary responses can be described by two random (Gaussian) distributions, one centered at zero, corresponding to the failures of release, and the other centered around 4 pA, corresponding both to successful responses and to the size of spontaneously released miniature synaptic potentials (Figure 63-12B). These distributions lend themselves to a surprisingly clear anatomical explanation. They suggest that a single CA3 neuron makes only a single functional synaptic contact on a CA1 neuron. This single synaptic contact appears to have only one active zone from which the transmitter content of only a single vesicle is released, in an all-or-none way, by a presynaptic action potential. In the basal state (where there are many more failures than responses) the probability of release of the vesicle is low. Thus, this situation is not very different from other central synaptic connections where a single release site typically releases only a single vesicle in an all-or-none fashion (Chapter 15).

What happens during the early phase of LTP? In the early phase the number of failures decreases and the number of successes increases, but the amplitude histograms of the responses and failures are still fit by two Gaussian distributions (corresponding to failures of release and successful responses). This indicates that the early phase of LTP produces no change in the number of synapses, the number of active zones, or the maximal number of vesicles released with each action potential (Figure 63-12C). Thus the early phase of LTP represents a functional change—an increase in the *probability* of transmitter release—without structural changes. An action potential still releases only one vesicle of transmitter from a single release site, but now it does so more reliably.

An equivalent of the late-phase LTP can be induced chemically, by exposing the neurons to permeant cAMP. After the late phase of LTP begins the distribution of successful responses changes dramatically and can no longer be approximated by two Gaussian functions. The responses now are not simply zero and 4 pA but are 8, 12, and even 20 pA in amplitude, so that several Gaussian curves are required to describe the distribution of responses (Figure 63-12D). This change suggests that during the late phase of LTP a single action potential in a single CA3 cell releases several vesicles of transmitter onto the CA1 neuron. Since each release site is thought to release only one vesicle in an all-or-none fashion, such an increase in the number of vesicles released would seem to entail growth of new pre- synaptic release sites as well as new clusters of post-synaptic receptors. Consistent with this idea, and with the properties of the late phase of LTP, the generation of these new distributions requires new protein synthesis (Figure 63-13).

Genetic Interference With Long-Term Potentiation Is Reflected in the Properties of Place Cells in the Hippocampus

LTP is an artificially induced change in synaptic strength produced by electrical stimulation of synaptic pathways. Does this form of synaptic plasticity exist naturally? If so, how does it affect the normal processing of information for memory storage in the hippocampus?

In 1971 John O'Keefe and John Dostrovsky made the remarkable discovery that the hippocampus contains a cognitive map of the spatial environment in which an animal moves. The location of an animal in a particular space is encoded in the firing pattern of individual pyramidal cells, the very cells that undergo LTP when their afferent pathways are stimulated electrically.

A mouse's hippocampus has about a million pyramidal cells. Each of these cells is potentially a *place cell*, encoding a position in space. When an animal moves around, different place cells in the hippocampus fire. For example, one cell will fire only when the animal's head enters a particular area in the north end of the space, while other cells will fire when the animal takes other positions at the south end of that space (Figure 63-14A,63-14B). Thus, the mouse's whereabouts are signaled by the discharge of a unique population of hippocampal place cells.

By this means the animal is thought to form a "place field," an internal representation of the space that it occupies. When the animal enters a new environment, new place fields are formed within minutes and are stable

for weeks to months. The same pyramidal cells may signal different information in different environments and can therefore be used in more than one spatial map.

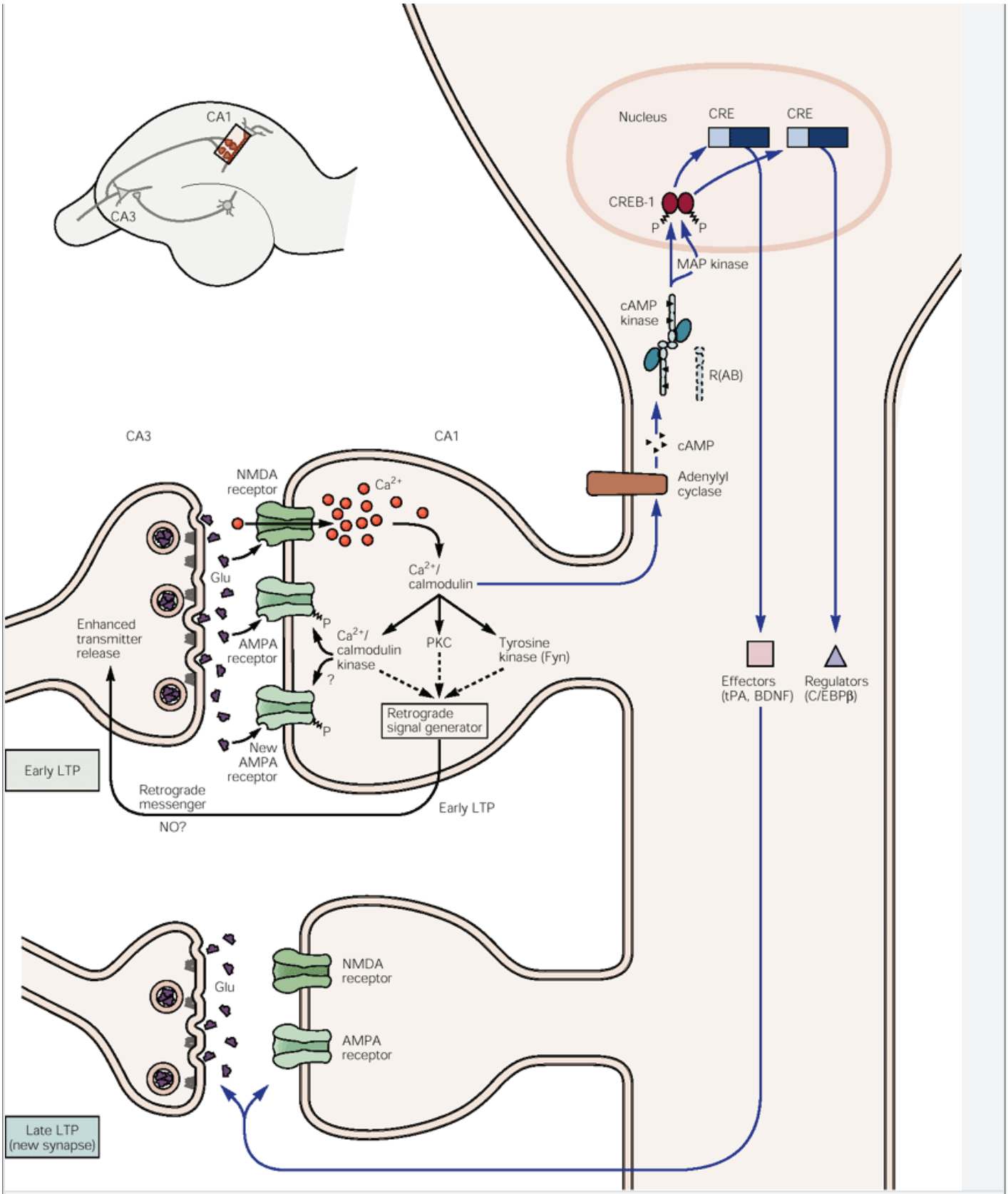


Figure 63-13 A model for the early and late phase of LTP. A single train of action potentials leads to early LTP by activating NMDA receptors, Ca^{2+} influx into the postsynaptic cell, and a set of second messengers. With repeated trains the Ca^{2+} influx also recruits an adenylyl cyclase, which activates the cAMP-dependent protein kinase (cAMP kinase) leading to its translocation to the nucleus, where it phosphorylates the CREB protein. CREB in turn activates targets that are thought to lead to structural changes. Mutations in mice that block PKA or CREB reduce or eliminate the late phase of LTP. The adenylyl cyclase can also be modulated by dopaminergic and perhaps other modulatory inputs. **BDNF** = brain-derived neurotrophic factor; **C/EBP**β = transcription factor; **P** = phosphate; **R(AB)** dominant negative PKA; **tPA** tissue plasminogen activator.

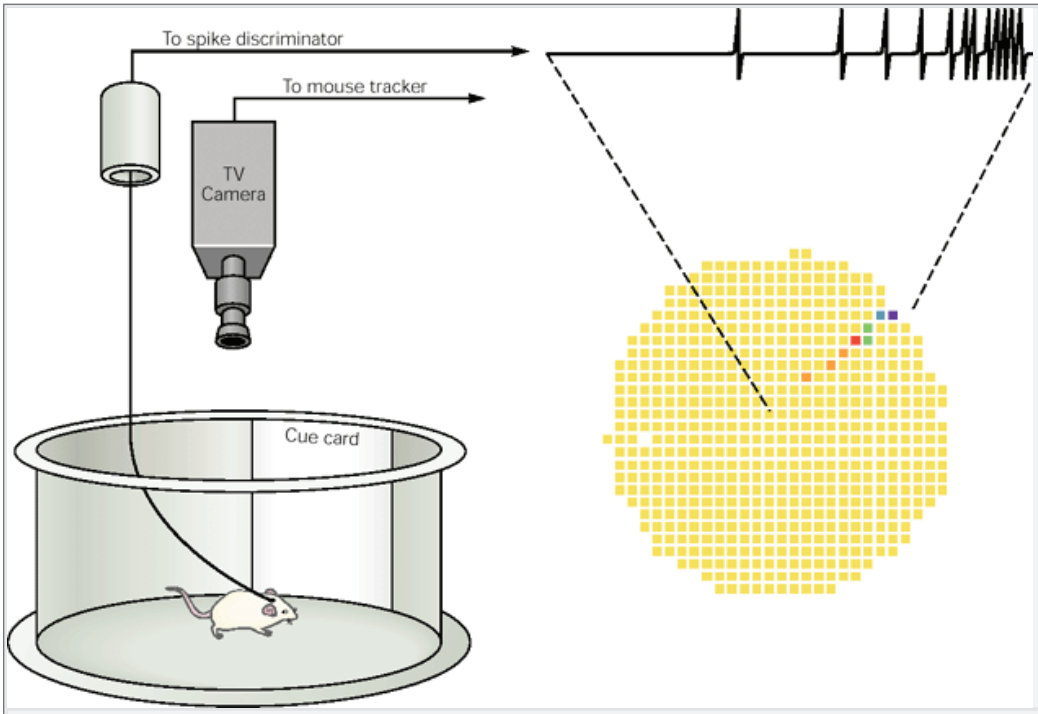


Figure 63-14A The firing patterns of pyramidal cells in the hippocampus create an internal representation of the animal's location within its surrounding. A mouse is attached to a recording cable and placed inside a cylinder (49 cm in diameter by 34 cm high). The other end of the cable goes to a 25-channel commutator attached to a computer-based spike-discrimination system. The cable is also used to supply power to a light-emitting diode mounted on the headstage the mouse carries. The entire apparatus is viewed with an overhead TV camera whose output goes to a tracking device that detects the position of the mouse. The output of the tracker is sent to the same computer used to detect spikes, so that parallel time series of positions and spikes are recorded. The occurrence of spikes as a function of position is extracted from the basic data and is used to form two-dimensional firing-rate patterns that can be numerically analyzed or visualized as color-coded firing-rate maps. (Based on Muller et al. 1987.)

The rapid formation of place fields and their persistence for weeks offer an excellent opportunity to ask, How are place fields formed and maintained? Is LTP important for the formation or maintenance of place fields? To address these questions two types of mutations have been examined in mice, each of which interferes with LTP in a different way.

One mutation was produced by Joe Tsien, Susumu Tonegawa, and their colleagues by selectively knocking out NMDA R1, one of the subunits of the NMDA receptor, in the pyramidal cells of the CA1 region. This restricted knockout, limited to just the CA1 region, disrupts LTP in the Schaffer collateral pathway completely (Box 63-1).

In the other mutation, produced by Mark Mayford and his colleagues, a persistently active form of the Ca^{2+}/calmodulin-dependent kinase is expressed throughout the hippocampus. This mutated gene product does not affect LTP produced at 100 Hz stimulation, the frequency commonly used in the laboratory, but it does interfere with LTP produced at low frequencies of stimulation, in the range of 1-10 Hz (Box 63-1). These lower frequencies are interesting because they are in the physiological range of a prominent spontaneous rhythm in the hippocampus, called the *theta rhythm*, that occurs in an animal as it moves in the environment.

In both types of mutants the interference with LTP does not prevent the formation of place fields. Although the place fields formed in the absence of LTP are larger and fuzzier in outline than normal, LTP is not required for the basic transformation of sensory information into

place fields. LTP is required for fine-tuning the properties of place cells and ensuring their stability over time.

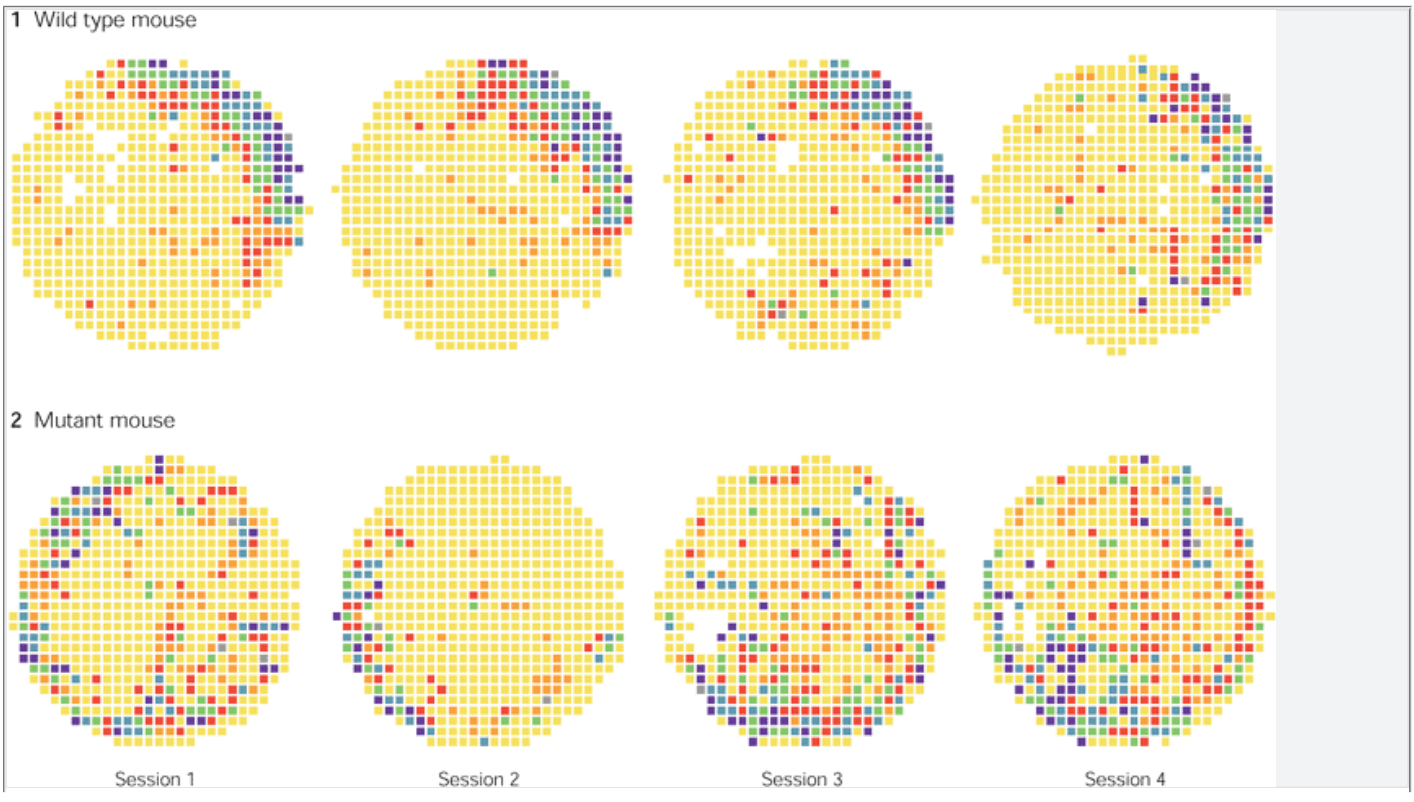


Figure 63-14B The firing-rate patterns from four successive recording sessions in a single cell from a wild-type mouse and a mouse carrying a gene for a persistently active Ca^{2+}/calmodulin-dependent kinase. Before each recording session the animal was taken out of the cylinder and reintroduced into it. In each of the four sessions the positional firing pattern for the wild-type cell is stable. By contrast, the pattern of the mutant cell is unstable in sessions 2 and 3.

For example, pyramidal cells that encode for overlapping positions in space fire synchronously when the animal enters the space represented by those cells. Thomas McHugh and his colleagues found that this correlated firing is lost in mice lacking a functional NMDA receptor in the CA1 region. Alex Rotenberg found that the defect is more severe in mice that overexpress an activated form of the Ca^{2+}/calmodulin-dependent protein kinase. As noted earlier, in wild-type mice a place field forms within minutes after an animal enters a new en-vironment and, once formed, remains stable in that environment for months. In contrast, when mutant mice are removed from a space and then put back into the same space, cells that were previously active in that space form *different* place fields. This instability of place cells is reminiscent of the memory defect in H.M. and other patients with lesions of the medial temporal lobe. Each time these patients enter the same space, they act as if they have never seen it before.

Associative Long-Term Potentiation Is Important for Spatial Memory

If LTP is a synaptic mechanism for maintaining a coherent spatial map over time, then defects in LTP should interfere with spatial memory. One spatial memory test is a water maze in which a mouse must find a platform hidden under opaque water in a pool. When released at random locations around the pool, the mouse must use contextual (spatial) cues—markings on the walls of the room in which the pool is located—to find the platform. This task requires the hippocampus. In a simple noncontextual

P.1268

P.1269

P.1270

P.1271

P.1272

(nonspatial) version of this test the platform is raised above the surface of the water or marked with a flag so that it is visible, permitting the mouse to navigate to the platform directly. This task does not require the hippocampus.

Box 63-1 Restricting Gene Knockout and Regulating Transgene Expression

Biological analysis of learning requires the establishment of a causal relation between specific molecules and learning. This relationship, which has been difficult to demonstrate in mammals, can now be studied in mice either by the use of transgenes or the selective knockout of genes. With *transgenes* a new gene is introduced into the brain under the control of a promoter that expresses that gene in a specific region of the brain. With *gene knockout* specific gene deletions are induced in embryonic stem cells through homologous recombination (see Figure 3-8).

Experiments using transgenes and gene knockout have made it possible to examine the roles of NMDA receptors and different second-messenger kinases in a variety of learning mechanisms in the hippocampus, including long-term potentiation, spatial learning, and the development and maintenance of a cognitive map of space.

With conventional gene knockout techniques animals inherit the genetic deletions in all of their cell types. Global genetic deletion may cause developmental defects that interfere with the later functioning of neural circuits important for memory storage. As a result, interpretation of the results from conventional knockout mice runs into two types of difficulties. First, it is often difficult to exclude the possibility that the abnormal phenotype observed in mature animals results directly or indirectly from a developmental defect. Second, global gene knockout makes it difficult to attribute abnormal phenotypes to a particular type of cell within the brain.

To improve the utility of gene knockout and transgene technology, methods have been developed to restrict gene expression locally or temporally. One method for restricting gene knockout locally exploits the *Cre/loxP* system, a site-specific recombination system, derived from the P1 phage, in which the Cre recombinase catalyzes recombination between 34 bp *loxP* recognition sequences (Figure 63-15). The *loxP* sequences can be inserted into the genome of embryonic stem cells by homologous recombination such that they flank one or more exons of the gene one is interested in. Thus the gene encoding the R1 subunit of the NMDA receptor can be floxed and then ablated selectively in the CA1 region of the hippocampus. Loss of the R1 subunit leads to loss of LTP. Moreover, mice that lack this gene only in the CA1 region of the hippocampus show a deficiency in spatial memory (Figure 63-17). In contrast, these mice show no deficit in tasks that do not involve the hippocampus, such as learning simple visual discrimination.

In addition to regional restriction of gene expression, effective use of genetically modified mice requires control over the timing of gene expression. The ability to turn a transgene on and off gives the investigator an additional degree of flexibility and can exclude the possibility that any abnormality observed in the phenotype of the mature animal is the result of a developmental defect produced by the transgene. This can be done in mice by constructing a gene that can be turned on or off by giving a drug.

One starts by creating two lines of mice (Figure 63-16). One line carries a particular transgene, for example CaMKII-Asp 286, a mutated form of the gene CaMK-II coding for a constitutively active kinase (line 1). Instead of being attached to its normal promoter, the transgene is attached to the promoter *tet-O* that is ordinarily found only in bacteria. This promoter cannot, by itself, turn on the gene; it needs to be activated by a specific transcriptional regulator. Thus, a second transgene expressed in the second line of mice encodes a hybrid transcriptional regulator called *tetracycline transactivator* (tTA) that recognizes and binds to the *tet-O* promoter (line 2). Expression of tTA is placed under the control of a region-specific promoter, such as the promoter for CaMKII.

When the two lines of mice are mated some of the offspring will carry both transgenes. In these mice the tTA binds to the *tet-O* promoter and activates the mutated *CaMKII* gene. This mutant causes abnormalities in long-term potentiation. But when the animal is given doxycycline, the drug binds to the transcription factor tTA causing it to undergo a change in shape that makes it come off the promoter (Figure 63-16). The cell stops expressing *CaMKII* and long-term potentiation returns to normal, demonstrating that the transgene acts on the adult synapse and does not interfere with the development of the synapse.

One can also generate mice that express a mutant form of tTA called reverse tTA (rtTA). This transactivator will *not* bind to *tet-O* unless the animal is fed doxycycline. In this case the transgene is always turned off unless the drug is given.

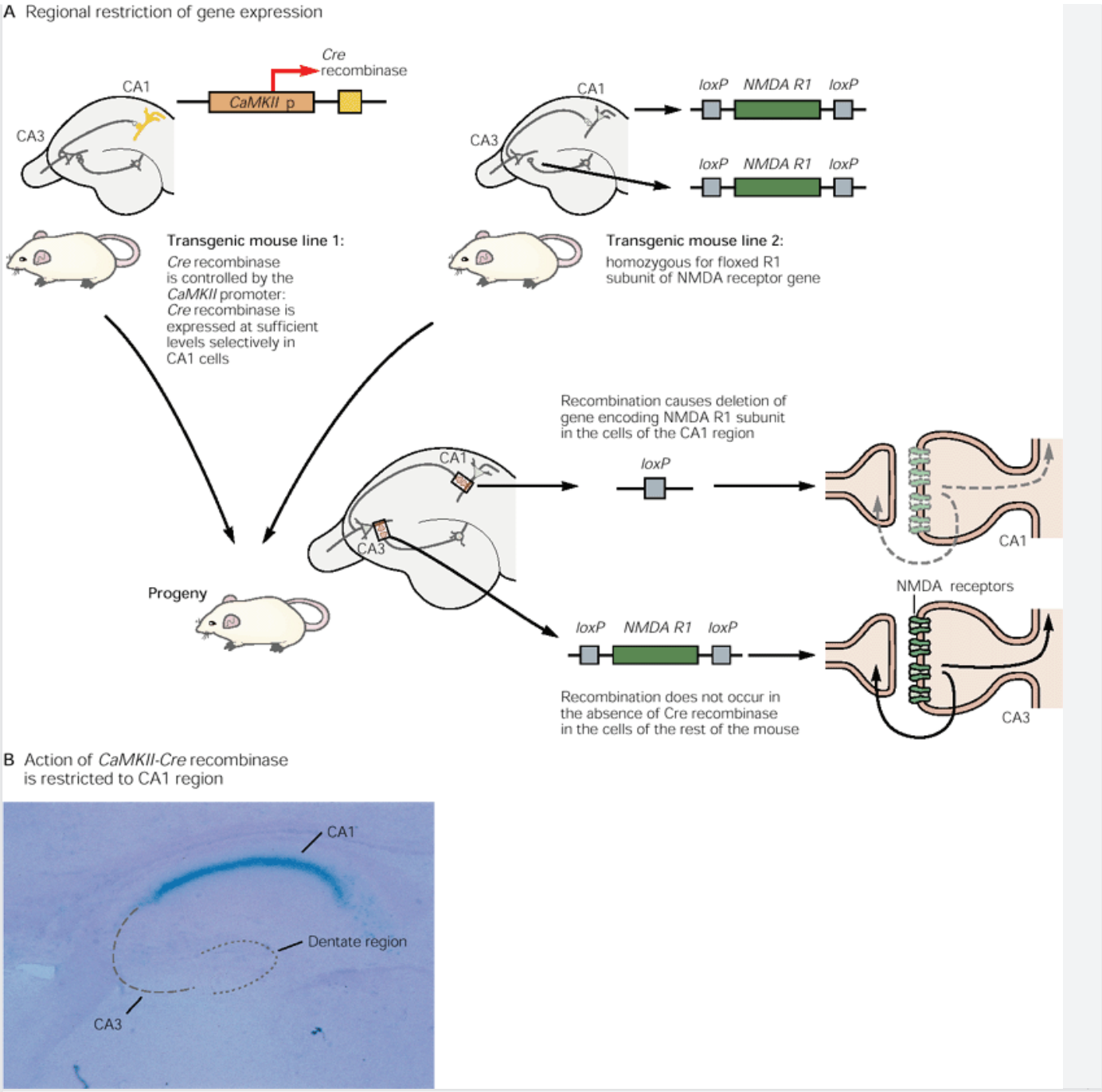


Figure 63-15 (Opposite) Using the *Cre/loxP* system to restrict the region of gene knockout.

A. A mouse homozygous for the floxed R1 subunit of the NMDA receptor (**line 1**) is generated from embryonic stem cells by conventional techniques and is crossed to a second mouse that contains a *Cre* transgene under the control of a transcriptional promoter from the *CamKII* gene (**line 2**). In progeny that are homozygous for the floxed gene and that carry the *Cre* transgene, the floxed gene will be deleted by *Cre/loxP* recombination but only in those cell types in which the *Cre* gene-associated promoter is active. By this means efficient gene knockout is accomplished in postmitotic neurons in a highly restricted manner, limited to the CA1 pyramidal cells of the hippocampus.

B. LacZ staining reveals the region where recombination was successful and led to the removal of a floxed stopper sequence that allows LacZ to be expressed. When the Cre recombinase is combined with the promoter for the Ca^{2+}/calmodulin-dependent kinase, the transgene is restricted to the CA1 region. This is evident in the section of the brain illustrated here. (From Tsien et al. 1996a.)

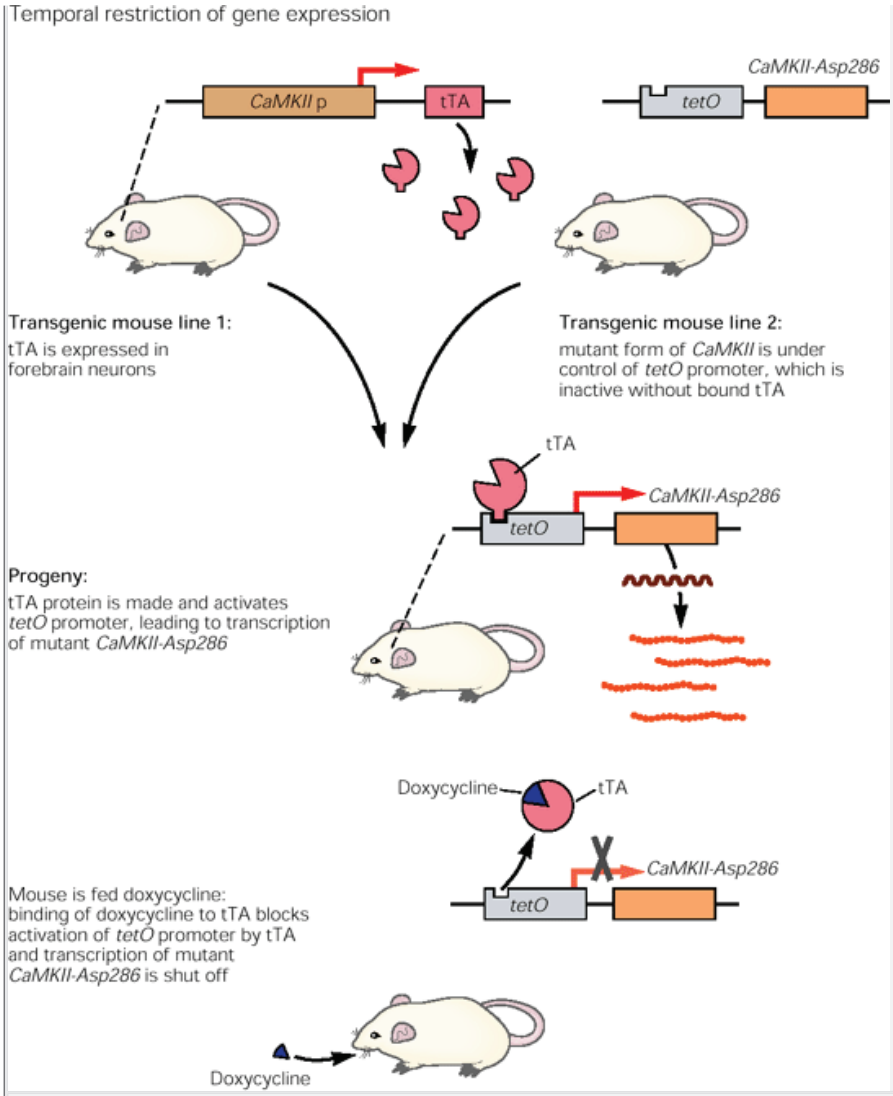
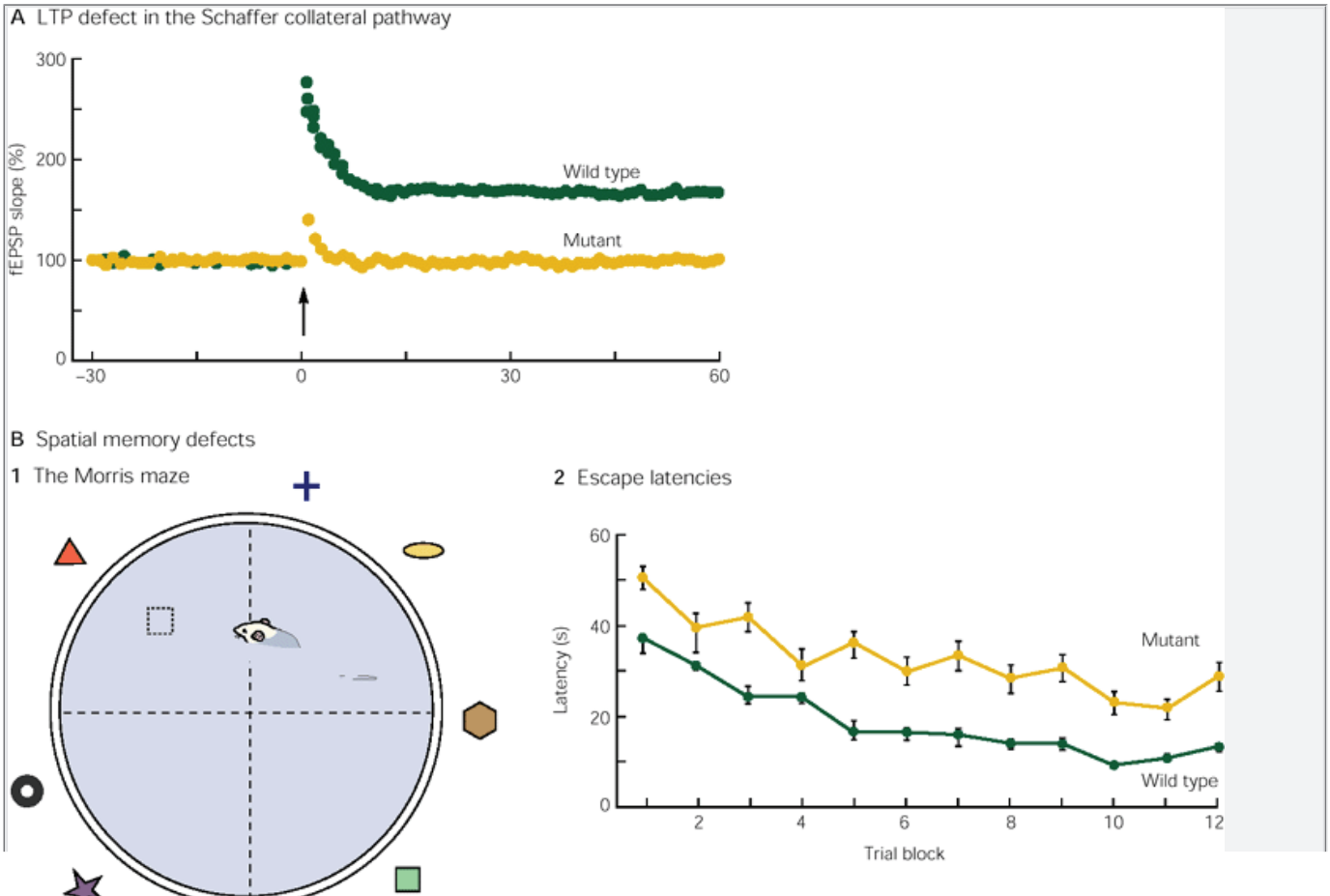


Figure 63-16 Using the tetracycline system to control the timing of gene expression. Two independent lines of transgenic mice are mated so that two transgenes are introduced into a single mouse. In the tetracycline system a bacterial transcription factor, the tetracycline transactivator (tTA), recognizes a bacterial promoter (the *tetO* promoter). When the transactivator binds to the promoter it activates its downstream gene, in this case a constitutively active form of CaMKII, *CaMKII-Asp286*. When the animal is given doxycycline the drug binds to the transcription factor, tTA, producing a conformational change in tTA that causes it to come off the promoter. (From Mayford et al. 1996)



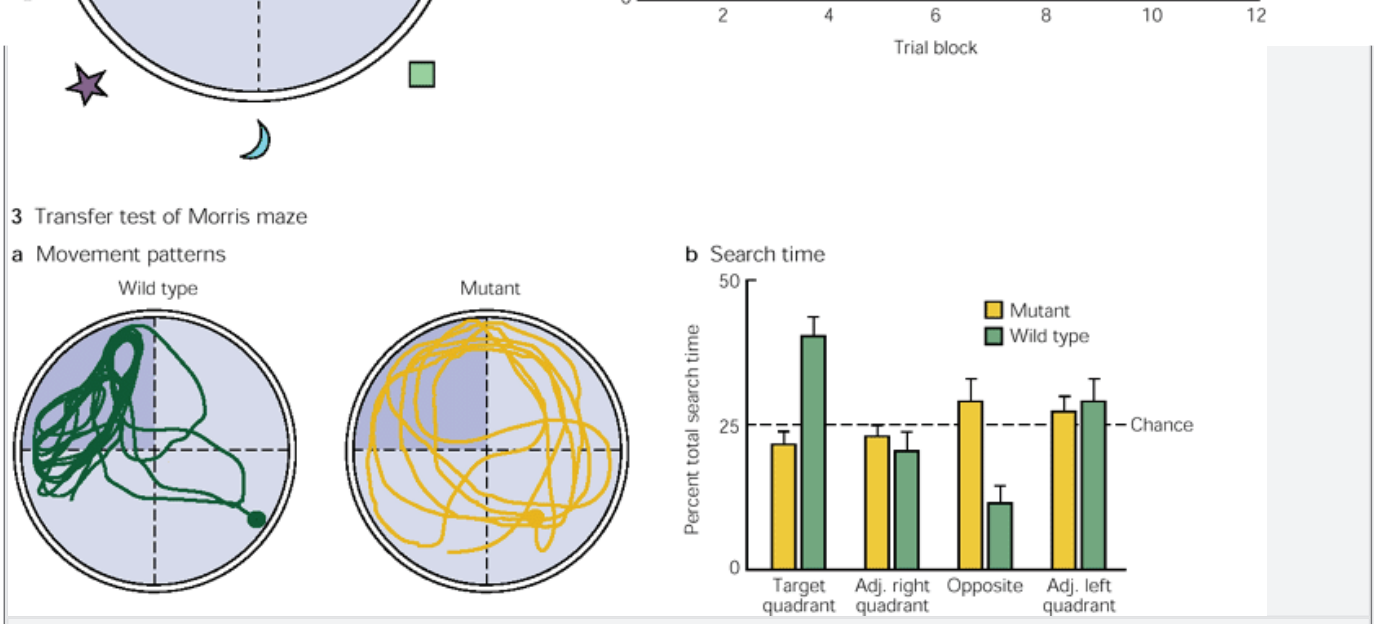


Figure 63-17 (Opposite) Mice that lack the NMDA receptor in the CA1 region of the hippocampus have a defect in LTP and in spatial memory.

A. Field excitatory postsynaptic potentials (fEPSP) were recorded in the stratum radiatum in the CA1 region of both mutant and wild-type mice. After a 30-min period of baseline recording a tetanus (100 Hz for 1 s) was applied (**arrow**). Activity in the pathway remained unchanged in the mutant but became potentiated in the wild-type, indicating that knockout of the NMDA receptor abolishes LTP.

B. Mice that lack the NMDA receptor in CA1 have impaired spatial memory. **1.** In the Morris maze a platform is submerged in one quadrant of an opaque fluid in a circular tank. To avoid remaining in the water the mice have to learn to find the platform and climb onto it. **2.** Mutant mice are slower in learning to find the submerged platform. The graph represents the escape latencies of mice trained to find the hidden platform using spatial (contextual) cues. The mutant mice display a longer latency in every block of trials (four trials per day) than do the wild-type mice. Also, mutant mice do not reach the optimal performance attained by the control mice, even though the mutants show some improvement. **3.** After the mice have been trained in the Morris maze, the platform is taken away and the mice are given a transfer test for memory. A wild-type mouse focuses its search in the quadrant that formerly contained the platform because it remembers the platform as being there. The mutant mouse, which does not remember where the platform was, spends an equal amount of time (25%) in all quadrants. **a.** Illustrative examples. **b.** Actual data. (From Tsien et al 1996b.)

Richard Morris demonstrated that when NMDA receptors are blocked by the injection of a pharmacological antagonist into the hippocampus, the animal can navigate to the visible platform in the noncontextual version of the water maze test but cannot find its way to the submerged platform in the contextual version. These experiments suggested that some mechanism involving NMDA receptors in the hippocampus, perhaps LTP, is involved in spatial learning. More direct evidence for this correlation has come from genetic experiments with the two types of mutant mice described in Box 63-1.

In the first type of mutant the R1 subunit of the NMDA receptor of the pyramidal cells of the CA1 region is selectively knocked out. As a result, basal transmission is normal but LTP is completely disrupted (Figure 63-17A). Although this disruption is restricted to the Schaffer collateral pathway, these mice nevertheless have a serious deficit in spatial memory when tested in the water maze (Figure 63-17B and C). These findings provide the most compelling evidence so far that the NMDA channels and NMDA-mediated synaptic plasticity in the Schaffer collateral pathway are important for spatial memory.

In the second type of mutant, expression of the persistently active form of the Ca^{2+}/calmodulin-dependent protein kinase can be turned on and off at will. Expression of the kinase selectively interferes with LTP when the frequency of electrical stimulation is between 1 and 10 Hz, and this results in an instability in the hippocampal place cells (Figure 63-18). These mutant mice also do not remember spatial tasks. But when the transgene is turned off, the LTP becomes normal and the animal's capability for spatial memory is restored!

These two sets of experiments based on the restricted knockout of the NMDA receptor and the regulated expression of the Ca^{2+}/calmodulin-dependent kinase make it clear that LTP in the Schaffer collateral pathway is important for spatial memory.

How are defects in the various phases of LTP reflected in defects in the phase of memory storage? Indeed, the memory deficits are surprisingly selective. Selective expression in the hippocampus of the transgene that blocks cAMP-dependent protein kinase selectively disrupts the late phase of LTP in the Schaffer collateral pathway. Animals with this deficit have normal learning abilities and normal short-term memory when tested at one hour, but they cannot convert short-term memory into stable long-term memory. They have defective long-term memory when tested at 24 hours after training. Essentially similar results are obtained in mice that have a knockout of several isoforms of CREB-1, as well as in normal wild-type mice that are exposed to pharmacological inhibitors of protein synthesis or camp-dependent protein kinase. These animals learn well and have normal short-term memory but poor long-term memory.

Thus, interference with the early and late component of LTP in the Schaffer collateral pathway interferes with both short- and long-term memory, whereas selective interference with the late component of LTP interferes only with the consolidation of long-term memory. Together, these experiments provide insight into the genetic chain of causation that connects molecules to LTP, LTP to place cells, and place cells to the outward behavior of the animal as reflected in both short- and long-term spatial memory.

Is There a Molecular Alphabet for Learning?

The changes in synaptic efficacy that we have encountered in studies of both implicit and explicit forms of storage raise three key issues in a neurobiological approach to learning.

First, the finding that the molecular mechanisms of some associative forms of synaptic plasticity are based on those of nonassociative forms in the same cell suggests that there may be a molecular alphabet for synaptic plasticity—simpler forms of plasticity might represent elements of more complex forms. Of course, all of these synaptic mechanisms are embedded in distributed neural circuits with considerable additional computational power, which can thus add substantial complexity to the actions of individual cells.

Second, the molecular mechanisms of elementary forms of associative memory storage used in both implicit and explicit learning are similar. In the two processes we have considered—activity-dependent presynaptic facilitation for storing implicit memory and associative LTP for storing explicit memory—the plasticity of neuronal function seems to derive from the associative properties of specific proteins: from the ability of proteins, such as adenylyl cyclase and the NMDA receptor, to respond conjointly to two independent signals.

Finally, despite clear differences in behavioral logic in the neural systems recruited for the task, implicit and explicit memory storage seem to use elements of a

common genetic switch, involving cAMP-dependent protein kinase, MAP kinase, and CREB, to convert labile short-term memory into long-term memory.

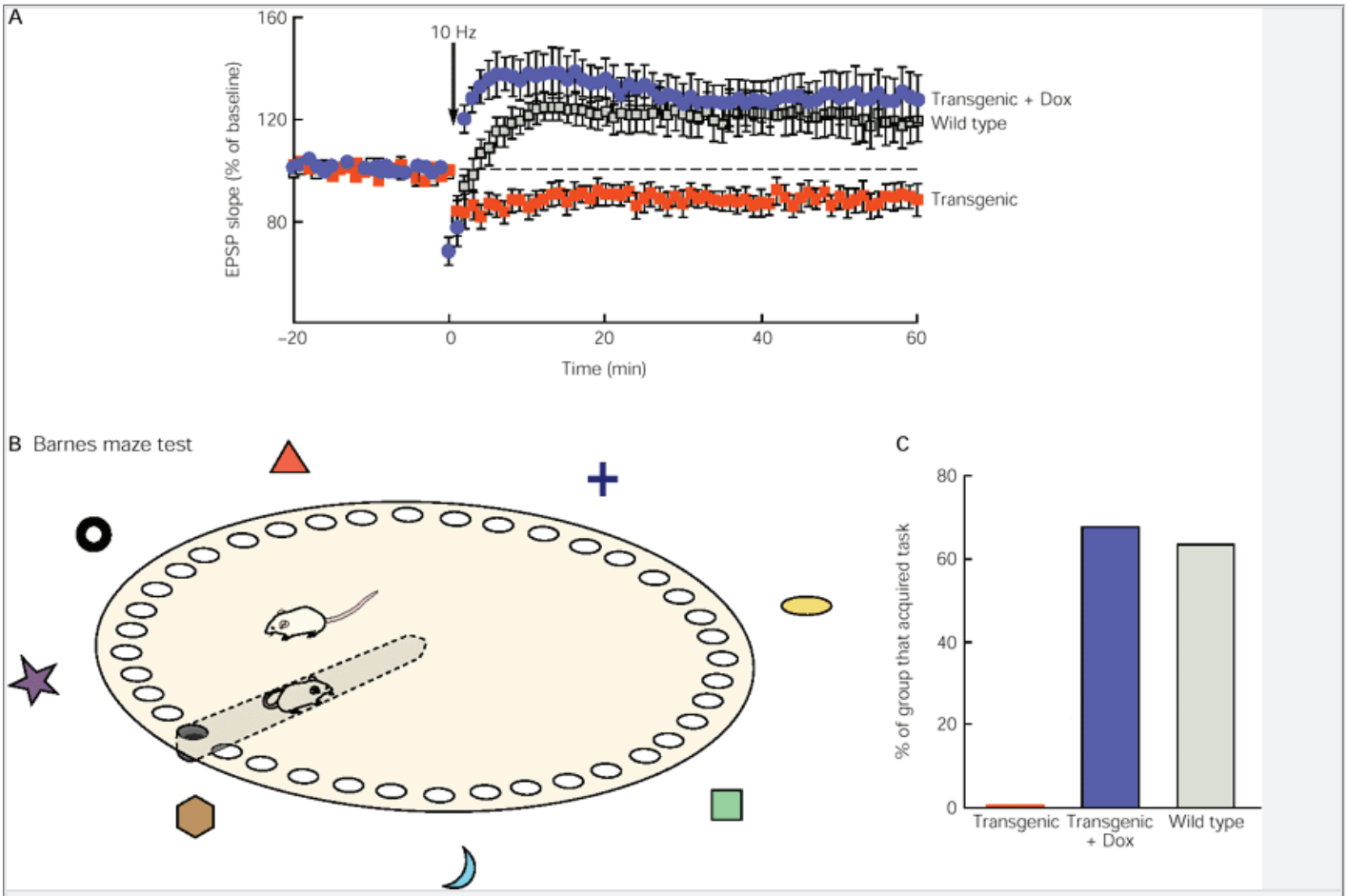
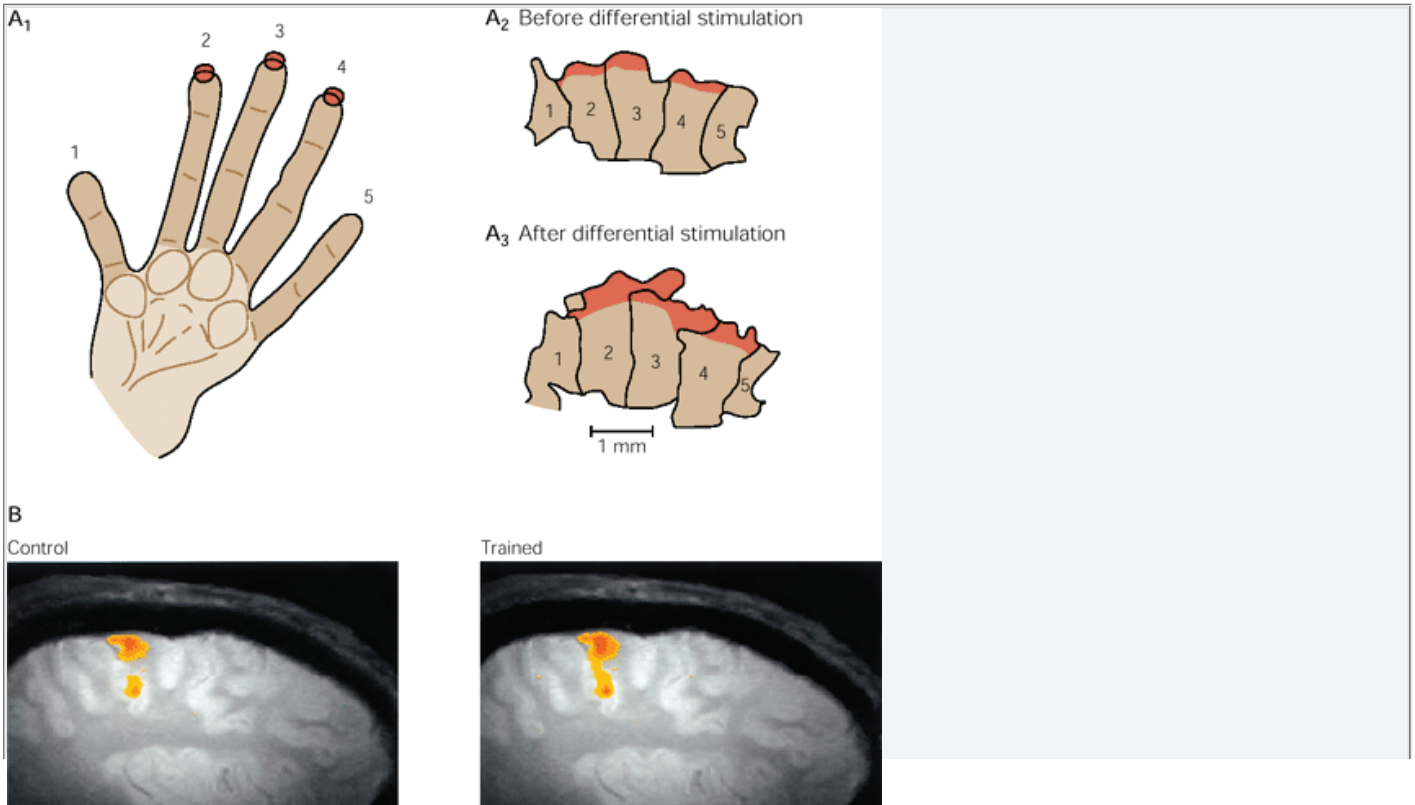


Figure 63-18 Reversal of the deficit in LTP and in spatial memory in the CA1 region of the hippocampus.

A. Field excitatory postsynaptic potentials (fEPSP) slopes before and after tetanic stimulation were recorded and expressed as the percentage of pretetanus baseline. Stimulation at 10 Hz (**arrow**) for 1.5 min induced a transient depression in activity followed by potentiation in wild-type mice but only a slight depression and no potentiation in transgenic mice. Doxycycline treatment (Dox) reversed the defect in the transgenic mice so that they show potentiation compared with wild type.

B. The Barnes maze measures the same spatial memory system explored in the Morris water maze. This maze consists of a platform with 40 holes, only one of which (marked in **black**) leads to an escape tunnel that allows the mouse to exit the platform. The mouse is placed in the center of the platform. Because mice do not like open, well-lit spaces, they try to escape from the platform. The only way they can do that is to find the one hole that leads to the escape tunnel. The most efficient way of doing that (and the only way of meeting the criteria set for the task by the experimenter) is to use the distinctive markings on the four walls.

C. Percentage of transgenic and wild-type mice that met the learning criterion on the Barnes circular maze. Three groups of mice were tested: transgenics with and without doxycycline and wild-types. Transgenics that received doxycycline performed as well as wild-types, whereas those without the doxycycline did not learn the task.



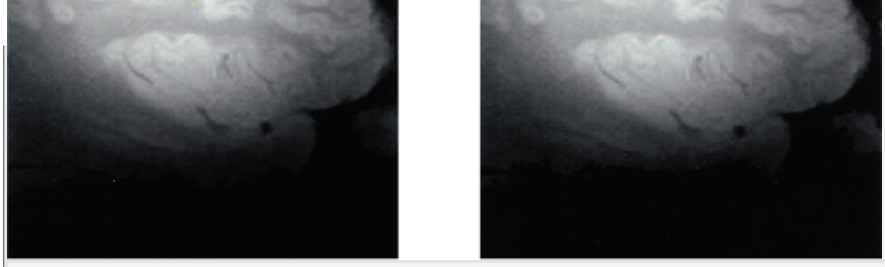


Figure 63-19 Training expands the existing representation of the fingers in the cortex.

A. A monkey was trained for 1 hour per day to perform a task that required repeated use of the tips of fingers 2, 3, and occasionally 4. After a period of repeated stimulation the portion of area 3b representing the tips of the stimulated fingers was substantially greater (**3**) than in untrained monkeys (**2)** measured 3 months before training). (Adapted from Jenkins et al. 1990.)

B. A human subject trained to do a rapid sequence of finger movements will improve in accuracy and speed after three weeks of daily (10-20 min) training. In these MRI scans of local blood oxygenation level-dependent signals in the primary motor cortex, the region activated in trained subjects (after three weeks of training) is larger than the region activated in control subjects who performed unlearned finger movements in the same hand. The change in cortical representation with the learned motor sequence persisted for several months.

Changes in the Somatotopic Map Produced by Learning May Contribute to the Biological Expression of Individuality

Learning, we have seen, can lead to structural alterations in the brain. How common are these changes in determining the functional architecture of the mature brain? This question has been examined in the somatosensory cortex.

The four maps of the body surface in the primary somatic sensory cortex vary among individuals in a manner that reflects different usage of sensory pathways. The connections of afferent pathways in the cortex can expand or retract depending on activity (see Chapter 20). Reorganization of afferent inputs is also evident at lower levels in the brain, specifically at the level of the dorsal column nuclei, which contain the first synapses of the somatic sensory system. Therefore, organizational changes probably occur throughout the somatic afferent pathway as well. As we grow up each of us is exposed to different combinations of stimuli and develops motor skills in different ways. Thus all brains—even the brains of identical twins who share the

same genes—are uniquely modified by experience. This distinctive modification of brain architecture, along with a unique genetic makeup, constitutes a biological basis for individuality.

The process by which experience alters the somatosensory maps in the cortex is illustrated in an experiment in which adult monkeys were trained to use only their middle three fingers to obtain food. After several thousand trials of this behavior the area of cortex devoted to the middle finger expanded greatly (Figure 63-19). Practice alone therefore may strengthen the effectiveness of existing patterns of connections.

Experiments such as this one suggest that early development of the input connections to cortical neurons in the somatic sensory system may depend on correlated activity in different afferent axons. As we have seen, cooperative activity in different afferent fibers stabilizes the development of ocular dominance columns in the visual system (Chapter 56). In turn, the use of correlated firing to fine-tune cortical connections may depend on mechanisms that require correlated firing similar to those of LTP. For example, when the skin of two adjacent fingers of a monkey were surgically connected to ensure that the connected fingers were always used together, the correlation of inputs from the joined fingers abolished the normally sharp discontinuity between areas in the somatosensory cortex that receive inputs from these digits (see Figure 19-8).

Neuronal Changes Associated With Learning Provide Insights Into Psychiatric Disorders

The demonstration that learning produces changes in the effectiveness of neural connections has revised our view of the relationship between social and biological processes in the shaping of behavior. Until recently the majority view in medicine and psychiatry was that biological and social determinants of behavior act on separate levels of the mind. For example, psychiatric illnesses were traditionally classified as either organic or functional. Organic mental illnesses included the dementias, such as Alzheimer disease, and the toxic psychoses, such as those that follow the chronic use of alcohol or cocaine. Functional mental illnesses included the various depressive syndromes, the schizophrenias, and the neuroses.

This distinction dates to the nineteenth century, when neuropathologists examined the brains of patients coming to autopsy and found gross and readily demonstrable distortions in the architecture of the brain in some psychiatric diseases but not in others. Diseases that produce anatomical evidence of brain lesions were called *organic;* those lacking these features were called *functional.*

The experiments reviewed in this chapter show that this distinction is no longer tenable. Everyday events—sensory stimulation, deprivation, and learning—can effectively weaken synaptic connections in some circumstances and strengthen them in others. We no longer think that only certain diseases ("organic diseases") affect mentation through biological changes in the brain. The basis of contemporary neural science is that all mental processes are biological and therefore any alteration in those processes is necessarily organic.

In the attempt to understand the biological basis of a particular mental illness we must ask, to what degree is this biological process determined by genetic and developmental factors? To what degree is it determined by a toxic or infectious agent, or by a developmental abnormality? To what degree is it socially determined?

Even those mental disturbances that are considered most heavily determined by social factors must have a biological aspect, since it is the activity of the brain that is being modified. Insofar as social intervention works, whether through psychotherapy, counseling, or the support of family or friends, it must work by acting on the brain, quite likely on the strength of connections between nerve cells. Just because structural changes may not initially be detected following cognitive therapy does not rule out the possibility that important biological changes are nevertheless occurring. They may simply be below the level of detection with the limited techniques available to us.

Demonstrating the biological nature of mental functioning requires more sophisticated anatomical methodologies than the light-microscopic histology of nineteenth century pathologists. We must develop a neuropathology of mental illness based on functional as well as structural analysis. The new imaging techniques—positron emission tomography and functional magnetic resonance imaging, among others—have opened the door to the noninvasive exploration of the human brain on a cell-biological level, the level of resolution that is required to understand the physical mechanisms of mentation and therefore of mental disorders. This approach is now being pursued in the study of schizophrenia and depression

It is intriguing to think that insofar as psychotherapy is successful in changing behavior, it may do so by producing alterations in gene expression. If so, successful psychotherapeutic treatment of neurosis or character disorders should also produce structural changes in the nervous system. Thus, we face the attractive possibility that as brain imaging techniques improve, these techniques might ultimately be useful not only for diagnosing various neurotic illnesses but also for monitoring the progress of psychotherapy.

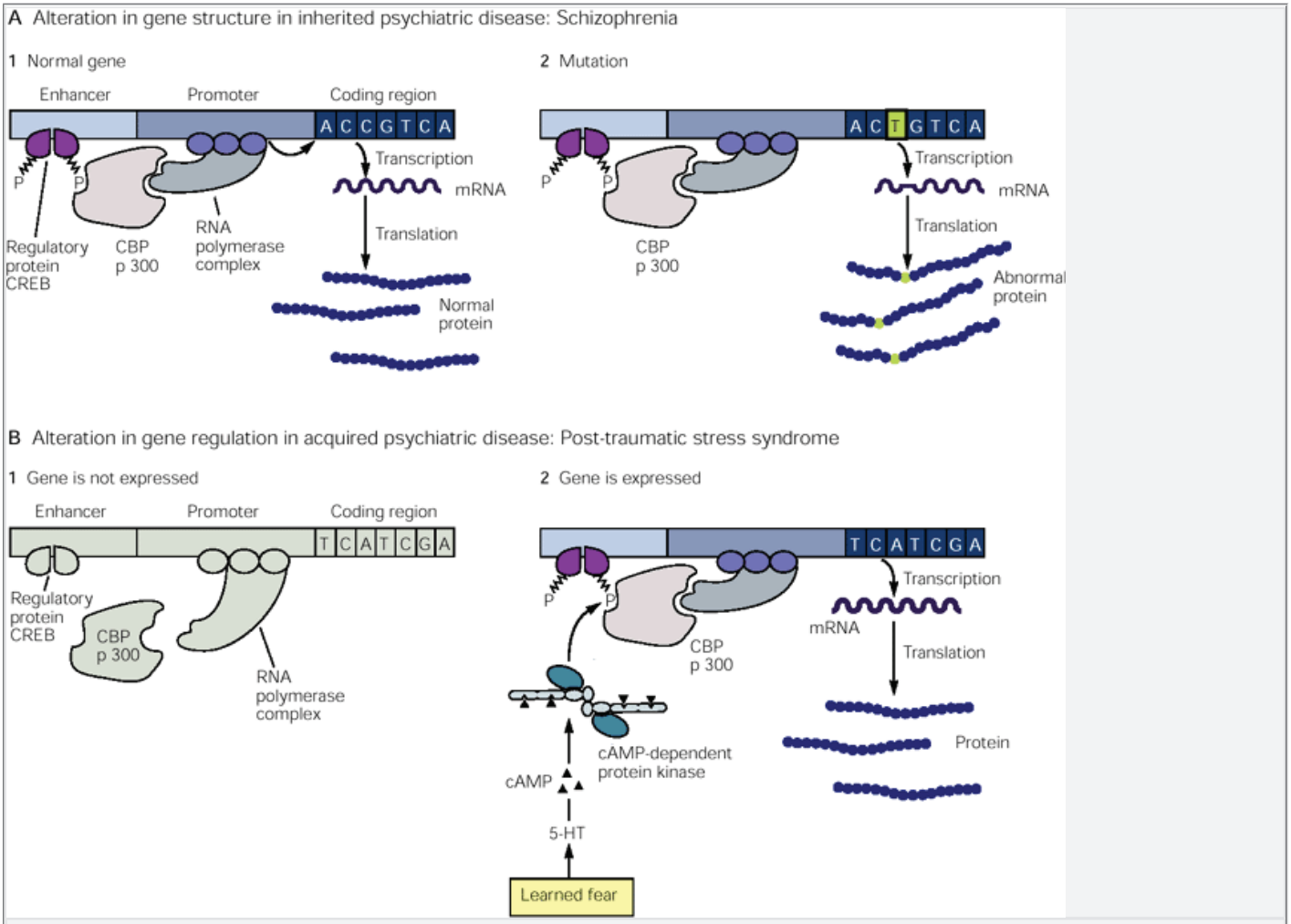


Figure 63-20 Genetic and acquired illnesses both have a genetic component. Genetic illnesses result from the expression of altered genes or allelic variants, whereas acquired illnesses (*neuroses*) involve the modulation of gene expression by environmental stimuli as a result of learning, leading to the transcription of a previously inactive gene. The gene is illustrated as having two segments. A coding region is transcribed into a messenger RNA (mRNA) by an RNA polymerase. The mRNA in turn is translated into a specific protein. A regulatory segment consists of an enhancer region and a promoter region (see Chapter 13, Box 13-1). In this example the RNA polymerase can transcribe the gene when the regulatory protein binds to the enhancer region. For the regulatory protein must first be phosphorylated before it can act on the enhancer. Only when it is phosphorylated can CREB bind the adapter protein, the *CREB binding protein* (CBP), which allows the phosphorylated form of CREB to interact with the transcriptional machinery.

A. Inherited disease such as schizophrenia. **1.** Under normal conditions the phosphorylated regulatory protein binds to the enhancer segment, thereby activating the transcription of the structural gene, leading to the production of the normal protein. **2.** A mutant form of the coding region of the structural gene, in which a thymidine (T) has been substituted for a cytosine (C), leads to transcription of an altered messenger RNA. This in turn produces an abnormal protein, giving rise to the disease state. This alteration in gene structure becomes established in the germ line, is heritable, and contributes to the disease.

B. Acquired disease such as post-traumatic stress syndrome. **1.** If the regulatory protein for a normal structural gene is not phosphorylated, it cannot bind to the promoter site and thus gene translation cannot be initiated. **2.** In this case a specific experience leads to the activation of serotonin (5-HT) and cAMP, which activate the cAMP-dependent protein kinase. The catalytic unit translocates to the nucleus and phosphorylates the regulatory protein CREB, which then can bind to the enhancer segment and thus initiate gene transcription. By this means an abnormal learning experience could lead to the expression of a protein that gives rise to symptoms of a neurotic disorder.

P.1276

P.1277

In studying the specific molecular changes that underlie memory storage, we should look for altered gene expression in abnormal as well as normal mental states. There is now substantial evidence that the susceptibility to major psychotic illnesses—schizophrenia and manic-depressive disorders—is heritable and is due to allelic variations. The cell-biological data on learning and long-term memory reviewed here suggest that neurotic illnesses, acquired by learning, also are likely to involve alterations of gene expression but alterations due to disordered regulation (Figure 63-20).

Development, hormones, stress, drug addiction, alcoholism, and learning are all factors that alter gene expression by modifying the binding of transcriptional regulatory proteins to each other and to the regulatory regions of genes. It is likely that at least some neurotic illnesses (or components of them), as well as various forms of drug addiction, result from reversible defects in gene regulation.

An Overall View

Studies on synaptic plasticity suggest there are two overlapping stages in the development and maintenance of synaptic strength. The first stage, the initial steps of synapse formation, occurs primarily early in development and is under the control of genetic and developmental processes. The second stage, the fine-tuning of developed synapses by experience, begins during late stages of development and continues to some degree throughout life. An attractive possibility is that the activity-dependent cellular mechanisms involved in the associative learning of the mature organism may be similar to the activity-dependent mechanisms at work during critical periods of development.

The ongoing modification of synapses throughout life means that all behavior of an individual is produced by genetic and developmental mechanisms acting on the brain—that everything the brain produces, from the most private thoughts to the most public acts, should be understood as a biological process. Environmental factors and learning bring out specific capabilities by altering either the effectiveness or the anatomical connections of existing pathways.

The synthesis of neurobiology, cognitive psychology, neurology, and psychiatry that we have emphasized throughout this book is filled with promise. Modern cognitive psychology has shown that the brain stores an internal representation of the world, while neurobiology has shown that this representation can be understood in terms

of individual nerve cells and their interconnections. This synthesis has given us a better perspective on perception, action, learning, and memory. It also provides us with a profound new biological insight into the nature of mental illnesses.

Although early behaviorist psychology led the way in exploring observable aspects of behavior, advances in modern cognitive psychology indicate that investigations that fail to consider brain mechanisms cannot adequately explain behavior. As recently as 10 years ago it would have been difficult to establish the existence of an internal representation directly, since internal mental processes were essentially inaccessible to experimental analysis. However, as we have seen throughout this book, cell biology, molecular biology, and brain imaging have now made biological experiments on elementary aspects of internal mental processes feasible.

Contrary to the expectations of some, biological analysis is unlikely to diminish our fascination with thinking or to make thinking trivial by reduction when we frame the issues in terms of molecular biology. Rather cell biology and molecular biology have expanded our vision, allowing us to perceive previously unanticipated interrelationships between biological and psychological phenomena.

The boundary between cognitive psychology and neural science is arbitrary and always changing. It has been imposed not by the natural contours of the disciplines, but by lack of knowledge. As our knowledge expands, the biological and behavioral disciplines will merge at certain points; it is at these points that our understanding of mentation will rest on more secure ground. As we have tried to illustrate in this book, the merger of biology and cognitive psychology is more than a sharing of methods and concepts. The joining of these two disciples represents the emerging conviction that scientific descriptions of mentation at several different levels will all eventually contribute to a unified biological understanding of behavior.

Selected Readings

Abel T, Nguyen PV, Barad M, Deuel TAS, Kandel ER, Bourtchouladze R. 1997. Genetic demonstration of a role for PKA in the late phase of LTP and in hippocampal-based long-term memory. Cell 88:615–626.

Dudai Y. 1989. *The Neurobiology of Memory: Concepts, Findings, Trends.* Oxford: Oxford Univ. Press.

Hawkins RD, Kandel ER, Siegelbaum SA. 1993. Learning to modulate transmitter release: themes and variations in synaptic plasticity. Annu Rev Neurosci 16:625–665.

Lisman J. 1994. The CaM kinase II hypothesis for the storage of synaptic memory. Trends Neurosci 17:406–412.

Merzenich MM, Recanzone EG, Jenkins WM, Allard TT, Nudo RJ. 1988. Cortical representational plasticity. In: PR Rakic, WR Singer (eds). *Neurobiology of Neocortex*, pp. 41-67. New York: Wiley.

Squire LR, Kandel ER. 1999. Memory: mind to molecules. New York: Sci Am Lib 1.

Tully T, Bolwig G, Christensen J, Connolly M, Del Vecchio J, DeZazzo J, Dubnau J, James C, Pinto S, Regulski M, Svedberg B, Velinzon K. 1997. A return of genetic dissection of memory in *Drosophila. Function and Dysfunction in the Nervous System.* Cold Spring Harb Symp Quant Biol 61:207–218.

References

Bailey CH, Chen MC. 1983. Morphological basis of long-term habituation and sensitization in *Aplysia.* Science 220:91–93.

Bliss TVP, Lom•' T. 1973. Long-lasting potentiation of synaptic transmission in the dentate gyrus of the anesthetized rabbit following stimulation of the perforant path. J Physiol (Lond) 232:331–356.

Bonhoeffer T, Staiger V, Aertsen A. 1989. Synaptic plasticity in rat hippocampal slice cultures: local "Hebbian" conjunction of pre- and postsynaptic stimulation leads to distributed synaptic enhancement. Proc Natl Acad Sci U S A 86:8113–8117.

Bolshakov VY, Golan H, Kandel ER, Siegelbaum SA. 1997. Recruitment of new sites of synaptic transmission during the cAMP-dependent late phase of LTP at CA3-CA1 synapses in the hippocampus. Neuron 19:635–651.

Bourtchouladze R, Frenguelli B, Blendy J, Cioffi D, Schutz G, Silva A. 1994. Deficient long-term memory in mice with a targeted mutation of the cAMP responsive element-binding protein. Cell 79:59–68.

Castellucci VF, Carew TJ, Kandel ER. 1978. Cellular analysis of long-term habituation of the gill-withdrawal reflex in *Aplysia californica.* Science 202:1306–1308.

Chain DG, Casadio A, Schacher S, Hedge AN, Valbrun M, Yamamoto N, Goldberg AL, Bartsch D, Kandel ER, Schwartz JH. 1999. Mechanisms for generating the autonomous cAMP-dependent protein kinase for long-term facilitation in *Aplysia.* Neuron 22:147–156.

Frost WN, Castellucci VF, Hawkins RD, Kandel ER. 1985. Monosynaptic connections made by the sensory neurons of the gill and siphon withdrawal reflex in *Aplysia* participate in the storage of long-term memory for sensitization. Proc Natl Acad Sci U S A 82:8266–8269.

Grant SGN, O'Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER. 1992. Impaired long-term potentiation, spatial learning, and hippocampal development in *fyn* mutant mice. Science 258:1903–1910.

Gustafsson B, Wigström H, Abraham WC, Huang Y-Y. 1987. Long-term potentiation in the hippocampus using depolarizing current as the conditioning stimulus to single volley synaptic potential. J Neurosci 7:774–780.

Hawkins RD, Abrams TW, Carew TJ, Kandel ER. 1983. A cellular mechanism of classical conditioning in *Aplysia*: activity-dependent amplification of presynaptic facilitation. Science 219:400–405.

Hebb DO. 1949. *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley.

Hegde AN, Inokuchi K, Pei W, Casadio A, Ghirardi M, Chain DG, Martin KC, Kandel ER, Schwartz JH. 1997. Ubiquitin C-terminal hydrolase is an immediate-early gene essential for long-term facilitation in *Aplysia.* Cell 89:115–126.

Huang Y-Y, Kandel ER. 1994. Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of hippocampus requires repeated tetanization. Learning & Memory 1:74–82.

Huang Y-Y, Kandel ER, Varshavsky L, Brandon EP, Qi M, Idzerda RL, McKnight GS, Bourtchouladze R. 1995. A genetic test of the effects of mutations in PKA on mossy fiber LTP and its relation to spatial and contextual learning. Cell 83:1211–1222.

Huang Y-Y, Li X-C, Kandel ER. 1994. cAMP contributes to mossy fiber LTP by initiating both a covalently-mediated early phase and a macromolecular synthesis-dependent late phase. Cell 79:69–79.

Jenkins WM, Merzenich MM, Ochs MT, Allard T, Guic-Robles E. 1990. Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. J Neurophsiol 63:82–104.

Kandel ER. 1989. Genes, nerve cells, and the remembrance of things past. J Neuropsychiatry 1:103–125.

Lin XY, Glanzman DL. 1994. Long-term potentiation of *Aplysia* sensorimotor synapses in cell culture: regulation by postsynaptic voltage. London: Proc R Soc 255:113–118.

Malinow R, Tsien RW. 1990. Presynaptic enhancement shown by whole cell recordings of LTP in hippocampal slices. Nature 357:134–139.

Mayford M, Bach ME, Huang Y-Y, Wang L, Hawkins RD, Kandel ER. 1996. Control of memory formation through regulated expression of a CaMKII transgene. Science 274:1678–1683.

McHugh TJ, Blum KI, Tsien JZ, Tonegawa S, Wilson MA. 1996. Impaired hippocampal representation of space in CA1-specific NMDAR1 knockout mice. Cell 87:1339–1349.

Muller RU, Kubie JL, Ranck JB Jr. 1987. Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. J Neurosci 7:1935–1950.

O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. Brain Res 34:171–175.

Pavlov IP. 1927. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.* GV Anrep (transl). London: Oxford Univ. Press.

Rotenberg A, Mayford M, Hawkins RD, Kandel ER, Muller RU. 1996. Mice expressing activated CaMKII lack low frequency LTP and do not form stable place cells in the CA1 region of the hippocampus. Cell 87:1351–1361.

Schuman EM, Madison DV. 1994. Locally distributed synaptic potentiation in hippocampus. Science 263:532–536.

Schuman EM, Madison DV. 1991. A requirement for the intercellular messenger nitric oxide in long-term potentiation. Science 254:1503–1506.

Sherrington CS. 1906. *The Integrative Action of the Nervous System.* New Haven, CT: Yale Univ. Press.

Silva AJ, Paylor R, Wehner JM, Tonegawa S. 1992. Impaired spatial learning in α-calcium-calmodulin kinase II mutant mice. Science 257:206–211.

Silva AJ, Stevens CF, Tonegawa S, Wang Y. 1992. Deficient hippocampal long-term potentiation in α-calcium-calmodulin kinase II mutant mice. Science 257:201–206.

Spencer AW, Thompson RF, Nielson DR Jr. 1966. Response decrement of the flexion reflex in the acute spinal cat and transient restoration by strong stimuli. J Neurophysiol 29:240–252.

Stevens C, Wang Y. 1994. Changes in reliability of synaptic function as a mechanism for plasticity. Nature 371:704–707.

Thompson RF, Spencer WA. 1966. Habituation: a model of phenomenon for the study of neural substrates of behavior. Psychol Rev 173:16–43.

Tsien JZ, Chen DF, Gerber D, Tom C, Mercer EH, Anderson DJ, Mayford M, Kandel ER, Tonegawa S. 1996a. Subregion- and cell type-restricted gene knockout in mouse brain. Cell 87:1317–1326.

Tsien JZ, Huerta PT, Tonegawa S. 1996b. The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. Cell 87:1327–1338.

Tsien RW, Malinow R. 1990. Long-term potentiation: Presynaptic enhancement following postsynaptic activation of Ca^{2+}-dependent protein kinases. Cold Spring Harb Symp Quant Biol 55:147–159.

Weisskopf MG, Castillo PE, Zalutsky RA, Nicoll RA. 1994. Mediation of hippocampal long-term potentiation by cyclic AMP. Science 265:1878–1882.

Yin JCP, Wallach JS, Del Vecchio M, Wilder EL, Zhuo H, Quinn WG, Tully T. 1994. Induction of a dominant negative CREB transgene specifically blocks long-term memory in *Drosophila.* Cell 79:49–58.

Zalutsky RA, Nicoll RA. 1990. Comparison of two forms of long-term potentiation in single hippocampal neurons. Science 248:1619–1624.

Resources

Appendices

Appendix A

Current Flow in Neurons

John Koester

THIS SECTION REVIEWS THE basic principles of electrical circuit theory. Familiarity with this material is important for understanding the equivalent circuit model of the neuron developed in Chapters 5,6,7,8,9. The section is divided into three parts:

- The definition of basic electrical parameters.
- A set of rules for elementary circuit analysis.
- Adescription of current flow in circuits with capacitance.

Definition of Electrical Parameters

Potential Difference (*V* or *E*)

Electrical charges exert an electrostatic force on other charges: Like charges repel, opposite charges attract. As the distance between two charges increases, the force that is exerted decreases. Work is done when two charges that initially are separated are brought together:

Negative work is done if their polarities are opposite, and *positive work* if they are the same. The greater the values of the charges and the greater their initial separation, the greater the work that is done (work $\int_r^0 f(r)\,dr$, where f is electrostatic force and r is the initial distance between the two charges). Potential difference is a measure of this work. The potential difference between two points is the work that must be done to move a unit of positive charge (1 coulomb) from one point to the other, ie, it is the potential energy of the charge. One volt (V) is the energy required to move 1 coulomb a distance of 1 meter against a force of 1 newton.

Current (*I*)

A potential difference exists within a system whenever positive and negative charges are separated. Charge separation may be generated by a chemical reaction (as in a battery) or by diffusion between two electrolyte solutions with different ion concentrations across a permeability-selective barrier, such as a cell membrane. If a region of charge separation exists within a conducting medium, then charges move between the areas of potential difference: positive charges are attracted to the

region with a more negative potential, and negative charges go to the regions of positive potential. The resulting movement of charges is current flow, which is defined as the net movement of positive charge per unit time. In metallic conductors current is carried by electrons, which move in the opposite direction of current flow. In nerve and muscle cells current is carried by positive and negative ions in solution. One ampere (A) of current represents the movement of 1 coulomb (of charge) per second.

Conductance (*g*)

Any object through which electrical charges can flow is called a conductor. The unit of electrical conductance is the siemen (S). According to Ohm's law, the current that flows through a conductor is directly proportional to the potential difference imposed across it.[1]

$$I = V \times g$$

$$\text{Current (A)} = \text{Potential difference (V)} \times \text{Conductance (S)}.$$

As charge carriers move through a conductor, some of their potential energy is lost; it is converted into thermal energy due to the frictional interactions of the charge carriers with the conducting medium.

Each type of material has an intrinsic property called conductivity (σ), which is determined by its molecular structure. Metallic conductors have very high conductivities; they conduct electricity extremely well. Aqueous solutions with high ionized salt concentrations have somewhat lower values of σ, and lipids have very low conductivities—they are poor conductors of electricity and are therefore good insulators. The conductance of an object is proportional to σ times its crosssectional area, divided by its length:

$$g = (\sigma) \times \frac{\text{Area}}{\text{Length}}$$

The length dimension is defined as the direction along. which one measures conductance.

For example, the conductance measured along the cytoplasmic core of an axon is reduced if its length is increased or its diameter decreased (Figure A-1).

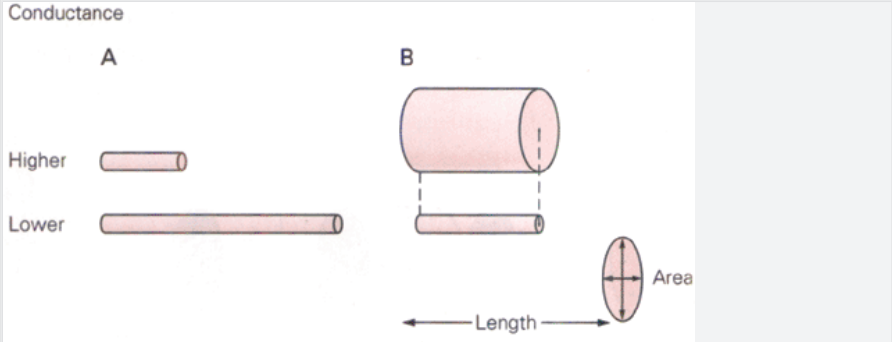


Figure A-1 Geometric factors, together with intrinsic conductivity, determine the conductance of an object.

A. Conductance is inversely proportional to the length of a conductor.

B. Conductance is directly proportional to the cross-sectional area of a conductor.

Electrical resistance (R) is the reciprocal of conductance, and is a measure of the resistance provided by an object to current flow. Resistance is measured in ohms (Ω):

$$1 \text{ ohm} = (1 \text{ siemen})^{-1}.$$

Capacitance (*C*)

A capacitor consists of two conducting plates separated by an insulating layer. The fundamental property of a capacitor is its ability to store charges of opposite sign: positive charge on one plate, negative on the other.

A capacitor made up of two parallel plates with its two conducting surfaces separated by an insulator (anair gap) is shown in Figure A-2A. There is a net excess of positive charges on plate x, and an equal number of excess negative charges on plate y, resulting in a potential difference between the two plates. One can measure thispotential difference by determining how much work is required to move a positive test charge from the surface of y to that of x. Initially, when the test charge is at y, it is attracted by the negative charges on y, and repelled less strongly by the more distant positive charges on x. The result of these electrostatic interactions is a force f that opposes the movement of the test charge from y to x. As the test charge is moved to the left across the gap, the attraction by the negative charges on y diminishes, but the repulsion by the positive charges on x increases, with the result that the net electrostatic force exerted on the test charge is constant everywhere between x and y (Figure A-2A). Work (W) is force (f) times the distance (D) over which the force is exerted:

$$W = f \times D.$$

Therefore, it is simple to calculate the work done in moving the test charge from one side of the capacitor to the

other. It is the shaded area under the curve in Figure A-2A. This work is equal to the difference in electrical potential energy, or potential difference, between x and y.

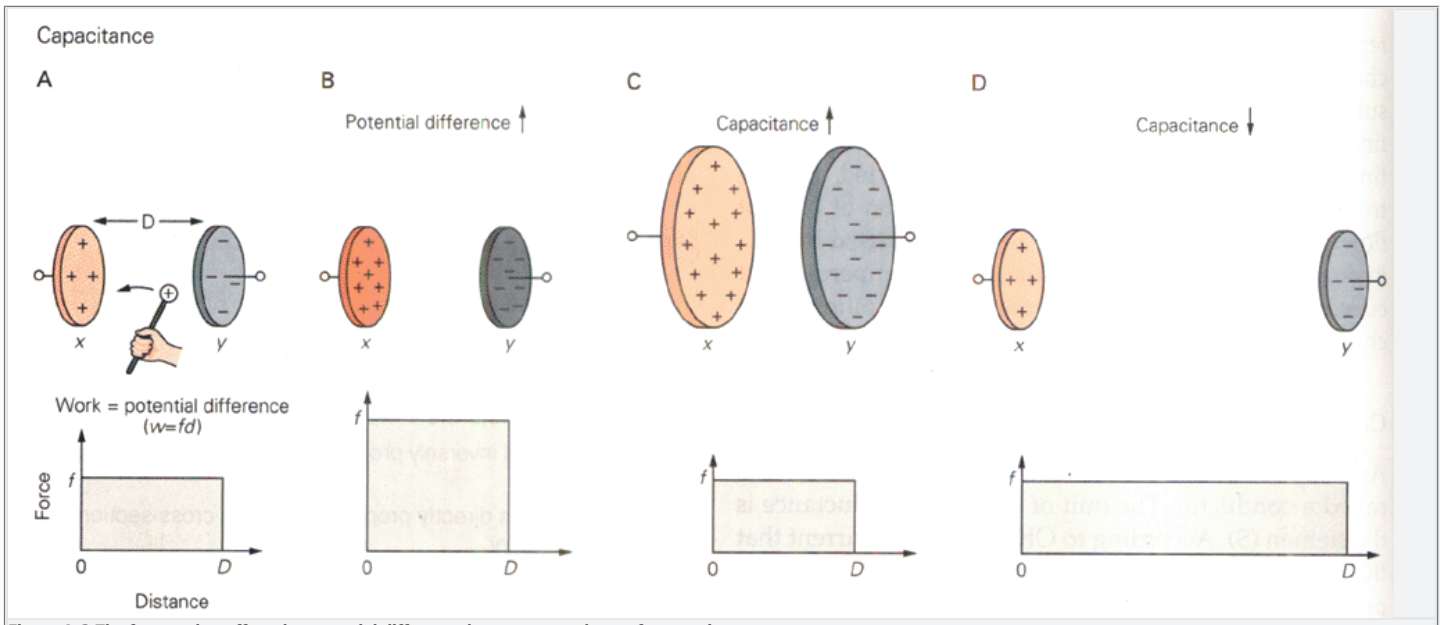


Figure A-2 The factors that affect the potential difference between two plates of a capacitor.

A. As a test charge is moved between two charged plates it must overcome a force. The work done against this force is the potential difference between the two plates.

B. Increasing the charge density increases the potential difference.

C. Increasing the area of the plates increases the number of charges required to produce a given potential difference.

D. Increasing the distance between the two plates increases the potential difference between them.

Capacitance is measured in farads (F). The greater the density of charges on the capacitor plates, the greater the force acting on the test charge, and the greater the resulting potential difference across the capacitor (Figure A-2B). Thus, for a given capacitor, there is a linear relationship between the amount of charge (Q) stored on its plates and the potential difference across it:

$$Q\ (\text{coulombs}) = C\ (\text{farads}) \times V\ (\text{volts}) \qquad \textbf{(A-1)}$$

where the capacitance, C, is a constant.

The capacitance of a parallel-plate capacitor is determined by two features of its geometry: the area (A) of the two plates, and the distance (D) between them. Increasing the area of the plates increases capacitance, because a greater amount of charge must be deposited on each side to produce the same charge density, which is what determines the force f acting on the test charge (Figure A-2A and C). Increasing the distance D between the plates does not change the force acting on the test charge, but it does increase the work that must be done to move it from one side of the capacitor to the other (Figure A-2A and D). Therefore, for a given charge separation between the two plates, the potential difference between them is proportional to the distance. Put another way, the greater the distance the smaller the amount of charge that must be deposited on the plates to produce a given potential difference, and therefore the smaller the capacitance (Equation A-1). These geometrical determinants of capacitance can be summarized by the equation:

$$C \propto \frac{A}{D}.$$

As shown in Equation A-1, the separation of positive and negative charges on the two plates of a capacitor results in a potential difference between them. The converse of this statement is also true: The potential difference across a capacitor is determined by the excess of positive and negative charges on its plates. In order for the potential across a capacitor to change, the amount of electrical charges stored on the two conducting plates must change first.

Rules for Circuit Analysis

A few basic relationships that are used for circuit analysis are listed below. Familiarity with these rules will help in understanding the electric circuit examples that follow.

Conductance

The symbol for a conductor is:

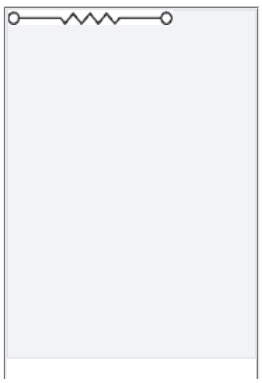


Figure. No Caption Available.

P.1283

A variable conductor is represented this way:

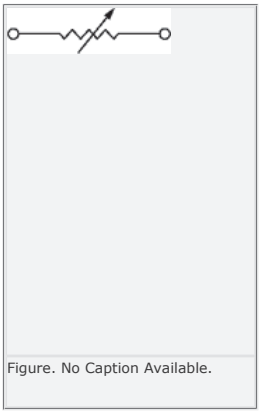


Figure. No Caption Available.

A pathway with infinite conductance (zero resistance) is called a short circuit, and is represented by a line:

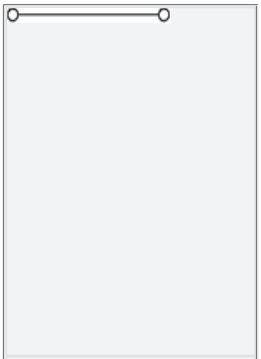


Figure. No Caption Available.

Conductances in parallel add:

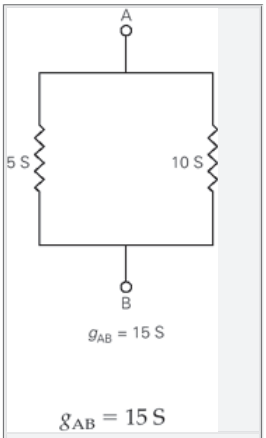


Figure. No Caption Available.

Conductances in series add reciprocally:

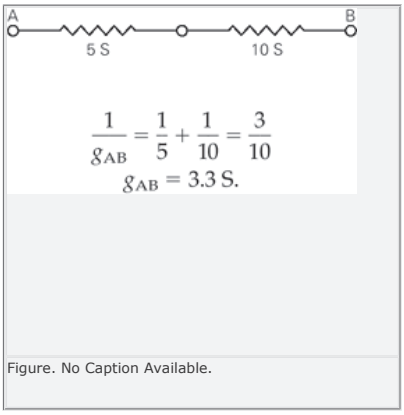


Figure. No Caption Available.

Resistances in series add, while resistances in parallel add reciprocally.

Current

An *arrow* denotes the direction of current flow (net movement of positive charge). Ohm's law is

$$I = V_g = \frac{V}{R}.$$

When current flows through a conductor, the end that the current enters is positive with respect to the end that it leaves:

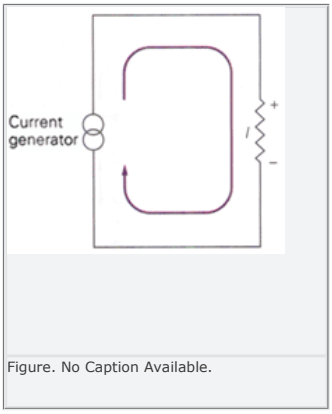


Figure. No Caption Available.

The algebraic sum of all currents entering or leaving a junction is zero (we arbitrarily define current approaching a junction as positive, and current leaving a junction as negative). In the following

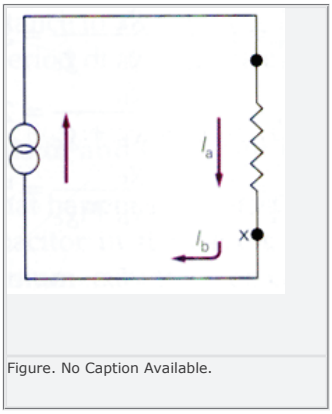


Figure. No Caption Available.

the currents for junction x are:

$$I_A = +5\,\text{A}$$
$$I_B = -5\,\text{A}$$
$$I_A + I_B = 0.$$

In the following circuit

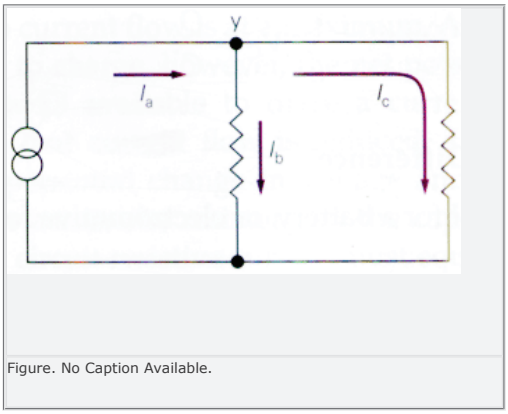


Figure. No Caption Available.

the currents for junction y are:

$$I_a = +3\,\text{A}$$
$$I_b = -2\,\text{A}$$
$$I_c = -1\,\text{A}$$
$$I_a + I_b + I_c = 0.$$

P.1284

Current follows the path of greatest conductance (least resistance). For conductance pathways in parallel, the current through each path is proportional to its conductance value divided by the total conductance of the parallel combination:

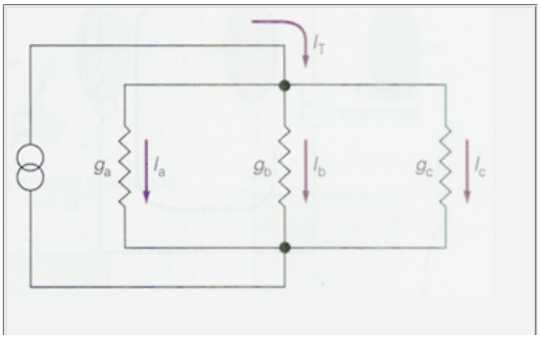


Figure. No Caption Available.

$$I_T = 10\,\text{A}$$
$$g_a = 3\,\text{S}$$
$$g_b = 2\,\text{S}$$
$$g_c = 5\,\text{S}$$
$$I_a = I_T \frac{g_a}{g_a + g_b + g_c} = 3\,\text{A}$$
$$I_b = I_T \frac{g_b}{g_a + g_b + g_c} = 2\,\text{A}$$
$$I_c = I_T \frac{g_c}{g_a + g_b + g_c} = 5\,\text{A}.$$

Capacitance

The symbol for a capacitor is:

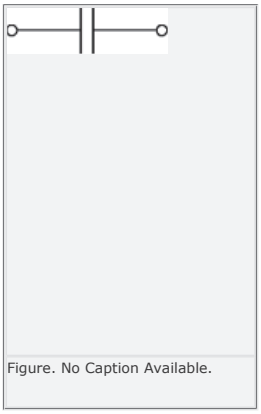


Figure. No Caption Available.

The potential difference across a capacitor is proportional to the charge stored on its plates:

$$V_C = \frac{Q}{C}.$$

Potential Difference

The symbol for a battery, or electromotive force is:

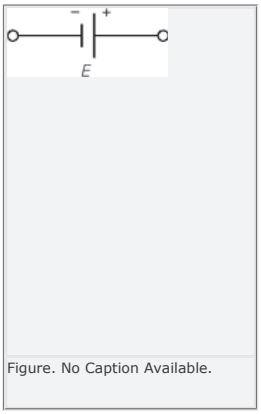


Figure. No Caption Available.

It is often abbreviated by the symbol E. The positive pole is always represented by the longer bar.

Batteries in series add algebraically, but attention must be paid to their polarities. If their polarities are the same, their absolute values add:

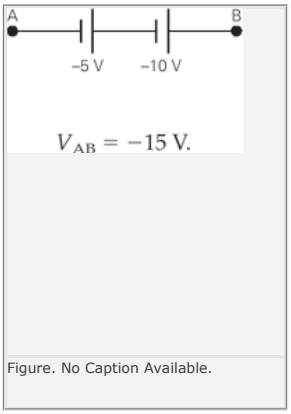


Figure. No Caption Available.

If their polarities are opposite, they subtract:

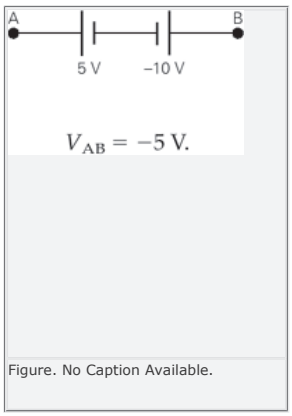


Figure. No Caption Available.

The convention used here for potential difference is that $V_{AB} = (V_A - V_B)$.

A battery drives a current around the circuit from its positive to its negative terminal:

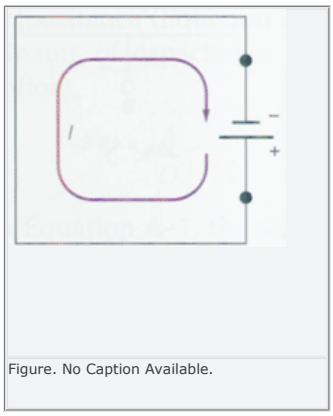


Figure. No Caption Available.

For purposes of calculating the total resistance of a circuit the internal resistance of a battery is set at zero.

The potential differences across parallel branches of a circuit are equal:

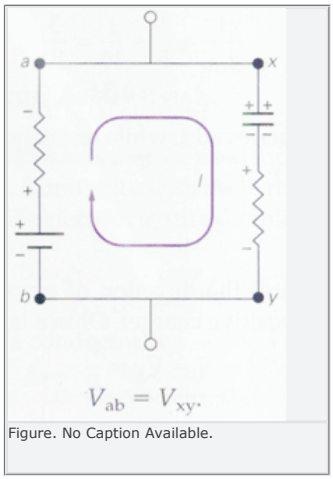


Figure. No Caption Available.

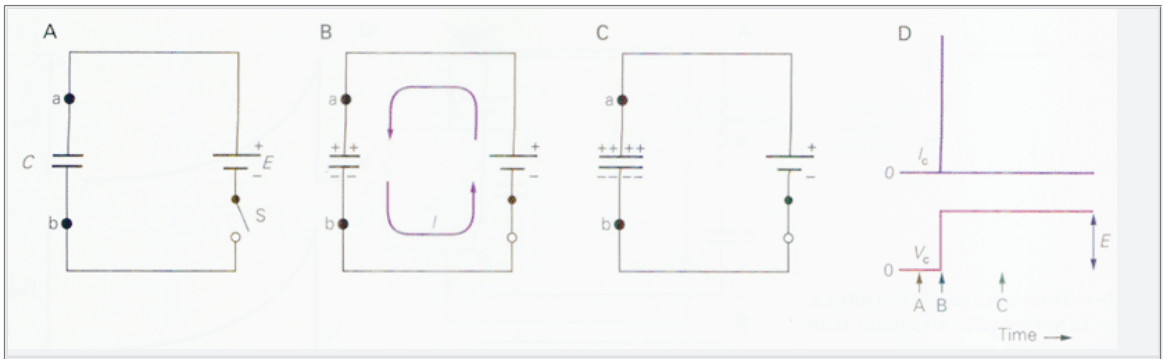


Figure A-3 Time course of charging a capacitor.

A. Circuit before the switch (S) is closed.

B. Immediately after the switch is closed.

C. After the capacitor has become fully charged.

D. Time course of changes in Ic and Vc in response to closing of the switch.

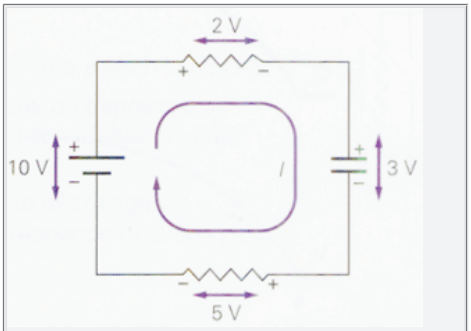


Figure. No Caption Available.

As one goes around a closed loop in a circuit, the algebraic sum of all the potential differences is zero:

$$2\,\mathrm{V} + 3\,\mathrm{V} + 5\,\mathrm{V} - 10\,\mathrm{V} = 0.$$

P.1285

Current Flow in Circuits With Capacitance

Circuits that have capacitive elements are much more complex than those that have only batteries and conductors. This complexity arises because current flow varies with time in capacitive circuits. The time dependence of the changes in current and voltage in capacitive circuits is illustrated qualitatively in the following three examples.

Circuit With Capacitor

Current does not actually flow across the insulating gap in a capacitor; rather it results in a build-up of positive and negative charges on the capacitor plates. However, we can measure a current flowing into and out of the terminal of a capacitor. Consider the circuit shown in Figure A-3A. When switch S is closed (Figure A-3B), a net positive charge is moved by the battery E onto plate a, and an equal amount of net positive charge is withdrawn from plate b. The result is current flowing counterclockwise in the circuit. Since the charges that carry this current flow into or out of the terminals of a capacitor, building up an excess of plus and minus charges on its plates, it is called a capacitive current (Ic). Because there is no resistance in this circuit, the battery E can generate a very large amplitude of current, which will charge the capacitance to a value Q × E × C in an infinitesimally short period of time (Figure A-3D).

Circuit With Resistor and Capacitor in Series

Now consider what happens if a resistor is added in series with the capacitor in the circuit shown in Figure A-4A. The maximum current that can be generated when switch S is closed (Figure A-4B) is now limited by Ohm's law (I = V/R). Therefore, the capacitor charges more slowly. When the potential across the capacitor has finally reached the value V_c = Q/C = E (Figure A-4C), there is no longer a difference in potential around the loop, ie, the battery voltage (E) is equal and opposite to the voltage across the capacitor, Vc. The two thus cancel out, and there is no source of potential difference left to drive a current around the loop. Immediately after the switch is closed the potential difference is greatest, so current flow is at a maximum. As the capacitor begins to charge, however, the net potential difference (V_c = E) available to drive a current becomes smaller, so that current flow is reduced. The result is that an exponential change in voltage and in current flow occurs across the resistor and the capacitor. Note that in this circuit resistive current must equal capacitative current at all times (see Rules for Circuit Analysis, above).

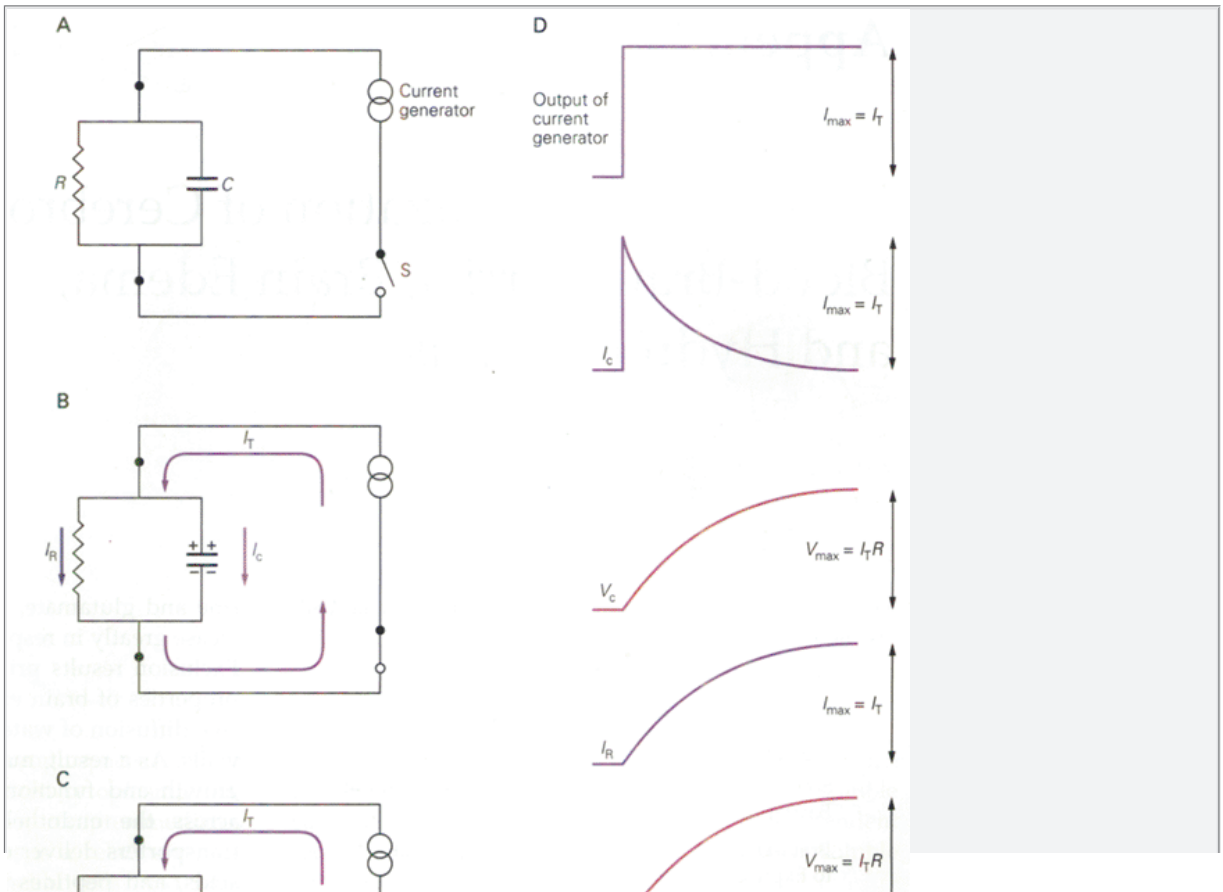
Circuit With Resistor and Capacitor in Parallel

Consider now what happens if we place a parallel resistor and capacitor combination in series with a constant

P.1286

P.1287

current generator that generates a current IT (Figure A-5A). When switch S is closed current starts to flow around the loop. Initially, in the first instant of time after the current flow begins, all of the current flows into the capacitor, ie, I_T=I_c. However, as charge builds up on the plates of the capacitor, a potential difference Vc is generated across it. Since the resistor and capacitor are in parallel, the potential across them must be equal; thus, part of the total current begins to flow through the resistors, such that $I_R R = V_R = V_c$ (Figure A-5B). As less and less current flows into the capacitor, its rate of charging will become slower; this accounts for the exponential shape of the curve of voltage versus time. Eventually, a plateau is reached at which the voltage no longer changes. When this occurs, all of the current flows through the resistor, and V_c=V_R=I_TR (Figure A-5C).



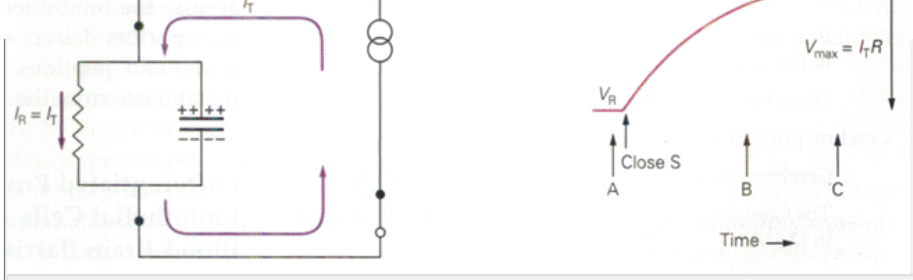


Figure A-4 Time course of charging a capacitor in series with a resistor from a constant voltage source (E).

A. Circuit before the switch (S) is closed.

B. Shortly after the switch is closed.

C. After the capacitor has settled at its new potential.

D. Time course of current flow, of the increase in charge deposited on the capacitor, and of the increased potential differences across the resistor and the capacitor.

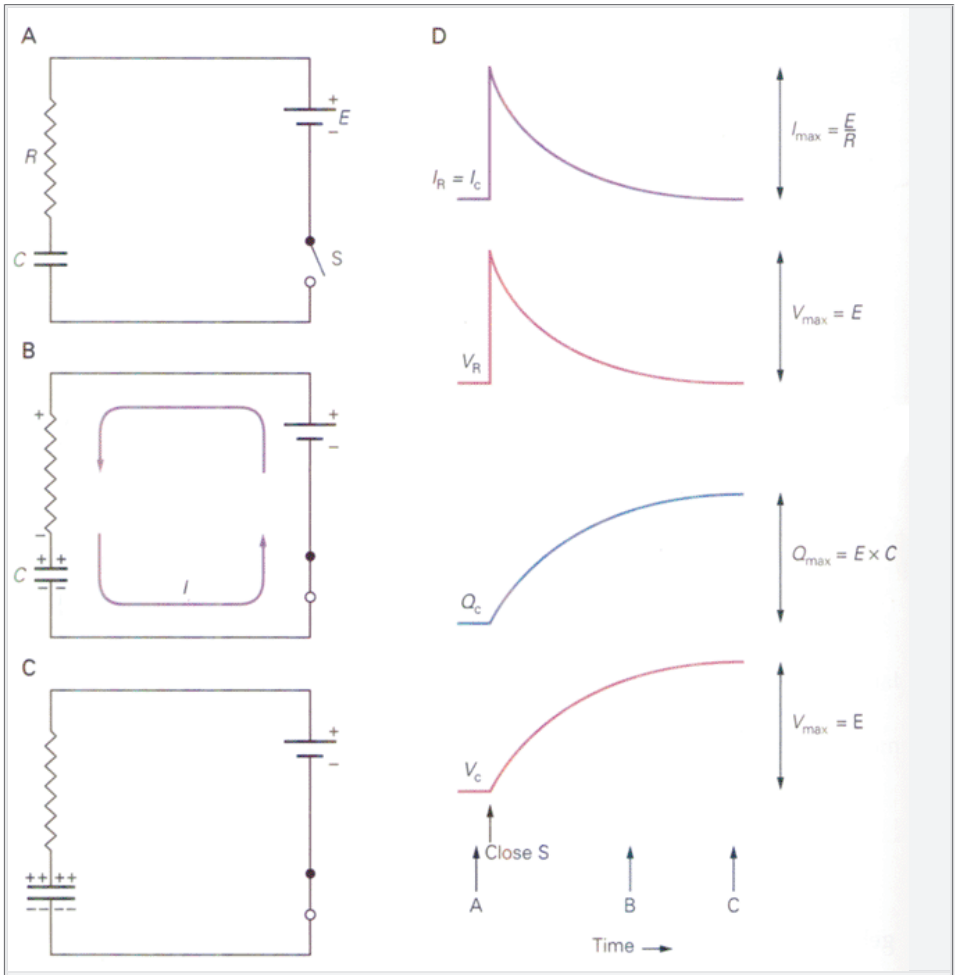


Figure A-5 Time course of charging a capacitor in parallel with a resistor from a constant current source.

A. Circuit before the switch (S) is closed.

B. Shortly after the switch is closed.

C. After the charge deposited on the capacitor has reached its final value.

D. Time course of changes in I_c, V_c, I_R, and V_R in response to closing of the switch.

Footnote

[1]Note the analogy of this formula for current flow to the other formulas for describing flow; eg, bulk flow of a liquid due to a hydrostatic pressure; flow of a solute in response to a concentration gradient; flow of heat in response to a temperature gradient, etc. In each case flow is proportional to the product of a driving force times a conductance factor.

Appendix B

Ventricular Organization of Cerebrospinal Fluid: Blood-Brain Barrier, Brain Edema, and Hydrocephalus

John Laterra

Gary W. Goldstein

A BARRIER TO THE ENTRY of blood-borne substances to the brain has been recognized for at least a century. Evidence for the blood-brain barrier was first obtained in the nineteenth century, when it was observed that acidic vital dyes stain the brain if the dye is injected into the cerebrospinal fluid (CSF) but not if injected into the blood stream. This barrier maintains a stable environment for neurons to function effectively. It excludes many toxic substances and protects neurons from circulating neurotransmitters such as norepinephrine and glutamate, the blood levels of which can increase greatly in response to stress or even after a meal. Exclusion results primarily from specialized anatomic properties of brain endothelial cells that limit the passive diffusion of water-soluble substances across vessel walls. As a result, many metabolites required for brain growth and function must be transported selectively across the endothelial surface. Specific endothelial transporters deliver energy substrates, essential amino acids, and peptides from blood to brain and remove metabolites from the brain.

Differentiated Properties of Brain Capillary Endothelial Cells Account for the Blood-Brain Barrier

Anatomy of the Blood-Brain Barrier

Brain microvessels are composed of endothelial cells, pericytes with smooth muscle-like properties that reside adjacent to capillaries, and astroglial processes that ensheath more than 95% of the abluminal microvessel surface (Figure B-1). It was originally thought that the glial foot processes formed the blood-brain barrier, but electron-microscopic studies identified the endothelial cell as the principal anatomic site of the blood-brain barrier (Figure B-2).

The blood-brain barrier results from specialized properties of the endothelial cells, their intercellular junctions, and a relative lack of vesicular transport. In capillaries of peripheral organs and in the relatively few brain capillaries that do not form a barrier (eg, the circumventricular organs) blood-borne polar molecules

P.1289

diffuse passively across vessels through spaces between endothelial cells, through specialized cytoplasmic fenestrations, or by fluid-phase or receptor-mediated endocytosis. *Fluid-phase endocytosis* is a relatively nonspecific process by which endothelial cells (and most other cells) engulf and then internalize molecules encountered in the extracellular space by vesicular endocytosis. *Receptor-mediated endocytosis* is a specific process in which a ligand first binds to a membrane receptor on the external surface of the cell, is internalized by means of clathrin-coated vesicles, and is transported across the cell membrane. Once within the cell, the vesicle may fuse with an endosome and release its ligand.

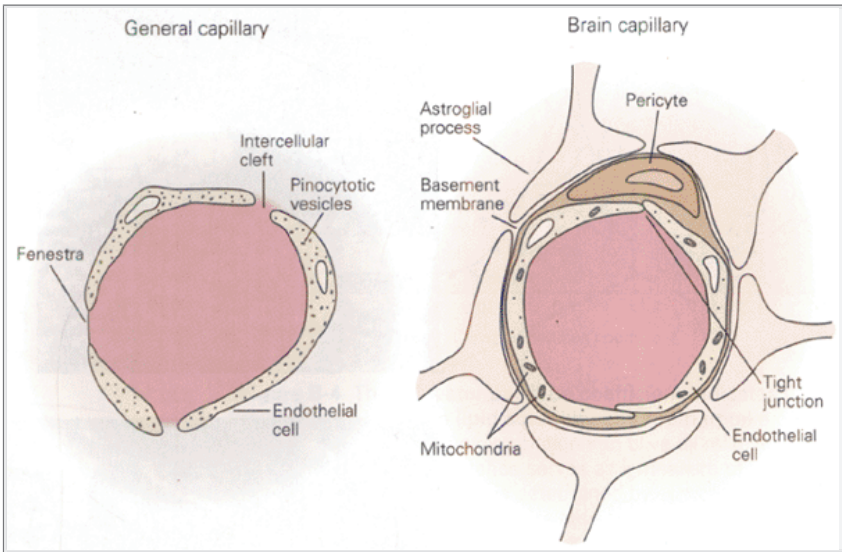


Figure B-1 The ultrastructural features of the capillary endothelial cells of the brain differ from those of general (systemic) capillaries. The endothelial cells of barrier capillaries are relatively lacking in pinocytotic vesicles, contain an increased number of mitochondria believed to support energy-dependent transport properties, and are interconnected by very complex interendothelial tight junctions. These anatomic features in conjunction with specific transport systems (see Figure B-5) result in highly selective transport of water-soluble compounds across the barrier endothelium. Astrocyte foot processes almost completely surround the blood-brain barrier capillaries and are thereby believed to influence barrier-specific endothelial differentiation. In contrast, systemic capillaries have interendothelial clefts, fenestrae, and prominent pinocytotic vesicles. These features of systemic capillaries allow relatively nonselective diffusion across the capillary wall. (From Goldstein and Betz 1986.)

Endothelial cells of blood-brain barrier vessels are relatively deficient in vesicular transport. They also are not fenestrated. Instead they are interconnected by complex arrays of tight junctions (Figure B-2C). These junctions between the endothelial cells block diffusion across the vessel wall. All endothelial cells are interconnected by tight junctional complexes but normally have low resistance (5–10 ω/cm2). In the vessels of the bloodbrain barrier, however, the resistance is very high (2000 ω/cm2), and molecules as small as K^+ ions are excluded (Figure B-3).

Selectivity of the Blood-Brain Barrier

Normal development and brain function require a large number of compounds that must be able to cross brain microvessels. Entry into the CSF is achieved primarily in three ways: (1) by diffusion of lipid-soluble substances, (2) by facilitative and energy-dependent receptor-mediated transport of specific water-soluble substances, and (3) by ion channels.

Lipid-Soluble Substances

The brain is separated from blood only by a very large surface of endothelial cell membranes (approximately 180 cm^2/g in gray matter). This permits the efficient exchange of lipid-soluble gases such as O_2 and CO_2, an exchange limited only by the surface area of the blood vessel and by cerebral blood flow. Barrier vessels are impermeable to poorly lipid-soluble molecules such a mannitol, as compared to very lipid-soluble compounds such as butanol. The permeability coefficient of the blood-brain barrier for many substances is directly proportional to the lipid solubility of the substance as measured by the oil-water partition coefficient (Figure B-4).

An example of this effect is in the correlation between the relative abuse potential of psychoactive drugs such as nicotine and heroin and the high oil-water partition coefficients of these drugs. Increasing the lipid solubility of pharmacologic agents within strict limits will enhance their delivery to the brain. Drugs with very high oil-water partition coefficients are poorly soluble in blood and bind to serum protein albumin, properties that reduce delivery to the brain. However, the permeability

P.1290

of some substances, such as glucose and vinca alkaloids, is not predicted accurately by the lipid solubility of the substance. This observation is explained by the presence of selective endothelial transport or enzyme systems that increase or inhibit substrate permeability.

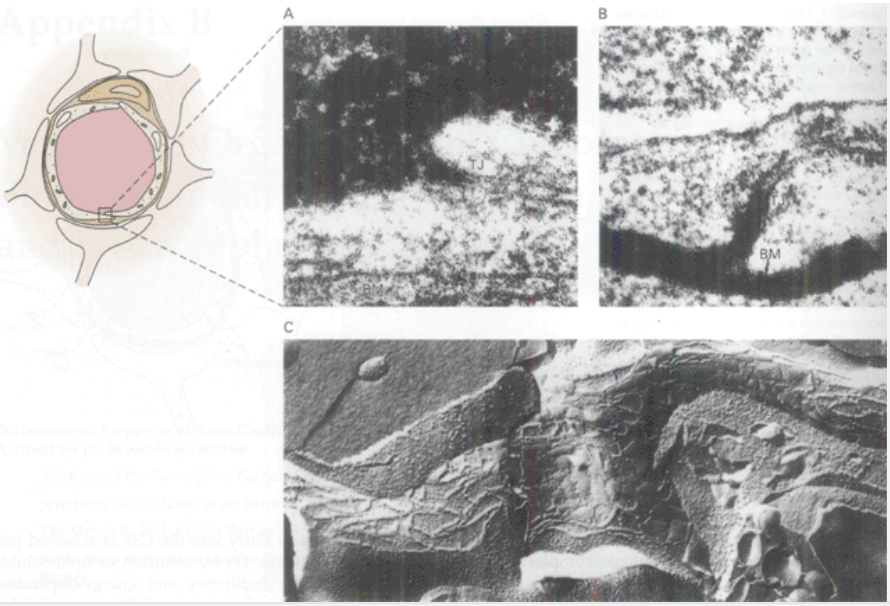


Figure B-2 Interendothelial tight junctions are the basis of the anatomic blood-brain barrier.

A. When injected into arteries that supply the brain, the electron-dense tracer horseradish peroxidase is easily visualized within the brain vessel lumens (**dark staining**) at top but is excluded from entering the brain by interendothelial tight junctions (**TJ**). (From Reese and Karnovsky 1967.)

B. When injected into the subarachnoid space, horseradish peroxidase readily diffuses between the perivascular astroglial foot processes and across the abluminal basement membrane (**BM**) but fails to penetrate interendothelial tight junctions (**TJ**). The luminal space appears at the top of the electron micrographs in both **A** and **B**. (From Brightman and Reese 1969.)

C. Freeze-fracture photomicrograph of isolated brain microvessels, showing complex arrays of interendothelial tight junctions. (From Shivers et al. 1984.)

Blood-Brain Barrier Transport Properties

Most substances that must cross the blood-brain barrier are not lipid soluble and therefore cross by specific carrier-mediated transport systems (Figure B-5). Because the brain uses glucose exclusively as its source of energy, the hexose transporter (glucose transporter isotype-1, Glut1) of the blood-brain barrier endothelial cells has been particularly well characterized.

Glut1 consists of 492 amino acids and has 12 putative transmembrane domains, similar to other transporters. It is a facilitative, saturable, and stereospecific transporter that functions at both the luminal and the abluminal endothelial cell membranes. Because it is not energy dependent, it cannot move glucose against a concentration gradient. The net flux of glucose is driven by the relatively higher concentration of glucose in plasma. Transport is half-saturated at glucose concentrations of 5 to 10 mM. More than 99% of the glucose molecules that enter blood-brain barrier endothelial cells are shuttled across for use by neurons and glia. The gene that codes for Glut1 resides on human chromosome 1. Whereas the complete absence of Glut1 is likely to be incompatible with life, deficient Glut1 expression

is thought to be responsible for a rare clinical syndrome consisting of mental retardation, epilepsy, and persistent low CSF glucose concentration (ie, hypoglycorrhachia).

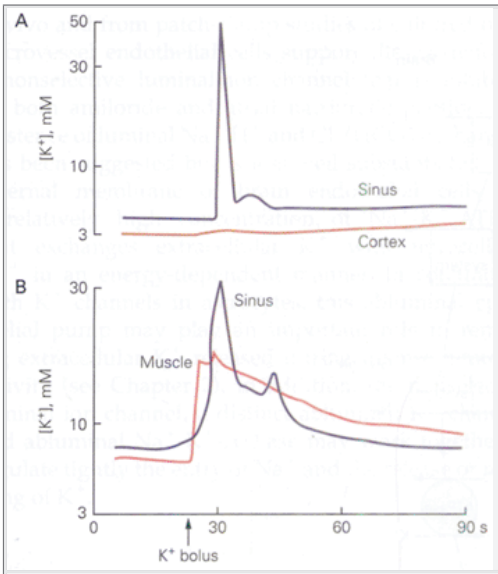


Figure B-3 Time course of K^+ concentration in the extracellular fluid of cerebral cortex (A), neck muscle (B), and sagittal sinus (A and B) following the bolus injection of KCl in the aortic arch. Note that the extracellular fluid K^+ concentration of cerebral cortex remains essentially unchanged because of the blood-brain barrier to ions. In contrast, K^+ diffuses rapidly across nonbarrier vessels into the extracellular space of muscle. (From Hansen et al. 1977.)

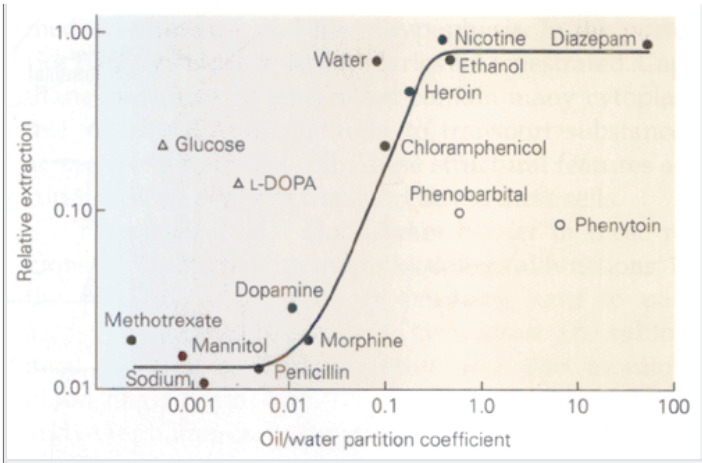


Figure B-4 The oil-water partition coefficient indicates the relationship between lipid solubility and brain uptake of selected compounds. The distribution into olive oil relative to water for each test substance serves as a measure of its lipid solubility. The brain uptake is determined by comparing the extraction of each test substance relative to a highly permeable tracer during a single passage through the cerebral circulation. In general, compounds with higher oil-water partition coefficients show increased entry into brain. Uptake of the anticonvulsants phenobarbital and phenytoin is lower than predicted from their lipid solubility partly because of their binding to plasma proteins. This explains the slower onset of anticonvulsant activity of these agents compared to diazepam. Uptake of glucose and L-DOPA is greater than predicted by their lipid solubility because specific carriers facilitate their transport into the brain capillary. (From Goldstein and Betz 1986.)

Amino acids are transported across barrier endothelial cells primarily by three distinct carrier systems: the L system, the A system, and the ASC system. These carriers are characterized by their different patterns and mechanisms of transport, and by their preference for different amino acid analogs. Large neutral amino acids with branched or ringed side chains, such as leucine and valine, are transported primarily by the L system. The L system is a Na^+-independent, facilitative transporter and is located at luminal and abluminal endothelial membranes. It can be inhibited experimentally by 2-aminobicycloheptane-2-carboxylic acid (BCH). Like glucose, large neutral amino acids are transported from blood down a concentration gradient. This carrier system transports systemically administered L-DOPA, the dopamine precursor that is the mainstay for treating Parkinson disease. Competitive inhibition of the transport of large neutral amino acids by elevated levels of phenylalanine may explain some of the neurotoxicity associated with untreated phenylketonuria, because phenylalanine is transported by the L system.

Glycine and neutral amino acids with short linear or polar side chains, such as alanine or serine, are preferentially transported by the A system. Unlike the L system, this carrier is energy dependent, Na^+ dependent, and experimentally inhibited by α-methylaminoisobutyric acid (MeAIB). The ASC system is also an energy-dependent and Na^+-dependent transporter that preferentially recognizes alanine, serine, and cysteine. ASC-mediated transport is insensitive to both BCH and MeAIB. In contrast to the L carrier, the Aand ASC transport systems are located exclusively at the abluminal endothelial cell surface. The physiological consequence of this localization is that the small neutral amino acids are transported primarily out of the brain up a concentration gradient through these carriers. The Asystem may limit the accumulation of the inhibitory neurotransmitter glycine within the spinal cord and the excitatory neurotransmitter glutamate within the brain. The energy and Na^+ exchange required for these systems is supplied by Na^+-

P.1292

K^+-ATPase, which also is localized to the abluminal endothelial membrane.

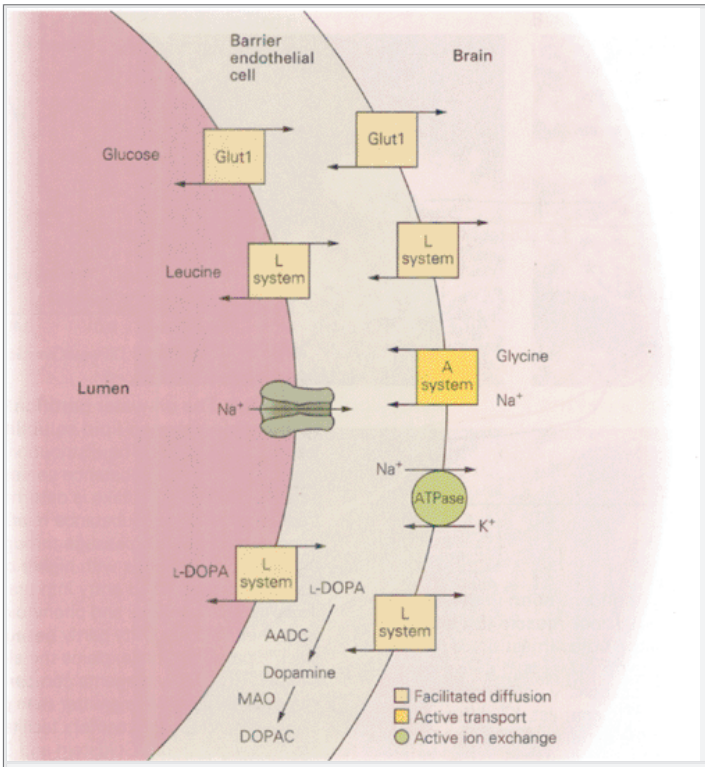


Figure B-5 A complex system of polarized transporter proteins and ionic channels determine the specific movement of water-soluble compounds and ions across barrier endothelial cells. Some transporters (eg, Glut1 and L system) facilitate the movement of substrates down concentration gradients, and others (eg, A system and Na^+-K^+-ATPase) actively transport substrates via energy-dependent mechanisms. Enzyme systems such as amino acid decarboxylase (AADC) and monoamine oxidase (MAO) function as a metabolic barrier by converting within the barrier endothelial cells substances such as L-DOPA to 3,4-dihydroxyphenylacetic acid.

Another transport system found to be most abundant in the blood-brain barrier microvessels belongs to a family of transmembranous proteins initially described for their ability to impart multiple drug resistance (MDR) to tumor cells. These transmembranous transporters remove a broad range of natural and synthetic hydrophobic toxins from cells that express the transporters. The MDR transporter (ie, p-glycoprotein) in the blood-brain barrier influences the delivery to the brain of many compounds used for cancer chemotherapy (eg, vinca alkaloids, actinomycin-D) and for other therapeutic purposes (eg, cyclosporin). Because p-glycoproteins pump certain steroid hormones, they may also have a physiological role. MDR transporters are expressed by blood-brain barrier microvessels but not by vessels of most other tissues. Mice genetically engineered to lack MDR1a gene expression no longer express blood-brain barrier p-glycoprotein and are much more sensitive than wild-type controls to centrally acting toxic compounds, indicating that MDR gene expression at the blood-brain barrier protects the brain from circulating neurotoxins.

Ion Channels and Exchangers

Specific ion channels and ion transporters mediate electrolyte movement across the blood-brain barrier. Evidence from transport studies across brain microvessels

P.1293

in vivo and from patch-clamp studies of cultured brain microvessel endothelial cells support the existence of a nonselective luminal ion channel that is inhibited by both amiloride and atrial natriuretic peptide. The existence of luminal Na^+/H^+ and Cl^-/HCO_3^- exchangers has been suggested but is less well substantiated. The external membrane of brain endothelial cells has a relatively high

concentration of Na^+-K^+-ATPase that exchanges extracellular K^+ with intracellular Na^+ in an energy-dependent manner. In conjunction with K^+ channels in astrocytes, this abluminal endothelial pump may play an important role in removing extracellular K^+ released during intense neuronal activity (see Chapter 2). In addition, the nonselective luminal ion channel, a distinct abluminal K^+ channel, and abluminal Na^+-K^+-ATPase may work together to regulate tightly the entry of Na^+ and the release or recycling of K^+.

The Metabolic Blood-Brain Barrier

Transport systems and carriers are not the only components that regulate the composition of the interstitial fluid. A similarly important role is also played by certain enzyme systems specific to the blood-brain barrier. The first recognized and still the best characterized example is the barrier to L-DOPA. Plasma L-DOPA enters brain endothelial cells by means of the L-system amino acid transporter. The relatively high amounts of DOPA decarboxylase and monoamine oxidase in endothelial cells of the barrier rapidly metabolize L-DOPA to 3,4-dihydroxyphenylacetic acid, thereby inhibiting the entry of L-DOPA to the brain (Figure B-5). This explains why effective Parkinson disease therapy requires that L-DOPA be given together with an inhibitor of DOPA decarboxylase. Other bloodborne amines, including catecholamines, are also inactivated by monoamine oxidases of the barrier endothelium. In addition to its proposed role in cystine transport, the abundant endothelial enzyme of the blood-brain barrier, γ-glutamyl transpeptidase, detoxifies glutathione-bound compounds and vasoactive leukotrienes.

Some Areas of the Brain Do Not Have a Blood-Brain Barrier

Not all cerebral blood vessels are entirely impermeable. Leaky areas include the posterior pituitary and circumventricular organs such as the area postrema, subfornical organ, the laminar terminalis, subcommissural organ, median eminence, and neurohypophysis. In the posterior pituitary most of the capillaries are fenestrated. Capillaries that are not fenestrated contain many cytoplasmic vesicles that are thought to transport substances across the endothelial cell. These structural features account for the enhanced transport across these cells.

The absence of a blood-brain barrier in these regions is consistent with their physiological functions. In the pituitary, neurosecretory products have to pass across endothelial cells into the circulation. The subfornical organ is a chemoreceptive area that monitors blood angiotensin II levels to regulate water balance and other homeostatic functions.

These leaky regions are isolated from the rest of the brain by specialized ependymal cells called *tanycytes* located along the ventricular surface close to the midline. The tanycytes are coupled by tight junctions and prevent free exchange between the circumventricular organs and the CSF.

Brain-Derived Signals Induce Endothelial Cells to Express Blood-Brain Barrier Properties

The cellular and molecular signals that induce endothelial expression of the blood-brain barrier phenotype have been elusive. The neural anlagen of the brain is vascularized very early in development through the invasion of proliferating vessels derived from an extraneural vascular plexus. These perineural vessels are composed of fenestrated endothelial cells that have no blood-brain barrier. Soon after penetrating neural tissue (within 2–4 days in the rat), fenestrations are lost. The subsequent expression of different blood-brain barrier properties on the endothelium occurs gradually, suggesting a maturational cascade.

The blood-brain barrier properties expressed by brain endothelial cells are believed to be induced by cells within the surrounding brain. Thus, barrier properties will develop in peripheral endothelial cells that invade brain tissue transplants but not in brain endothelial cells that invade peripheral tissue transplants. The cellular origin and biochemical nature of the brain parenchymal signals that induce endothelial barrier formation are not known but may originate from perivascular astroglia. Consistent with an important role for such parenchymal signals is the observation that brain endothelial cells rapidly lose their blood-brain barrier properties when isolated in culture and reexpress barrier properties when reimplanted in the central nervous system. Of the principal cell types that comprise the brain—neurons, oligodendroglia, and astroglia— astroglia are thought to be particularly important for

this function because of their close association with the abluminal endothelial surface in mature brain. Astroglial cells, particularly in conjunction with cyclic adenosine monophosphate agonists, specifically increase the complexity of interendothelial tight junctions.

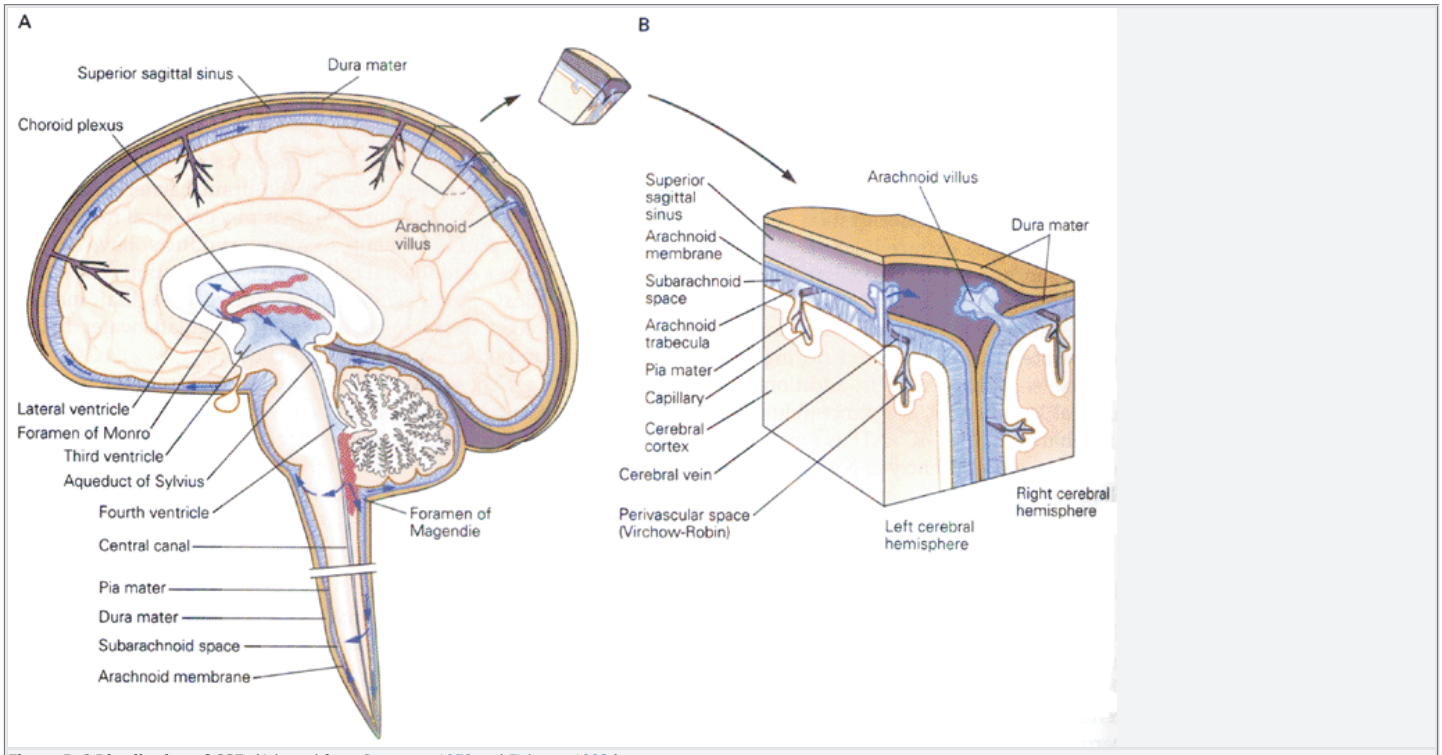


Figure B-6 Distribution of CSF. (Adapted from Carpenter 1978 and Fishman 1992.)

A. Sites of formation, circulation, and absorption of CSF. All spaces containing CSF communicate with each other. Choroidal and extrachoroidal sources of the fluid exist within the ventricular system. CSF circulates to the subarachnoid space and is absorbed into the venous system via the arachnoid villi. The presence of arachnoid villi adjacent to the spinal roots supplements the absorption into the intracranial venous sinuses. (Adapted from Fishman 1992.)

B. The subarachnoid space is bounded externally by the arachnoid membrane and internally by the pia mater, which extends along blood vessels that penetrate the surface of the brain. (Adapted from Carpenter 1978.)

Disorders of the Blood-Brain Barrier

A variety of pathological situations are associated with altered blood-brain barrier function. Many brain tumors contain vessels with a poorly developed blood-brain barrier. The least aggressive low-grade astrocytomas may contain vessels with close-to-normal barrier characteristics. In contrast, malignant primary tumors (ie, anaplastic astrocytoma, glioblastoma) and brain metastases of systemic cancers contain vessels that are excessively leaky and lack the differentiated transport properties of normal blood-brain barrier vessels. The abnormal vessel permeability accounts for the excessive accumulation of interstitial fluid (called *vasogenic edema*) commonly associated with brain tumors. The abnormal properties of tumor endothelial cells are presumably explained either by the absence of normal interactions between astrocytes and capillaries or by the secretion of factors by tumor cells. Vascular endothelial growth factor/vascular permeability factor is one such factor that may account for the excessive proliferation and permeability of the glioblastoma vessel.

Another condition in which the blood-brain barrier is altered is bacterial meningitis. The blood-brain barrier is normally impermeable to antibiotics such as penicillin. Bacterial meningitis, abscesses, and their associated inflammatory responses cause a partial breakdown of the barrier. This barrier response appears to be

mediated in part by the accumulation of vasoactive eicosanoids and inflammatory cytokines such as tumor necrosis factor. Although this barrier dysfunction accounts for some of the adverse

neurological effects of these infections, it also results in an enhanced ability to deliver antibiotics to the site of infection within the brain.

Because the development and function of the normal brain is linked closely to specific anatomical, biochemical, and transport properties of the blood-brain barrier, specific defects in genes that code for barrier endothelial proteins might account for inherited brain disorders. The first developmental disorder of blood-brain barrier transport has been described. Patients with this syndrome are normal at birth but soon develop poorly controlled seizures, diminished brain growth, and mental retardation in association with a substantially diminished concentration of glucose in the CSF. Glucose enters the CSF through Glut1. This transporter is also found in red blood cells. In these patients red blood cell Glut1 is reduced by approximately 50%, suggesting that this syndrome is due to reduced Glut1 gene expression and a subsequent decrease in blood-to-brain glucose transport. The genetic basis for this disorder has yet to be fully defined.

Cerebrospinal Fluids Has Several Functions

CSF communicates with brain interstitial fluid and is therefore important in maintaining a constant external environment for neurons and glia. The primarily oneway flow of CSF from the ventricular system, around the spinal cord, into the subarachnoid space around the brain, and into the venous sinuses is a major route for removing potentially harmful brain metabolites. CSF also provides a mechanical cushion to protect the brain from impact with the bony calvarium when the head moves. By its buoyant action, CSF allows the brain to float, thereby reducing its effective weight in situ to less than 50 g. CSF may also serve as a lymphatic system for the brain and as a conduit for polypeptide hormones that are secreted by hypothalamic neurons and that act at remote sites in the brain. The pH of CSF affects both pulmonary ventilation and cerebral blood flow—another example of the homeostatic role of CSF.

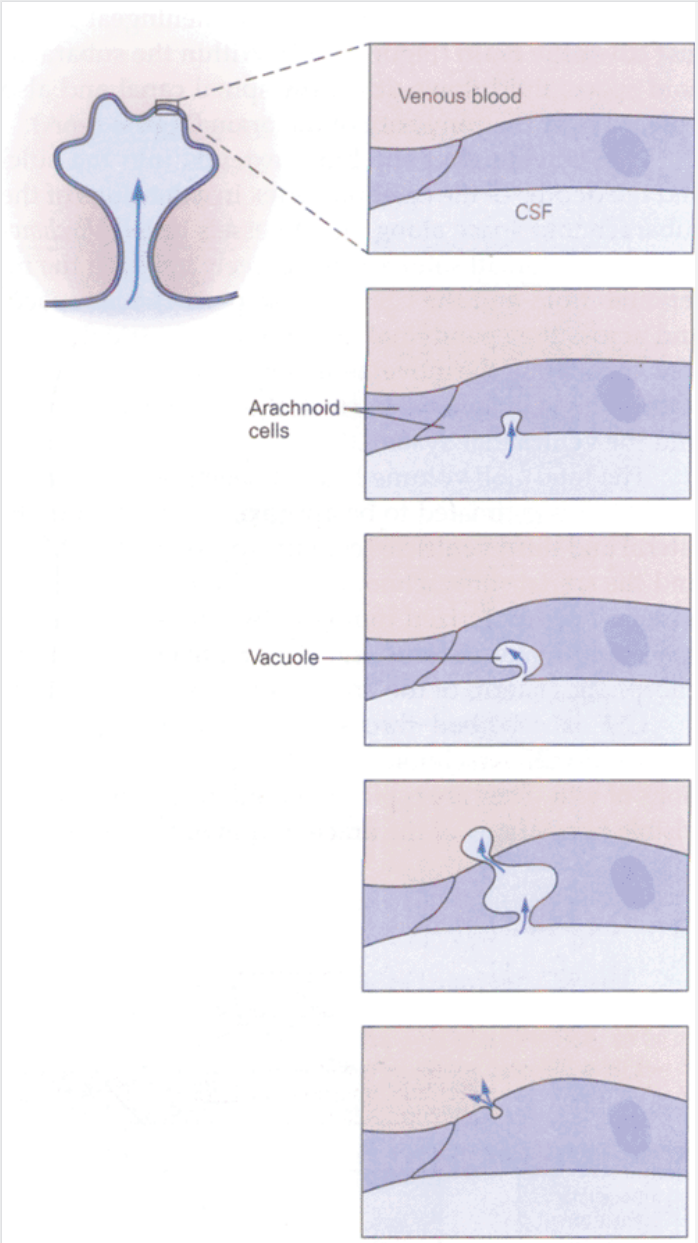


Figure B-7 Transport of CSF within the arachnoid villi is thought to be achieved by giant vacuoles. This mechanism could account for the one-way bulk flow of CSF from the subarachnoid space to the venous system. The arachnoid cells have tight intercellular junctions. Some vesicles are large enough to encompass red blood cells. (Adapted from Fishman 1992.)

Cerebrospinal Fluid Is Secreted by the Choroid Plexus

CSF is secreted in the lateral ventricles mainly by the choroid plexus (Figure B-6A), which consists of capillary networks surrounded by cuboidal or columnar epithelium. CSF flows from the lateral ventricles through the interventricular foramina of Monro into the third ventricle. From there it flows into the fourth ventricle through the cerebral aqueduct of Sylvius and then through the foramina of Magendie and Luschka into the subarachnoid space. The subarachnoid space lies between the arachnoid and the pia mater, which, together

with the dura mater, form the three meningeal layers that cover the brain (Figure B-6B). Within the subarachnoid space, fluid flows down the spinal canal and also upward over the convexity of the brain (Figure B-6A).

CSF flowing over the brain extends into the sulci and the depths of the cerebral cortex in extensions of the subarachnoid space along blood vessels called *Virchow-Robin spaces*. Small solutes diffuse freely between the interstitial fluid and the CSF in these perivascular spaces and across the ependymal lining of the ventricular system, facilitating the movement of metabolites from deep within the hemispheres to cortical subarachnoid spaces and the ventricular system.

The total CSF volume has not been measured accurately but is estimated to be approximately 140 ml. The lateral and third ventricles contain approximately 12 ml, and the spinal subarachnoid space about 30 ml as measured by computerized tomography. The subarachnoid space and major cisterns (eg, cisterna magna and mesencephalic cistern) of the brain contain most of the CSF.

CSF is absorbed through the arachnoid granulations and villi. Arachnoid granulations consist of collections of villi. They are typically found in clusters that are visible herniations of the arachnoid membrane through the dura and into the lumen of the superior sagittal sinus and other venous structures (see Figure B-6A). The villi themselves are visible microscopically and separate CSF from venous blood. Cells of the villus membrane may actively form vacuoles that transport fluid from one side of the cell to the other (Figure B-7).

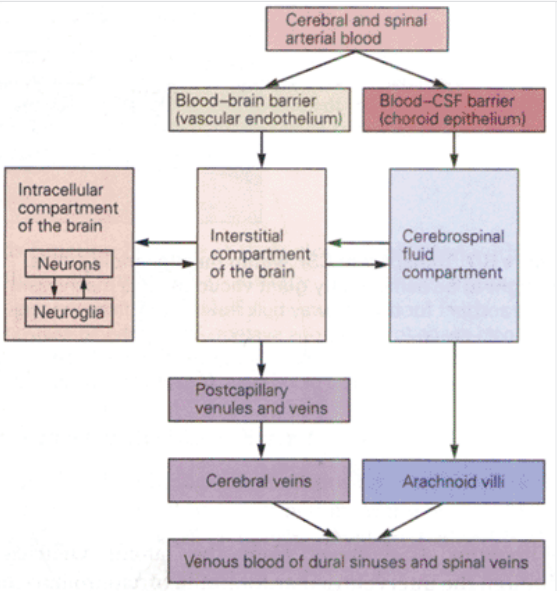


Figure B-8 Relationships between intracranial fluid compartments and the blood-brain and blood-CSF barriers. The tissue elements indicated in **parentheses** form the barriers. **Arrows** indicate the direction of fluid flow under normal conditions. (Adapted from Carpenter 1978.)

Table B-1 Comparision of Serum and Cerebrospinal Fluid

Component	CSF[1]	Serum[1]
Water content (%)	99	93
Protein (mg/dl)	35	7000
Glucose (mg/dl)	60	90
Osmolarity (mOsm/liter)	295	295
Na^+ (meq/liter)	138	138
K^+ (meq/liter)	2.8	4.5
Ca^{2+} (meq/liter)	2.1	4.8
Mg^{2+} (meq/liter)	0.3	1.7
Cl^- (meq/liter)	119	102
pH	7.33	7.41

[1] Average or representative values.(From Fishman 1980.)

The granulations appear to function as valves that allow one-way flow of CSF from the subarachnoid spaces into venous blood. This one-way flow of CSF is sometimes called *bulk flow* because all constituents of CSF leave with the fluid, including small molecules, proteins, microorganisms, and red blood cells. The rate of formation of CSF in adults is 0.35 ml/minute or about 500 ml/day, so that the entire volume of CSF is turned over three to four times per day (Figure B-8).

The choroid plexus is structurally similar to the distal and collecting tubules of the kidney, using capillary filtration and epithelial secretory mechanisms to maintain the chemical stability of the CSF. The capillaries that traverse the choroid plexus are freely permeable to plasma solutes. A barrier exists, however, at the level of the epithelial cells that make up the choroid plexus. This barrier is responsible for carrier-mediated active transport (Figure B-9). The secretory capacities of the choroid plexus are bidirectional, accounting for both continuous production of CSF and active transport of metabolites out of the central nervous system into the blood.

CSF and extracellular fluids of the brain are in a steady state under normal physiological circumstances. The concentrations of K^+, Ca^{2+}, bicarbonate, and glucose in CSF are lower than in blood plasma, and CSF is

also more acidic (Table B-1). These differences are due to regulation of the constituents of CSF by active transport. Under normal conditions, blood plasma and CSF are in osmotic equilibrium, because water follows the osmotic gradient created by active transport of solutes.

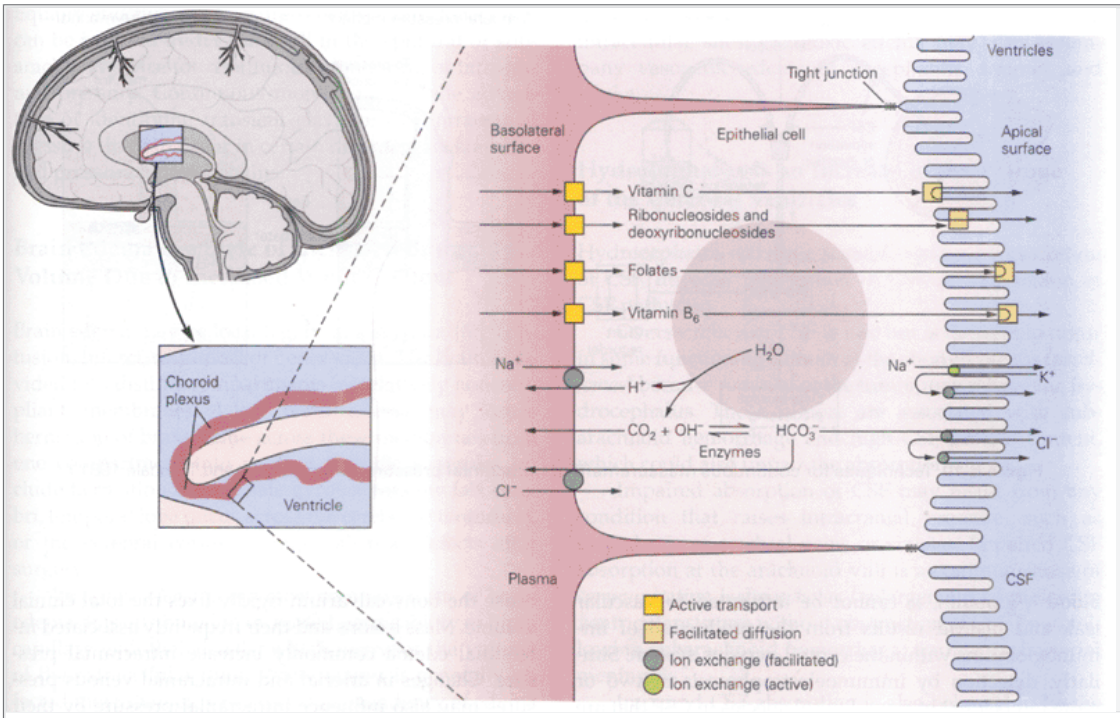


Figure B-9 Flow of molecules across the blood-CSF barrier is regulated by several mechanisms in choroid epithelial cells. Some micronutrients such as vitamin C are transported into epithelial cells by an energy-dependent active transporter located at the basolateral membrane and released into CSF at the apical surface by facilitated diffusion, which requires no energy. Essential ions are also exchanged between CSF and blood plasma. Transport of an ion in one direction is linked to the transport of a different ion in the opposite direction, as in the exchange of Na^+ ions for K^+ ions. (From Spector and Johanson 1989.)

The Composition of Cerebrospinal Fluid May Be Altered in Disease

The gross appearance of the CSF has clinical significance. Normally it is clear and colorless. It may appear cloudy when it contains many leukocytes or has a high protein content. It may also appear grossly bloody or yellow (*xanthochromia*) when blood pigments are left behind after a hemorrhage or when its protein content is greater than 150 mg/dl, indicating that bilirubin (bound to albumin) has been brought from the plasma to the CSF.

Normal CSF does not contain red blood cells. White blood cell counts greater than 4/mm^3 are pathological. In acute bacterial meningitis the count may be a thousandfold greater. Cells may be increased moderately in viral infections or in response to cerebral infarction, brain tumors, or other damage to cerebral tissue. Tumor cells in CSF can be collected and identified by their characteristic morphology and cell type-specific proteins.

Protein content may be increased by many pathological processes of the brain or spinal cord, because of changes in vascular permeability or changes in CSF dynamics due to altered function of choroid plexus or arachnoid granulations. In multiple sclerosis and a few other inflammatory diseases the γ-globulin content of CSF is disproportionately increased to more than 13% of total protein. When this increment in γ-globulin content of CSF is present without a corresponding increase in

P.1298

blood γ-globulin, it cannot be attributed to vascular leak and instead results from the production of immunoglobulins within the central nervous system. Similarly, detection by immunoelectrophoresis of two or more oligoclonal immunological species in CSF that are not present in serum indicates immune-mediated brain disease. Oligoclonal immunoglobulins are present in CSF in about 90% of patients with multiple sclerosis. Protein content greater than 500 mg/dl is usually a manifestation of a block in the spinal subarachnoid space by a solid tumor, meningeal cancer, or other compressing lesion.

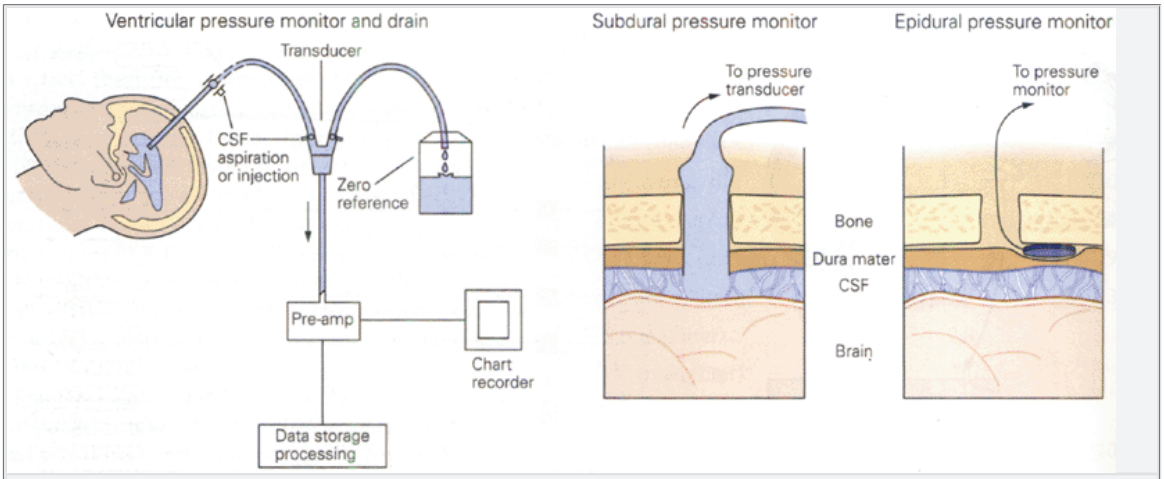


Figure B-10 Techniques for continuous measurement of intracranial pressure. (From Jennett and Teasdale 1981.)

The concentration of glucose in the CSF is decreased in acute bacterial infections but only rarely in viral infections. In chronic diseases a CSF glucose content less than 40 mg/dl is a symptom of a tumor in the meninges; of fungal, yeast, or tuberculosis infection; or of sarcoidosis. The basis for the reduced CSF glucose content is not yet understood. It may be due to impaired transport into CSF; excessive glycolysis by organisms, blood cells, tumor cells, or the brain itself; or combinations of these mechanisms.

Increased Intracranial Pressure May Harm the Brain

In considering the factors that regulate intracranial pressure, the cranium and spinal canal can be regarded as a closed system. According to the Monro-Kellie doctrine, an increase in the volume of any one of the contents of the calvarium—brain tissue, blood, CSF, or other brain fluids—will produce increased intracranial pressure because the bony calvarium rigidly fixes the total cranial volume. Mass lesions and their frequently associated interstitial edema commonly increase intracranial pressure. Changes in arterial and intracranial venous pressures may also influence intracranial pressure by their actions on intracranial blood volume and CSF dynamics. Acute changes in arterial or venous pressures can change intracranial pressure dramatically. Chronic changes may be compensated for by several mechanisms, including venous collateralization and increased absorption or decreased formation of CSF.

CSF pressure is ordinarily measured by *lumbar puncture*, a procedure in which a needle is inserted through the skin, between the fourth and fifth lumbar vertebrae, and into the lumbar subarachnoid space, with the patient lying sideways (lateral decubitus position). Because the spinal cord extends only to the first lumbar vertebra, there is no risk of injuring the cord. When the CSF flows freely through the needle, the hub of the needle is attached to a manometer and the fluid is allowed to rise. The normal pressure is 65–195 mm CSF (or water) or 5–15 mm Hg.

In measuring the lumbar CSF pressure as a guide to intracranial pressure, it is assumed that pressures are equal throughout the neuraxis. Normally this is a reasonable assumption; however, in many disease states (eg, brain tumor or obstruction of CSF pathways) this may not be true. For this reason, and also because the lumbar needle cannot be left in place for prolonged periods, catheters are sometimes inserted into the lateral ventricles to measure the pressure there (Figure B-10).

P.1299

Equally effective are pressure-sensitive transducers that can be inserted under the skull in the epidural or subarachnoid space for continuous monitoring of intracranial pressure. Continuous monitoring has the advantage of identifying transient elevations in intracranial pressure that can occur in certain disorders such as normal pressure hydrocephalus.

Brain Edema Is A State Of Increased Brain Volume Due To Increased Water Content

Brain edema may be local (eg, from a surrounding contusion, infarct, or tumor) or generalized. The brain is divided into distinct compartments by relatively noncompliant membranes. Local brain edema may cause herniation of brain tissue across these membranes from one compartment into another. Specific examples include herniation of cingulate gyrus across the falx cerebri, temporal lobe uncus across the cerebellar tentorium, or the cerebral cortex through calvarial defects after surgery.

Vasogenic edema is the most common form of brain edema. It is attributed to increased permeability of brain capillary endothelial cells, which increases the volume of the extracellular fluid. White matter is generally affected more than gray because of the tendency of edema fluid to accumulate along tracts of white matter. Vasogenic edema is most easily visualized using T2-weighted magnetic resonance imaging. Pathological increases in blood-brain barrier permeability also can be visualized by computerized tomography and magnetic resonance imaging after intravenous administration of contrast agents that selectively enter the brain through the affected vessels. Pathologically increased blood brain barrier permeability, particularly that associated with brain tumors, can be corrected by systemically administered glucocorticoids. The mechanism by which glucocorticoids enhance blood-brain barrier function is poorly understood.

Cytotoxic edema refers to the intracellular swelling of injured neurons, glia, and endothelial cells. It occurs in hypoxia from asphyxia or global cerebral ischemia after cardiac arrest because failure of the ATP-dependent Na^+-K^+ pump allows Na^+, and therefore water, to accumulate within cells. Another cause of cytotoxic edema is water intoxication, a consequence of the acute systemic hypoosmolarity caused by excessive ingestion of water or administration of hypotonic intravenous fluids. Acute hyponatremia, induced for example by inappropriate secretion of antidiuretic hormone or renal saltwasting from secretion of atrial natriuretic hormone, can cause cellular swelling and brain edema. Under these circumstances water moves from extracellular to intracellular sites. Cytotoxic edema may also accompany vasogenic edema in encephalitis, trauma, and stroke.

Hydrocephalus Is An Increase In The Volume Of The Cerebral Ventricles

Hydrocephalus has three possible causes: oversecretion of CSF, impaired absorption of CSF, or obstruction of CSF pathways.

Oversecretion of CSF is rare but is thought to occur in some functioning tumors of the choroid plexus (*papillomas*) because removal of the tumor may relieve the hydrocephalus. These tumors are associated with subarachnoid hemorrhage and high CSF protein content, which could also impair the absorption of CSF.

Impaired absorption of CSF may result from any condition that raises intracranial pressure, such as thrombosis of cerebral veins or sinuses. Impaired CSF absorption at the arachnoid villi is a common cause of *communicating hydrocephalus* (enlargement of the entire ventricular system without obstruction of CSF flow) following subarachnoid hemorrhage, trauma, or bacterial meningitis.

Impaired CSF absorption is also believed to be the cause of **normal-pressure hydrocephalus**, which is characterized by dementia, urinary incontinence, and a disorder of gait called *apraxia*. Brain imaging reveals communicating hydrocephalus, and routine lumbar puncture typically shows normal intracranial pressure. Continuous intracranial pressure monitoring reveals episodic elevations in intracranial pressure, however, suggesting that intermittent intracranial hypertension causes the condition.

Cisternography reveals alterations in the flow of CSF in these patients. In normal patients technetiumlabeled albumin injected into the lumbar subarachnoid space can be traced by a gamma camera up to the cortical convexities where the arachnoid granulations are located; however, the label does not enter the ventricles. In patients with normal-pressure hydrocephalus the isotopic label reaches the cortical convexities after a prolonged delay and may reflux into the ventricles. If identified early, normal-pressure hydrocephalus can be treated surgically by shunting CSF using a ventriculoperitoneal catheter.

Obstruction of CSF pathways may result from tumors, congenital malformations, or scarring. A particularly vulnerable site for all three mechanisms is the narrow aqueduct of Sylvius. Aqueductal stenosis may result from congenital malformations or gliosis due to

intrauterine infection or hemorrhage. Later in life the aqueduct may be occluded by tumor. The outlets of the fourth ventricle may be obstructed by congenital atresia of the foramina of Luschka and Magendie, which may lead to enlargement of all four ventricles (Dandy-Walker syndrome). In early life the cranial vault enlarges with the ventricles; after the sutures fuse, cranial volume is fixed and hydrocephalus develops at the expense of brain volume.

Selected Readings

Bradbury MWB (ed). 1992. *Physiology and Pharmacology of the Blood-Brain Barrier*. New York: Springer.

Del-Bigio MR. 1993. Neuropathological changes caused by hydrocephalus. Acta Neuropathol (Berlin) 85:573–585.

Doczi T. 1993. Volume regulation of the brain tissue—a survey. Acta Neurochir (Wien) 121:1–9.

Doyle DJ, Mark PW. 1992. Analysis of intracranial pressure. J Clin Monit 8:81–90.

Fishman RA. 1975. Brain edema. N Engl J Med 293:706–711.

Fishman RA. 1992. *Cerebrospinal Fluid in Diseases of the Nervous System*. Philadelphia, PA: Saunders.

Goldstein GW, Betz AL. 1986. The blood-brain barrier. Sci Am 255(3):74–83.

Greer M. 1988. Carrier drugs: presidential address, American Academy of Neurology, 1987. Neurology 38:628–632.

Kalaria RN, Gravina SA, Schmidley JW, Perry G, Harik SI. 1988. The glucose transporter of the human brain and blood-brain barrier. Ann Neurol 24:757–764.

Katzman R, Pappius HM. 1973. *Brain Electrolytes and Fluid Metabolism. Baltimore, MD: Williams & Wilkins*.

Keep RF, Xiang J, Betz AL. 1993. Potassium transport at the blood-brain and blood-CSF barriers. Adv Exp Med Biol 331:43–54.

Laterra J, Stewart PA, Goldstein GW. 1991. Development of the blood-brain barrier. In: RA Polin, WW Fox (eds). *Neonatal and Fetal Medicine—Physiology and Pathophysiology*, pp. 1525–1531. Philadelphia, PA: Saunders.

Lyons MK, Meyer FB. 1990. Cerebrospinal fluid physiology and the management of increased intracranial pressure. Mayo Clin Proc 65:684–707.

Mooradian AD, Morin AM, Cipp LJ, Haspel HC. 1991. Glucose transport is reduced in the blood-brain barrier of aged rats. Brain Res 551:145–149.

Nilsson C, Lindvall-Axelsson M, Owman C. 1992. Neuroendocrine regulatory mechanisms in the choroid plexus— cerebrospinal fluid system. Brain Res Rev 17:109–138.

Pardridge W (ed). 1993. The Blood-Brain Barrier: Cellular & *Molecular Biology*. New York: Raven.

Rapoport SI. 1976. *Blood-Brain Barrier in Physiology and Medicine*. New York: Raven.

Ropper AH, Kennedy SF (eds). 1988. *Neurological and Neurosurgical Intensive Care*. Rockville, MD: Aspen.

Segal MB. 1993. Extracellular and cerebrospinal fluids. J Inherit Metab Dis 16:617–638.

References

Betz AL, Goldstein GW. 1986. Specialized properties and solute transport in brain capillaries. Annu Rev Physiol 48:241–250.

Borgesen SE, Gjetrris F. 1982. The predictive value of conductance to outflow of CSF in normal pressure hydrocephalus. Brain 105:65–86.

Borgesen SE, Gjetrris F. 1987. Relationships between intracranial pressure, ventricular size, and resistance to CSF outflow. J Neurosurg 67:535–539.

Brightman MW, Reese TS. 1969. Junctions between intimately apposed cell membranes in the vertebrate brain. J Cell Biol 40:648–677.

Carpenter MB. 1978. *Human Neuroanatomy*, 7th ed. Baltimore, MD: William & Wilkins.

DeVivo DC, Trifiletti R, Jacobson RI, Harik SI. 1990. Glucose transporter deficiency causing persistent hypoglycorrhachia: a unique cause of infantile seizures and acquired microcephaly. Ann Neurol 29:414–415.

Ehrlich P. 1885. *Das Sauerstoff-Bedörfnis des Organismus. Eine Farbenanalytische Studie*. Berlin: Hirschwold, cited by Friedemann 1942.

Friedemann U. 1942. Blood-brain barrier. Physiol Rev 22:125–145.

Goldman E. 1909. Die äussere und innere Sekretion des gesunden und kranken Organismus im Lichte der "vitalen Färbung." Beitr Klin Chirurg 64:192–265.

Guerin C, Laterra J, Hruban R, Brem H, Drewes LR, Goldstein GW. 1990. The glucose transporter and blood-brain barrier of human brain tumors. Ann Neurol 28:758–765.

Guerin C, Laterra J, Drewes L, Brem H, Goldstein GW. 1992. Vascular expression of glucose transporter in experimental brain neoplasms. Am J Pathol 140:114–125.

Hansen AJ, Lund-Andersen H, Crone C. 1977. K^+- permeability of the blood-brain barrier, investigated by aid of a K^+-sensitive microelectrode. Acta Physiol Scand 101:438–445.

Jennett B, Teasdale G. 1981. *Management of Head Injuries*. Philadelphia, PA: Davis.

Lal B, Indurti RR, Couraud P-O, Goldstein GW, Laterra J. 1994. Endothelial cell implantation and survival within experimental gliomas. Proc Natl Acad Sci U S A 21:9695–9699.

Reese TS, Karnovsky MJ. 1967. Fine structural localization of a blood-brain barrier to exogenous peroxidase. J Cell Biol 34:207–217.

P.1301

Resnick L, Berger JR, Shapshak P, Tourtellotte WW. 1988. Early penetration of the blood-brain barrier by HIV. Neurology 38:9–14.

Rubin LL, Hall DE, Porter S, Barbu K, Cannon C, Horner HC, Janatpour M, Liaw CW, Manning K, Morales J, Tanner LI, Toamselli KJ, Bard F. 1991. A cell culture model of the blood-brain barrier. J Cell Biol 115:1725–1735.

Schinkel AH, Smit JJM, Vantellingen O, Beijnen JH, Wagenaar E, Vandeemter L, Mol CAAM, Vandervalk MA, Robanusmaandag EC, Teriele HPJ, Berns AJM, Borst P. 1994. Disruption of the mouse *mdr1a* P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell 77:491–502.

Shivers RR, Betz AL, Goldstein GW. 1984. Isolated rat brain capillaries possess intact, structurally complex, interendothelial tight junctions; freeze-fracture verification of tight junction integrity. Brain Res 324:313–322.

Spector R, Johanson CE. 1989. The mammalian choroid plexus. Sci Am 261(5):68–74.

Stewart PA, Wiley MJ. 1981. Developing nervous tissue induces formation of blood-brain barrier characteristics in invading endothelial cells: a study using quail-chick transplantation chimeras. Dev Biol 84:183–192.

Svengaard N, Bjorklund A, Hardebo JE, Stenevi U. 1975. Axonal degeneration associated with a defective blood-brain barrier in cerebral implants. Nature 255:334–336.

Wahl M, Unterberg A, Baerthmann A, Schilling L. 1988. Mediators of blood-brain barrier dysfunction and formation of vasogenic brain edema. J Cereb Blood Flow Metab 8:621–634.

Wolburg H, Neuhaus J, Kniesel U, Kraub B, Schmid E-M, Ocalan M, Farrell C, Risau W. 1994. Modulation of tight junction structure in blood-brain barrier endothelial cells: Effects of tissue culture, second messengers, and cocultured astrocytes. J Cell Sci 107:1347–1357.

P.1302

Appendix C

Circulation of the Brain

John C.M. Brust

THE BRAIN IS HIGHLY vulnerable to disturbance of its blood supply. Anoxia and ischemia lasting only seconds cause neurological symptoms and when they last minutes they can cause irreversible neuronal damage.

Blood flow to the central nervous system must efficiently deliver oxygen, glucose, and other nutrients and remove carbon dioxide, lactic acid, and other metabolic products. The cerebral vasculature has unique anatomical and physiological features that protect the brain from circulatory compromise. When these protective mechanisms fail, the result is a stroke. Broadly defined, the term *stroke*, or *cerebrovascular accident*, refers to the neurological symptoms or signs, usually focal and acute, that result from diseases involving blood vessels.

The Blood Supply Of The Brain Can Be Divided Into Arterial Territories

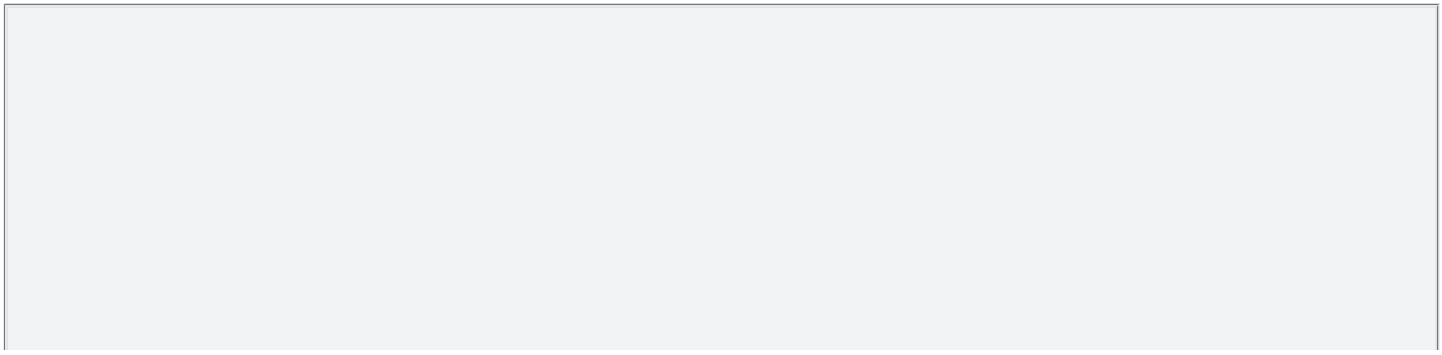
Each cerebral hemisphere is supplied by an *internal carotid artery*, which arises from a common carotid artery beneath the angle of the jaw, enters the cranium through the carotid foramen, traverses the cavernous sinus (giving off the *ophthalmic artery*), penetrates the dura, and divides into the anterior and middle cerebral arteries (Figure C-1).

The large surface branches of the *anterior cerebral artery* supply the cortex and white matter of the inferior frontal lobe, the medial surface of the frontal and parietal lobes, and the anterior corpus callosum. Smaller penetrating branches—including the so-called *recurrent artery of Heubner*—supply the deeper cerebrum and diencephalon, including limbic structures,

P.1303

P.1304

the head of the caudate, and the anterior limb of the internal capsule.



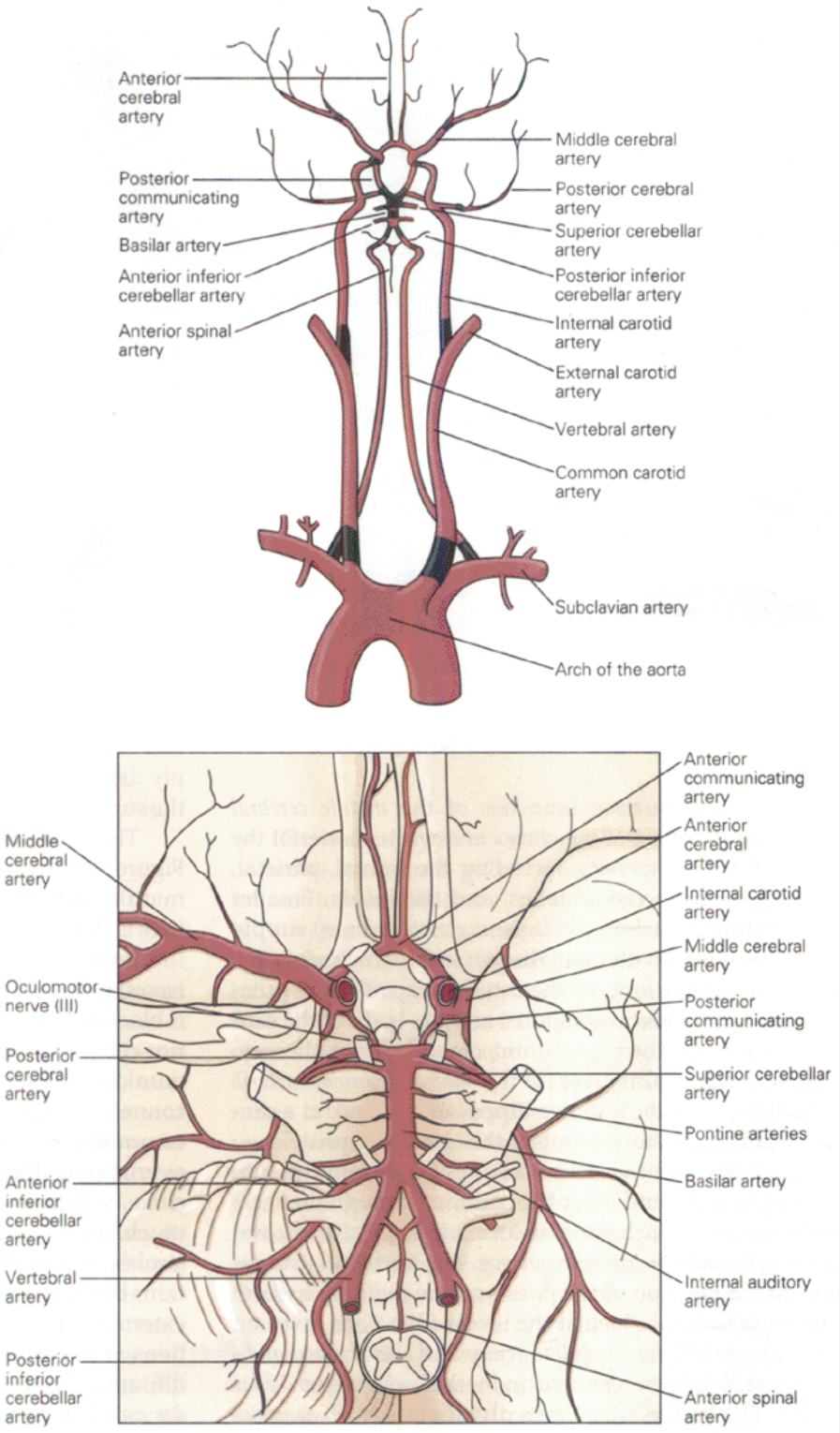


Figure C-1 The blood vessels of the brain. The circle of Willis is made up of the proximal posterior cerebral arteries, the posterior communicating arteries, the internal carotid arteries just before their bifurcations, the proximal anterior cerebral arteries, and the anterior communicating artery. *Dark areas* show common sites of atherosclerosis and occlusion.

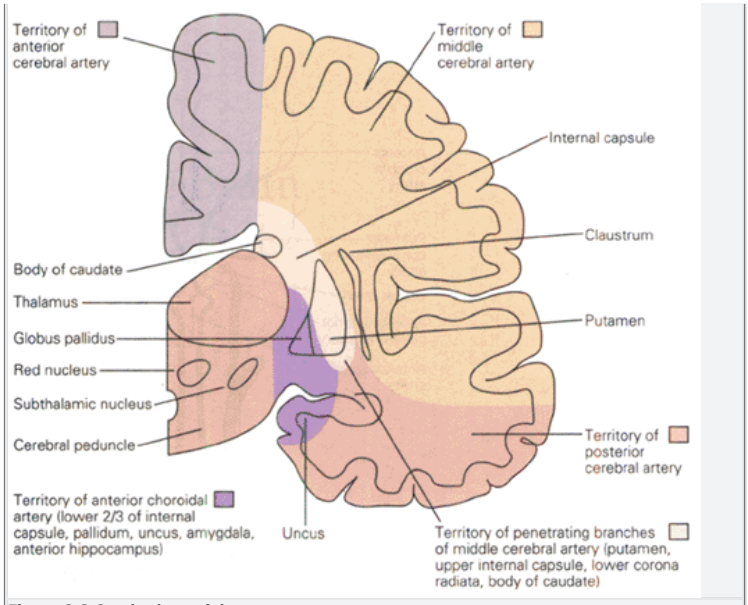


Figure C-2 Cerebral arterial areas.

The large surface branches of the *middle cerebral artery* supply most of the cortex and white matter of the hemisphere's convexity, including the frontal, parietal, temporal, and occipital lobes, and the insula. Smaller penetrating branches (the *lenticulostriate arteries*) supply the deep white matter and diencephalic structures, such as the posterior limb of the internal capsule, the putamen, the outer globus pallidus, and the body of the caudate. After the internal carotid emerges from the cavernous sinus, it also gives off the *anterior choroidal artery*, which supplies the anterior hippocampus and, at a caudal level, the posterior limb of the internal capsule.

Left and right vertebral arteries arise from the subclavian arteries and enter the cranium through the foramen magnum. Each gives off an *anterior spinal artery* and a *posterior inferior cerebellar artery*. The vertebral arteries join at the junction of the pons and the medulla to form the *basilar artery*, which at the level of the pons gives off the *anterior inferior cerebellar artery* and the *internal auditory artery* and at the midbrain the *superior cerebellar* artery. The basilar artery then divides into the two posterior *cerebral arteries*, which supply the inferior temporal and medial occipital lobes and the posterior corpus callosum. The smaller penetrating branches of these vessels (the *thalamoperforate and thalamogeniculate arteries*) supply diencephalic structures, including the thalamus and the subthalamic nuclei, as well as parts of the midbrain.

These arterial territories are shown schematically in Figure C-2. Infarctions in the territories of the anterior, middle, and posterior cerebral arteries are shown in Figures C-3, C-4, and C-5.

Interconnections between blood vessels (*anastomoses*) protect the brain when part of its vascular supply is blocked (Figure C-6). At the *circle of Willis* the two anterior cerebral arteries are connected by the anterior communicating artery, and the posterior cerebral arteries are connected to the internal carotid arteries by the posterior communicating arteries. The circle of Willis provides an overlapping blood supply. A congenitally incomplete circle, which is common in the general population, is much more frequent among patients who have had strokes. Other important anastomoses include connections between the ophthalmic artery and branches of the external carotid artery through the orbit, and connections at the brain surface between branches of the middle, anterior, and posterior cerebral arteries (*sharing border zones or watersheds*). The small penetrating vessels arising from the circle of Willis and proximal major arteries tend to lack anastomoses. The deep brain regions they supply are therefore referred to as *end zones*.

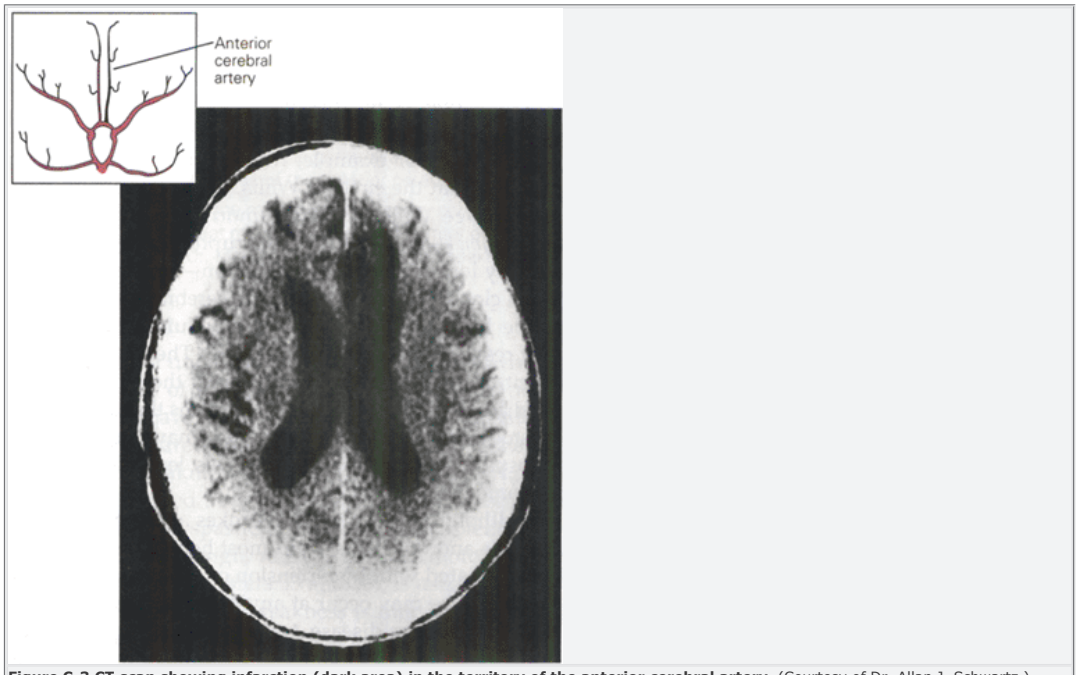


Figure C-3 CT scan showing infarction (dark area) in the territory of the anterior cerebral artery. (Courtesy of Dr. Allan J. Schwartz.)

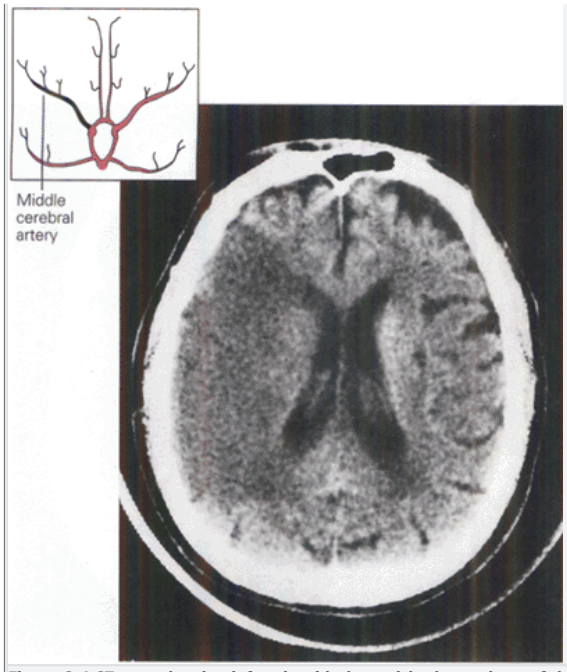


Figure C-4 CT scan showing infarction (dark area) in the territory of the middle cerebral artery. (Courtesy of Dr. Allan J. Schwartz.)

P.1305

The Crebral Vessels Have Unique Physiological Responses

Although the human brain constitutes only 2% of total body weight, it receives about 15% of the cardiac output and consumes approximately 20% of the oxygen used by the entire body. These values reflect the high metabolic rate and oxygen requirements of the brain. The total blood flow to the brain is about 750–1000 ml/min; about 350 ml of this amount flows through each carotid artery and about 100–200 ml flows through the vertebrobasilar system. Flow per unit mass of gray matter is approximately four times that of white matter.

Cerebral vessels are capable of altering their own diameter and can respond in a unique fashion to altered physiological conditions. Two main types of autoregulation exist. Brain arterioles constrict when the systemic blood pressure is raised and dilate when it is lowered. These adjustments help to maintain optimal cerebral blood flow. The result is that normal individuals have a constant cerebral blood flow between mean arterial pressures of approximately 60–150 mm Hg. Above or below these pressures cerebral blood flow rises or falls linearly.

The second type of autoregulation involves blood or tissue gases and pH. When arterial CO_2 is raised, brain arterioles dilate and cerebral blood flow increases; with hypocarbia, vasoconstriction results and cerebral blood flow decreases. The response is very sensitive: inhalation of 5% CO_2 increases cerebral blood flow by 50%; 7% CO_2 doubles it. Changing arterial O_2 causes an opposite and less pronounced response: Breathing 100% O_2 lowers cerebral blood flow by about 13%; 10% O_2 raises it by 35%. The mechanism of these responses is uncertain. The vasodilatory action of arterial CO_2 is probably mediated by alterations in extracellular pH. Local concentrations of K^+ and adenosine, both of which cause vasodilation in animals, may play a role.

Whatever the mechanism, these responses protect the brain by increasing the delivery of oxygen and the removal of acid metabolites in the presence of hypoxia, ischemia, or tissue damage. They also allow nearly instantaneous adjustments of regional cerebral blood flow to meet the demands of rapidly changing oxygen and glucose metabolism that accompany normal brain activities. For example, viewing a complex scene will increase oxygen and glucose consumption in the visual

P.1306

cortex of the occipital lobes. The resulting increased CO_2 concentration and lowered pH in the area cause an immediate local increase in blood flow.

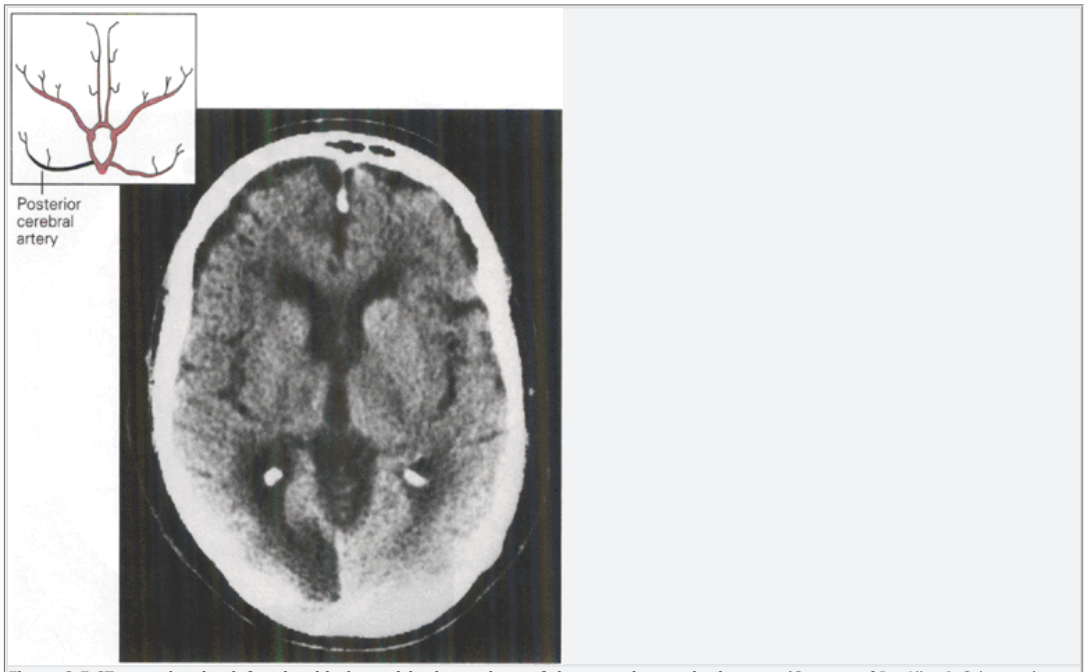


Figure C-5 CT scan showing infarction (dark area) in the territory of the posterior cerebral artery. (Courtesy of Dr. Allan J. Schwartz.)

A Stroke Is the Result of Diseases Involving Blood Vessels

Diseases of the blood vessels are among the most frequent serious neurological disorders, ranking third as a cause of death in the adult population in the United States and probably first as a cause of chronic functional incapacity. Approximately two million people in the United States today are impaired by the neurological consequences of cerebrovascular disease. Many of them are between 25 and 64 years of age.

Strokes are either *occlusive* (due to closure of a blood vessel) or *hemorrhagic* (due to bleeding from a vessel). Insufficiency of blood supply is termed *ischemia;* if it is temporary, symptoms and signs may clear with little or no pathological evidence of tissue damage. Ischemia is not synonymous with *anoxia*, for a reduced blood supply deprives tissue not only of oxygen but also of glucose. In addition, it prevents the removal of potentially toxic metabolites such as lactic acid. When ischemia is sufficiently severe and prolonged, neurons and other cellular elements die; this condition is called

infarction.

Hemorrhage may occur at the brain surface (*extraparenchymal*), for example, from rupture of congenital aneurysms at the circle of Willis, causing subarachnoid hemorrhage. Alternatively, hemorrhage may be *intraparenchymal* —for example, from rupture of vessels damaged by long-standing hypertension—and may cause a blood clot or *hematoma* within the cerebral hemispheres, in the brain stem, or in the cerebellum. Hemorrhage may result in ischemia or infarction. The mass effect of an intracerebral hematoma may limit the blood supply of adjacent brain tissue. By mechanisms that are not understood, subarachnoid hemorrhage may cause reactive vasospasm of cerebral surface vessels, leading to further ischemic brain damage.

Although most occlusive strokes are due to atherosclerosis and thrombosis and most hemorrhagic strokes are associated with hypertension or aneurysms, strokes of either type may occur at any age from many causes, including cardiac disease, trauma, infection, neoplasm, blood dyscrasia, vascular malformation, immunological disorder, and exogenous toxins. Diagnostic strategies and treatment should vary accordingly; however, in this appendix we examine the anatomical and physiological principles relevant to *any* occlusive or hemorrhagic stroke.

Clinical Vascular Syndromes May Follow Vessel Occlusion, Hypoperfusion, Or Hemorrhage

Infarction Can Occur in the Middle Cerebral Artery Territory

Infarction in the territory of the middle cerebral artery (cortex and white matter) causes the most frequently encountered stroke syndrome, with contralateral weakness, sensory loss, and visual field impairment (*homonymous hemianopsia*), and, depending on the hemisphere involved, either language disturbance or impaired spatial perception. Weakness and sensory loss affect the face and arm more than the leg because of the somatotopy of the motor and sensory cortex (pre- and postcentral gyri). The face- and arm-control areas are on the convexity, whereas the leg-control area is on the medial surface of the hemisphere. Motor and sensory loss are greatest in the hand, for the more proximal limbs and the trunk tend to have greater representation in both hemispheres. Paraspinal muscles, for example, are

rarely weak in unilateral cerebral lesions. Similarly, the facial muscles of the forehead and the muscles of the pharynx and jaw are represented in both hemispheres and therefore are usually spared. Tongue weakness is variable. If weakness is severe (*plegia*), muscle tone is usually decreased at first but gradually increases over days or weeks to spasticity with hyperactive tendon reflexes. A Babinski sign, reflecting upper motor neuron disturbance (see Chapters 33 and 38), is usually present from the outset. When weakness is mild, or during recovery, there may be clumsiness or slowness of movement out of proportion to loss of strength; such motor disability may resemble parkinsonian bradykinesia or even cerebellar ataxia.

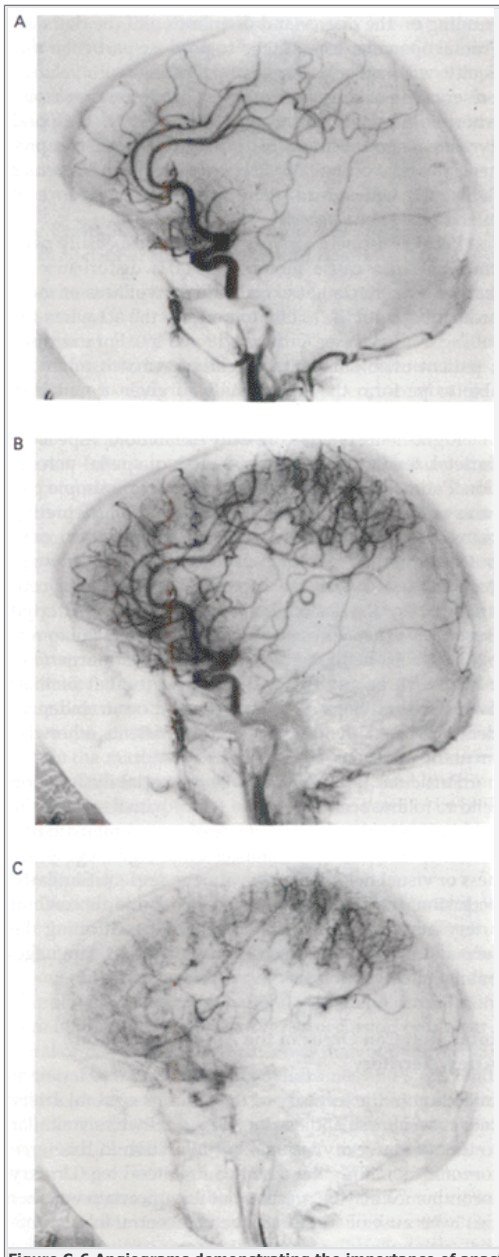


Figure C-6 Angiograms demonstrating the importance of anastomoses in that they allow retrograde filling after occlusion of the middle cerebral artery.

A. Occlusion of the middle cerebral artery results in no filling in the middle cerebral distribution.

B. Retrograde filling of the middle cerebral artery has begun via distal anastomotic branches of the anterior cerebral artery.

C. Retrograde filling of the middle cerebral artery continues at a time when little contrast material is seen in the anterior cerebralartery. (Courtesy of Dr. Margaret Whelan and Dr. Sadek K. Hilal.)

Acutely, paresis of contralateral conjugate gaze often occurs as a result of damage to the convexity of the cortex anterior to the motor cortex (the frontal eye field). The reason this gaze palsy persists for only 1 or 2 days, even when other signs remain severe, is not clear.

Sensory loss tends to involve discriminative and proprioceptive modalities more than affective modalities. Pain and temperature sensation may be impaired or altered but are usually not lost. Joint position sense, however, may be severely disturbed, causing limb ataxia, and there may be loss of two-point discrimination, *astereognosis* (inability to recognize a held object by tactile sensation), or *extinction* (failure to appreciate a touch stimulus if a comparable stimulus is delivered simultaneously to the unaffected side of the body).

Homonymous hemianopsia is the result of damage to the optic radiations, the deep fiber tracts connecting the thalamic lateral geniculate body to the visual (*calcarine*) cortex. If the parietal radiation is

primarily affected, the visual field loss may be an inferior quadrantanopia, whereas in temporal lobe lesions quadrantanopia may be superior.

In more than 95% of right-handed persons and in the majority of those who are left handed, the left hemisphere is dominant for language. Destruction of left opercular (*perisylvian*) cortex in left-dominant individuals causes aphasia, which may take several forms de-

P.1308

pending on the degree and distribution of the damage. Frontal opercular lesions tend to produce particular difficulty with speech output and writing with relative preservation of language comprehension (*Broca aphasia*), whereas infarction of the posterior superior temporal gyrus tends to cause severe difficulty in speech comprehension and reading (*Wernicke aphasia*). When opercular damage is widespread, there is severe disturbance of mixed type (*global aphasia*).

Left-hemisphere convexity damage, especially parietal, may also cause motor apraxia, a disturbance of learned motor acts not explained by weakness or incoordination, with the ability to perform the act when the setting is altered (see Chapters 19 and 59). For example, a patient unable to imitate lighting a match might be able to perform the act normally if given a match to light.

Right-hemisphere convexity infarction, especially parietal, tends to cause disturbances of spatial perception. Patients may have difficulty in copying simple pictures or diagrams (*constructional apraxia*), in interpreting maps or finding their way about (*topographagnosia*), or in putting on their clothing properly (*dressing apraxia*). Awareness of space and the patient's own body contralateral to the lesion may be particularly affected (*hemi-inattention or hemineglect*). Patients may fail to recognize their hemiplegia (*anosognosia*), left arm (*asomatognosia*), or any external object to the left of their own midline. Such phenomena may occur independently of visual field defects and in patients otherwise mentally intact (see Chapter 20).

Particular types of language or spatial dysfunction tend to follow occlusion not of the proximal stem of the middle cerebral artery but of one of its several main pial branches. In such circumstances other signs (eg, weakness or visual field defect) may not be present. Similarly, occlusion of the rolandic branch of the middle cerebral artery may cause motor and sensory loss affecting the face and arm without disturbance of vision, language, or spatial perception.

Infarction Can Occur in the Anterior Cerebral Artery Territory

Infarction in the territory of the anterior cerebral artery causes weakness and sensory loss qualitatively similar to that of convexity lesions, but infarction in this territory affects mainly the distal contralateral leg. Urinary incontinence may be present, but it is uncertain whether this is because of a lesion of the paracentral lobule (medial hemispheric motor and sensory cortices) or of a more anterior region necessary for the inhibition of bladder emptying. Damage to the supplementary motor cortex may cause speech disturbance, considered aphasia by some and a type of motor inertia by others. Involvement of the anterior corpus callosum may cause apraxia of the left arm (*sympathetic apraxia*), which is attributed to disconnection of the left (languagedominant) hemisphere from the right motor cortex.

Bilateral anterior cerebral artery territory infarction (occurring, for example, when both arteries arise anomalously from a single trunk) may cause a severe behavioral disturbance, known as *abulia*, consisting of profound apathy, motor inertia, and muteness, and attributable to destruction, in variable combinations, of the inferior frontal lobes (orbitofrontal cortex), deeper limbic structures, supplementary motor cortices, and cingulate gyri.

Infarction Can Occur in the Posterior Cerebral Artery Territory

Infarction in the territory of the posterior cerebral artery causes contralateral homonymous hemianopsia by destroying the calcarine cortex. Macular (central) vision tends to be spared because the occipital pole, where macular vision is represented, receives blood supply from the middle cerebral artery. If the lesion is on the left and the posterior corpus callosum is affected, *alexia*—inability to read—may be present without aphasia or agraphia. A possible explanation for the alexia is the disconnection of the seeing right occipital cortex from the language-dominant left hemisphere. If infarction is bilateral (eg, following thrombosis at the point where both posterior cerebral arteries arise from the basilar artery), there may be cortical blindness with failure of the patient to recognize that he cannot see (*Anton syndrome*), or memory disturbance may occur as a result of bilateral damage to the inferomedial temporal lobes.

If posterior cerebral artery occlusion is proximal, the lesion may include, or especially affect, the following structures: the thalamus, causing contralateral hemisensory loss and sometimes spontaneous pain and dysesthesia (*thalamic pain syndrome*); the subthalamic nucleus, causing contralateral severe proximal chorea (*hemiballism*); or even the midbrain, with ipsilateral oculomotor palsy and contralateral hemiparesis or ataxia from involvement of the corticospinal tract or the crossed superior cerebellar peduncle (dentatothalamic tract).

The Anterior Choroidal and Penetrating Arteries Can Become Occluded

Anterior choroidal artery occlusion can cause contralateral hemiplegia and sensory loss from involvement of

P.1309

the posterior limb of the internal capsule and homonymous hemianopsia from involvement of the thalamic lateral geniculate nucleus.

As mentioned earlier in this chapter, the deeper cerebral white matter and diencephalon are supplied by small penetrating arteries, variously called the *lenticulostriates*, the *thalamogeniculates*, or the *thalamoperforates*, which arise from the circle of Willis or the proximal portions of the middle, anterior, and posterior cerebral arteries. These end-arteries lack anastomotic interconnections, and occlusion of individual vessels, usually in association with hypertensive damage to the vessel wall, causes small (less than 1.5 cm in diameter) infarcts (*lacune*s), which, if critically located, are followed by characteristic syndromes. For example, lacunes in the pyramidal tract area of the internal capsule cause *pure hemiparesis*, with arm and leg weakness of equal severity and with little or no sensory loss, visual field disturbance, aphasia, or spatial disruption. Lacunes in the ventral posterior nucleus of the thalamus produce pure hemisensory loss, with involvement of pain, temperature, proprioceptive, and discriminative modalities and with little motor, visual, language, or spatial disturbance. Most lacunes occur in redundant areas (eg, nonpyramidal corona radiata) and so are asymptomatic. If bilateral and numerous, however, they may cause a characteristic syndrome (état lacunaire) of progressive dementia, shuffling gait, and pseudobulbar palsy (spastic dysarthria and dysphagia, with lingual and pharyngeal paralysis and hyperactive palate and gag reflexes, plus lability of emotional response, with abrupt crying or laughing out of proportion to mood).

Infarction restricted to structures supplied by the recurrent artery of Heubner or other deep penetrating branches of the anterior cerebral artery (the anterior caudate nucleus and less predictably the anterior putamen and anterior limb of the internal capsule) results in varying combinations of psychomotor slowing, dysarthria, agitation, contralateral neglect, and, when left hemispheric, language disturbance.

The Carotid Artery Can Become Occluded

Atherothrombotic vessel occlusion often occurs in the internal carotid artery rather than the intracranial vessels. Particularly in a patient with an incomplete circle of Willis, infarction may include the territories of both the middle and anterior cerebral arteries, with arm and leg weakness and sensory loss equally severe. Alternatively, infarction may be limited to the distal shared territory (border zones of these vessels), producing, by destruction of the motor cortex at the upper cerebral convexity, weakness limited to the arm or the leg. Another cause of leg weakness and sensory loss in association with a convexity syndrome is occlusion of the middle cerebral artery at its proximal stem; the internal capsule and other diencephalic structures supplied by the middle cerebral artery's lenticulostriate branches are then affected in addition to the cortex of the cerebral convexity.

The Brain Stem and Cerebellum Are Supplied by Branches of the Vertebral and Basilar Arteries

Branches of the vertebral and basilar arteries consist of three sets: (1) Paramedian branches, including the anterior spinal artery, supply midline structures; (2) short circumferential branches supply more lateral structures, including the inferior, middle, and superior cerebral peduncles; and (3) long circumferential arteries—the posterior inferior, anterior inferior, and superior cerebellar arteries—also supply lateral brain stem structures and the cerebellar peduncles, as well as the cerebellum itself. Most of the midbrain is supplied by branches of the posterior cerebral artery. The interpeduncular branches, the most medial branches located between the basilar artery bifurcation and the posterior communicating arteries, supply paramedian midbrain structures. Lateral to this group are the thalamoperforate branches, which supply the inferior, medial, and anterior thalamus and the subthalamic nucleus. Further laterally are the thalamogeniculate branches, which supply lateral and dorsal structures in the midbrain and thalamus. In some people the midbrain also receives blood from the superior cerebellar, posterior communicating, and anterior choroidal arteries. After the posterior cerebral artery passes around the midbrain, it enters the middle fossa to supply the occipital and inferior temporal lobes. It does not supply the cerebellum.

Various syndromes resulting from damage to specific brain stem structures have been defined (Figure C-7). With the exception of the lateral medullary syndrome of Wallenberg, however, most of the original descriptions were based on patients with neoplasms. Brain stem infarction more often follows occlusion of the vertebral or basilar arteries themselves than of their medial or lateral branches; resulting syndromes and signs tend to be less stereotyped than classical descriptions imply.

Generally speaking, a lesion of the posterior fossa is suggested by (1) bilateral long tract (motor or sensory) signs, (2) crossed (eg, left face and right limb) motor or sensory signs, (3) cerebellar signs, (4) stupor or coma (from involvement of the ascending reticular activating system), (5) disconjugate eye movements or nystagmus, and (6) involvement of cranial nerves not usually affected by unilateral hemispheric infarcts (eg, unilateral

P.1310

deafness or pharyngeal weakness). Sometimes a lesion involving only a single tiny structure can be localized accurately by symptomatology. For example, internuclear opthalmoplegia implicates a lesion of the median longitudinal fasciculus. Other lesions produce more ambiguous symptoms. For example, infarction limited to the upper pontine corticospinal tract can produce contralateral face, arm, and leg weakness indistinguishable from that caused by a small infarct in the internal capsule.

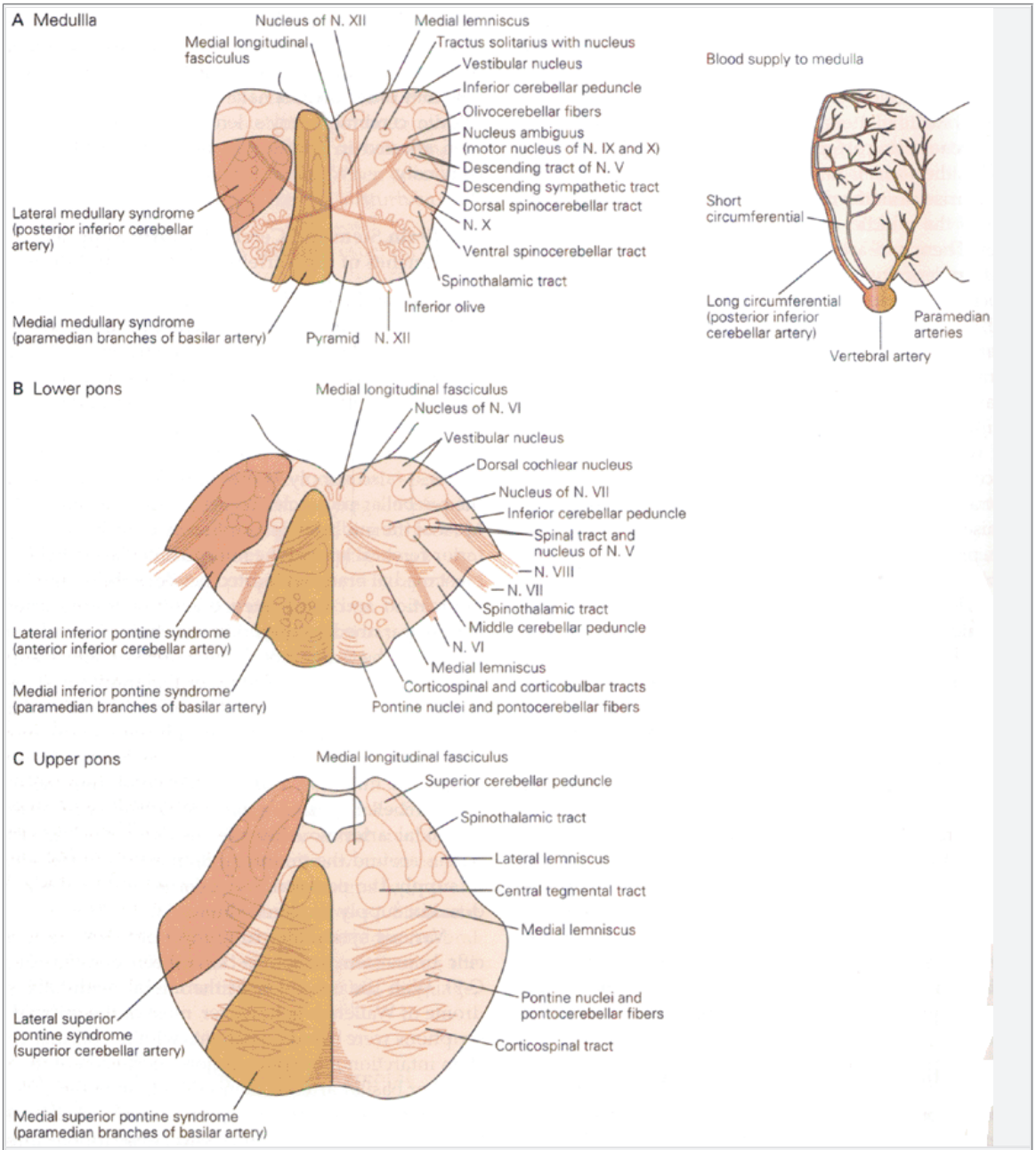
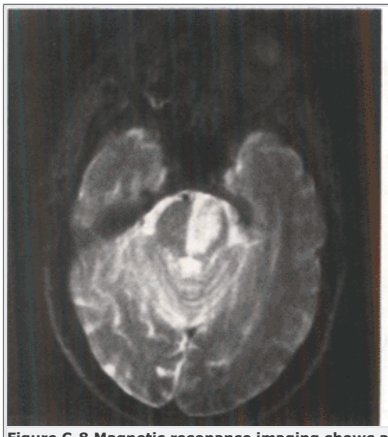


Figure C-7 Syndromes of brain stem vascular lesions (indicated on the left in each figure).



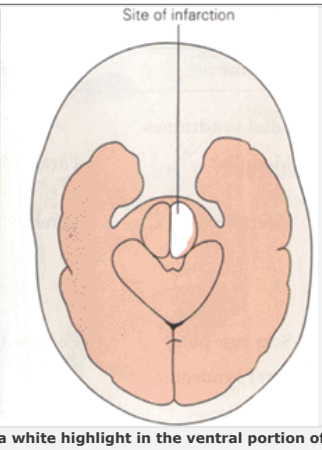


Figure C-8 Magnetic resonance imaging shows a white highlight in the ventral portion of one half of the pons. The lesion stops abruptly at the midline, suggesting unilateral occlusion of one or more paramedian vessels.

Infarcts Affecting Predominantly Medial or Lateral Brain Stem Structures Produce Characteristic Syndromes

Infarction of the lateral medulla follows occlusion of the vertebral artery or less often the posterior inferior cerebellar artery. Symptoms and signs include (1) vertigo, nausea, vomiting, and nystagmus (from involvement of the vestibular nuclei); (2) ataxia of gait and ipsilateral limbs (inferior cerebellar peduncle or the cerebellum itself); (3) decreased pain and temperature (but not touch) sensation on the ipsilateral face (descending tract and nucleus of the trigeminal nerve) and the contralateral body (spinothalamic tract); (4) dysphagia, hoarseness, ipsilateral weakness of the palate and vocal cords, and ipsilaterally decreased gag reflex (nucleus ambiguus, or glossopharyngeal and vagus outflow tracts); and (5) ipsilateral Horner syndrome (descending sympathetic fibers). Involvement of the nucleus solitarius can cause ipsilateral loss of taste, and hiccups are often present (Table C-1).

Infarction of the medial medulla causes contralateral hemiparesis (from involvement of the corticospinal tract), ipsilateral tongue weakness with dysarthria and deviation toward the paretic side (hypoglossal nucleus or outflow tract), and contralateral impaired proprioception and discriminative sensation with preserved pain and temperature sensation (medial lemniscus).

Infarction of the lateral pons affects caudal structures when the anterior inferior cerebellar artery is occluded and rostral structures when the superior cerebellar artery is occluded (Figure C-8). Symptoms of caudal damage resemble those of lateral medullary infarction, with vertigo, nystagmus, gait and ipsilateral limb ataxia, crossed face-and-body pain and temperature loss, Horner syndrome, and ipsilateral loss of taste. There is also unilateral tinnitus and deafness (from involvement of the cochlear nuclei). Involvement of more medial structures can cause ipsilateral gaze paresis or facial weakness. Symptoms of rostral damage include gait and ipsilateral limb ataxia, Horner syndrome, and crossed sensory loss, which at this level includes touch, pain, and temperature sensation on the ipsilateral face (from involvement of the primary sensory nucleus or entering sensory fibers of the trigeminal nerve). There may also be ipsilateral jaw weakness with deviation to the paretic side (trigeminal motor nucleus and outflow tract). Vertigo, deafness, and face weakness are not present.

Infarction of the medial pons, whether caudal or rostral, causes contralateral hemiparesis (from involvement of the corticospinal tract). Caudal lesions affecting the facial nucleus or outflow tract cause ipsilateral facial weakness. Rostral lesions result in contralateral facial weakness. There may also be ipsilateral gaze paresis (paramedian pontine reticular formation or abducens nucleus, together comprising the *pontine gaze center*) or abducens paresis (sixth nerve outflow tract); internuclear ophthalmoplegia and limb and gait ataxia are often present. Contralateral impairment of proprioception and discriminative touch is most prominent with caudal lesions. Rapid involuntary movements of the palate—so-called *palatal myoclonus*—has been attributed to involvement of the central tegmental tract; theinvoluntary movements may spread to include the pharynx, larynx, face, eyes, or respiratory muscles.

Table C-1 Signs That indicate the Level of Brain Stem Vascular Syndromes

Syndrome	Artery affected	Structure involved	Manifestations
Medial syndromes			
Medulla	Paramedian branches	Emerging fibers of twelfth nerve	Ipsilateral hemiparalysis of tongue
Inferior pons	Paramedian branches	Pontine gaze center Emerging fibers of sixth nerve	Paralysis of gaze to side of lesion Ipsilateral abduction paralysis
Superior pons	Paramedian branches	Medial longitudinal fasciculus	Internuclear ophthalmoplegia
Lateral syndromes			
Medulla	Posterior inferior cerebellar	Emerging fibers of ninth or tenth nerves Vestibular nuclei Descending tract and nucleus of fifth nerve Solitary nucleus and tract	Dysphagia; hoarseness; ipsi-lateral paralysis of vocal cord; ipsilateral loss of pharyngeal reflex Vertigo; nystagmus Ipsilateral facial analgesia Taste loss on ipsilateral half of tongue posteriorly
Inferior pons	Anterior inferior cerebellar	Emerging fibers of seventh nerve Solitary nucleus and tract Cochlear nuclei	Ipsilateral facial paralysis Taste loss on ipsilateral half of tongue anteriorly Deafness; tinnitus
Mid pons		Motor nucleus of fifth nerve Emerging sensory fibers of fifth nerve	Ipsilateral jaw weakness Ipsilateral facial numbness

Syndromes of midbrain infarction are more conveniently characterized as peduncular (ventral), tegmental, or pretectal/tectal (dorsal) (Figure C-9). Unilateral peduncular lesions cause *Weber syndrome*, characterized by ipsilateral paresis of adduction and vertical gaze and pupillary dilation (involvement of oculomotor nerve outflow tract) and contralateral face, arm, and leg paresis (corticospinal and corticobulbar tracts). Unilateral tegmental lesions cause *Claude syndrome*, characterized by oculomotor paresis (oculomotor nucleus) and contralateral ataxia and tremor (often referred to as *rubral tremor* but probably the result of damage to the brachium conjunctivum). Lesions affecting both the peduncle and tegmentum produce combinations of oculomotor paresis, ataxia, and weakness (*Benedikt syndrome*). Dorsal midbrain lesions, which are infrequently vascular and rarely unilateral, cause *Parinaud syndrome*, characterized by impaired vertical gaze—especially upward (posterior commissure and the rostral interstitial nucleus of the median longitudinal fasciculus)—and loss of the pupillary light reflex (pretectal structures).

Bilateral Brain Stem Lesions Can Have Devastating Consequences

Bilateral paramedian infarction of the upper brain stem can involve the reticular activating system and cause obtundation or coma. Bilateral damage of the proximal posterior cerebral artery produces altered consciousness and various combinations of mesencephalic, diencephalic, and cortical signs (*top of the basilar artery syndrome*), affecting eye movements, pupils, vision

(including, even without visual loss, vivid formed hallucinations), sensation, coordination, memory, and behavior (including agitated delirium).

Bilateral corticospinal tract infarction causes quadriparesis or quadriplegia; additional cranial weakness depends on how rostral or caudal the lesion is. Bilateral infarction of the rostral basis pontis (with sparing of the tegmentum) produces paralysis of all muscles except eye movements, the so-called *locked-in state*. Such a patient may appear comatose yet is awake and able to communicate through eye movements. Less severe damage to the corticospinal tracts in the rostral brain stem can cause pseudobulbar palsy, with spastic quadriparesis, dysarthria, dysphagia, a hyperactive gag reflex, and for reasons unclear, labile emotional responses with explosive crying or laughing.

Infarction Can Be Restricted to the Cerebellum

Infarcts of the inferior cerebellum, which has extensive vestibular connections, can cause vertigo, nausea, and nystagmus without other symptoms, suggesting disease of the inner ear or vestibular nerve. (Similar symptoms, with or without tinnitus or deafness, can occur following occlusions of the internal auditory artery, arising from the basilar artery and supplying the peripheral labyrinth.) More superior cerebellar infarcts produce gait and ipsilateral limb ataxia.

Infarction Can Affect the Spinal Cord

The ventral spinal cord is supplied by a single anterior spinal artery; the dorsal spinal cord is supplied by two or more posterior spinal arteries (Figure C-10). Except most rostrally, where the anterior spinal artery arises from the joining of two vertebral artery branches, the anterior and posterior spinal arteries are fed along their course by several radicular arteries, which arise from segmental branches of the aorta and iliac arteries. Moreover, whereas the posterior spinal arteries are longitudinally continuous and anastomotic, the anterior spinal artery's continuity is tenuous, making the anterior spinal cord more dependent on its segmental supply. The upper thoracic spinal cord is especially vulnerable in this regard.

Vascular anatomy explains the characteristic pattern of spinal cord infarction. Vessel occlusion is usually in a proximal segmental artery; because of the anastomotic continuity of the posterior spinal arteries, infarction tends to be limited to the anterior spinal artery territory. There is paraparesis or quadriparesis (corticospinal tracts), loss of bladder and bowel control, and loss of pain and temperature sensation below the lesion (spinothalamic tracts), but proprioception and discriminative touch (dorsal columns) are spared. If the cervical or lumbar spinal cord is involved, atrophic weakness of upper or lower extremity muscles (anterior horns) can occur. Because the anterior spinal artery gives off sulcal arteries that alternately enter the left and right halves of the spinal cord, infarction can sometimes produce a *Brown-Séquard syndrome*, with ipsilateral weakness and contralateral loss of pain and temperature sensation.

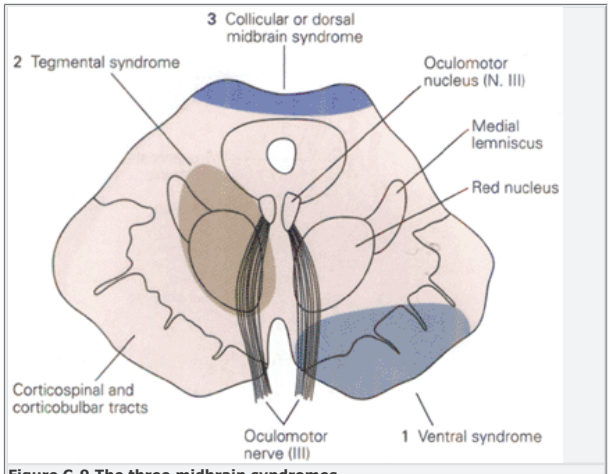


Figure C-9 The three midbrain syndromes.

Diffuse Hypoperfusion Can Cause Ischemia or Infarction

Brain ischemia or infarction may accompany diffuse hypoperfusion (shock). In such circumstances the most vulnerable regions are often the border zones between large arterial territories and the end

zones of deep penetrating vessels. Whatever the cause of reduced cerebral perfusion, signs tend to be bilateral. Paralysis and sensory loss may be present in both arms (from bilateral infarction of the cortex at the junction of the middle and anterior arterial supply, affecting the arm-control area of the motor and sensory cortex).

Disturbed vision or memory may result (from infarction of the occipital or temporal lobes at the junction of the middle and posterior cerebral arterial supply). There may also be ataxia (from cerebellar border-zone infarction) or abnormal movements such as chorea or myoclonus (presumably from involvement of the basal ganglia). Such signs may exist alone or in combination and may be accompanied by a variety of aphasic or other cognitive disturbances.

Hypotension can also cause spinal cord infarction, most often upper thoracic, affecting either the territory of the anterior spinal artery or the border zone between the anterior and posterior spinal arteries.

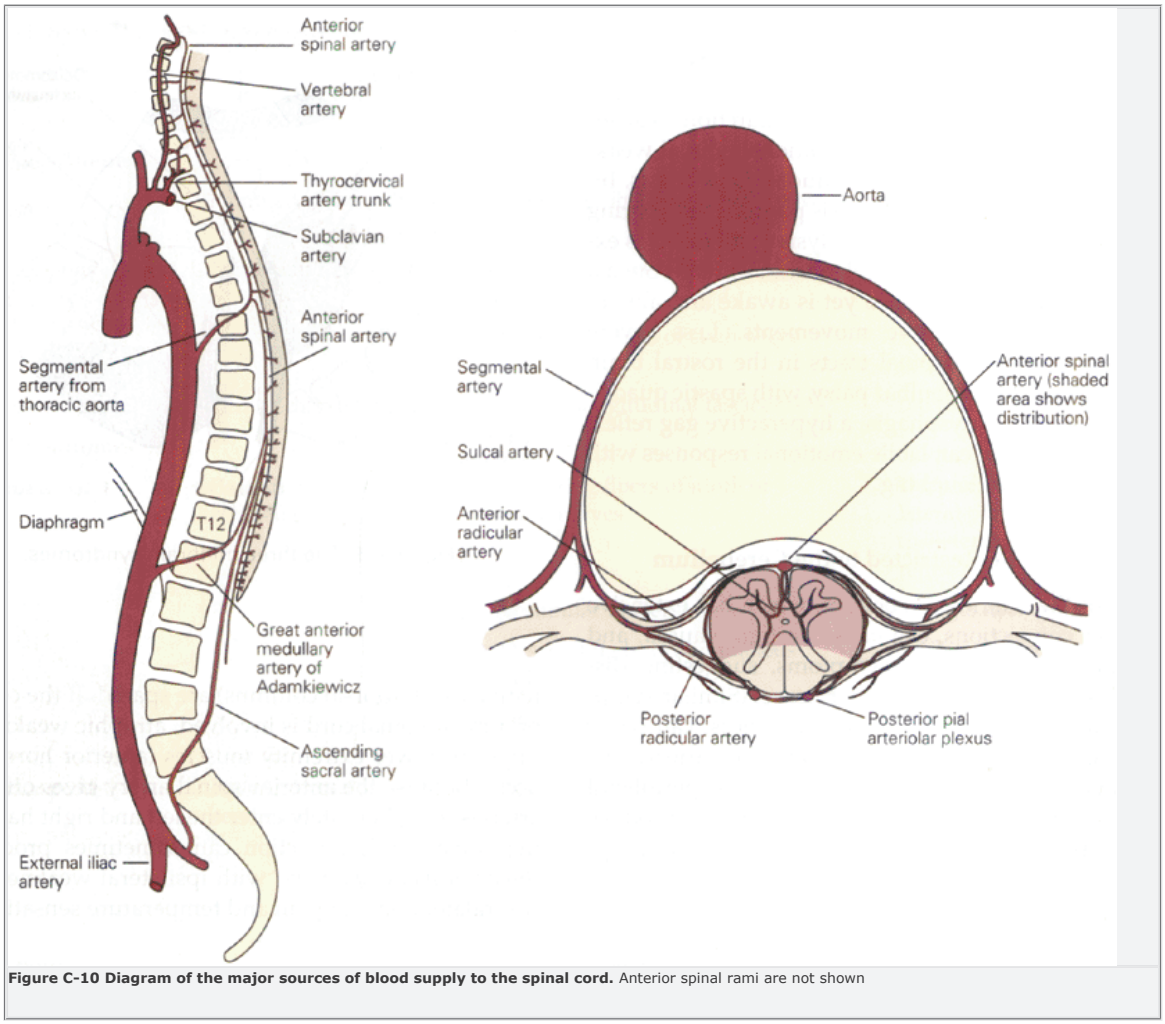


Figure C-10 Diagram of the major sources of blood supply to the spinal cord. Anterior spinal rami are not shown

Cerebrovascular Disease Can Cause Dementia

Cerebral infarction causes dementia by a number of mechanisms, including:

- Infarcts may be critically located. For example, thalamic or inferomedial temporal damage (posterior cerebral artery, usually bilateral) can cause amnesia; hemispheric convexity damage (middle cerebral artery) can cause cognitive or behavioral impairment not explained by disruption of language or spatial discrimination; and bilateral inferomedial frontal lobe damage (anterior cerebral artery) can cause abulia and impaired memory.
- Multiple scattered infarcts, none sufficient to cause significant cognitive loss, can produce additive effects culminating in dementia. In such patients at least 100 cc of brain volume has usually been destroyed.
- Small vessel disease, affecting especially the deep cerebral white matter, can cause either scattered lacunes or more diffuse ischemic lesions. When such lesions are severe enough to cause dementia— so-called *Binswanger disease*—altered behavior, pseudobulbar palsy, pyramidal signs, disturbed gait, andurinary incontinence usually occur.

The Rupture of Microaneurysms Causes Intraparenchymal Stroke

The two most common causes of hemorrhagic stroke, hypertensive intra-axial hemorrhage and rupture of a saccular aneurysm, tend to occur at particular sites and to cause recognizable syndromes. Hypertensive intercerebral hemorrhage is the result of damage to the same small penetrating vessels that, when occluded, cause lacunes; in the case of hemorrhage, however, the damaged vessels develop weakened walls (*Charcot-Bouchard microaneurysms*) that eventually rupture. The most common sites are the putamen, thalamus, pons, internal capsule and corona radiata, and cerebellum. Large diencephalic hemorrhages tend to cause stupor and hemiplegia and have a high mortality rate.

With lesions of the putamen the eyes are usually deviated ipsilaterally (because of disruption of capsular pathways descending from the frontal eye field), whereas with thalamic hemorrhage the eyes tend to be deviated downward and the pupils may not react to light (because of involvement of midbrain pretectal structures essential for upward gaze and pupillary light reactivity). Small hemorrhages may not impair alertness, and, with small thalamic hemorrhages, the motor sensory loss may exceed the weakness. Moreover, computerized tomography has shown that small thalamic hemorrhages may cause aphasia when on the left and hemi-inattention when on the right. Figures C-11 and C-12 show a large putamenal and a small thalamic hemorrhage, respectively.

Pontine hemorrhage, unless quite small, usually causes coma (by disrupting the reticular activating system) and quadriparesis (by transecting the corticospinal tracts). Eye movements, spontaneous or reflex (eg, to ice water in either external auditory canal) are absent, and pupils are pinpoint in size, perhaps in part from transection of descending sympathetic pathways and in part from destruction of reticular inhibitory mechanisms on the Edinger-Westphal nucleus of the midbrain. Pupillary light reactivity, however, is usually preserved, for the pathway subserving this reflex, from the retina to midbrain, is intact. Respirations may be irregular, presumably because of reticular formation involvement. These strokes are nearly always fatal.

Cerebellar hemorrhage, which tends to occur in the region of the dentate nucleus, typically causes a sudden inability to stand or walk (*astasia-abasia*), with ipsilateral limb ataxia. There may be ipsilateral abducens palsy, or horizontal gaze palsy, or facial weakness, presumably from pontine compression. Long-tract motor and sensory signs, however, are usually absent. As swelling increases, further brain stem damage may cause coma, ophthalmoplegia, miosis, and irregular respiration, with fatal outcome.

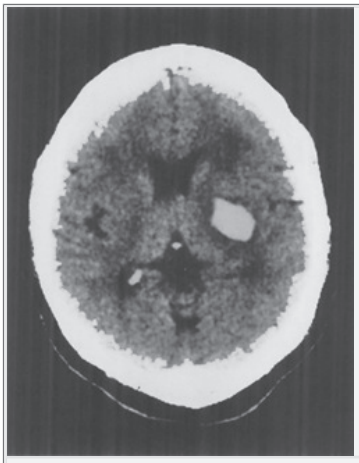


Figure C-11 CT scan showing hemorrhage (white area) in the putamen. (Courtesy of Dr. Allan J. Schwartz.)

The Rupture of Saccular Aneurysms Causes Subarachnoid Hemorrhage

Congenital saccular aneurysms (not to be confused with hypertensive Charcot-Bouchard microaneurysms) are most often found at the junction of the anterior communicating artery with an anterior cerebral artery, at the junction of a posterior communicating artery with an internal carotid artery, and at the first bifurcation of a middle cerebral artery in the sylvian fissure. Each aneurysm, upon rupture, tends to cause not only sudden severe headache but also a characteristic syndrome. By producing a hematoma directly over the oculomotor nerve as it traverses the base of the brain, a ruptured posterior communicating artery aneurysm often causes ipsilateral pupillary dilation with loss of light reactivity. A middle cerebral artery aneurysm may, by either hematoma or secondary infarction, cause a clinical picture resembling that of middle cerebral artery occlusion. After rupture of an anterior communicating artery aneurysm, there may be no focal signs but only decreased alertness or behavioral changes.

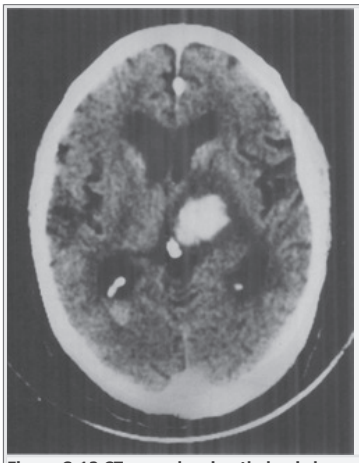


Figure C-12 CT scan showing thalamic hemorrhage. Hematoma is the *white area* surrounded by a darker zone of edema or infarction. (Courtesy of Dr. Allan J. Schwartz.)

Posterior fossa aneurysms most often occur at the rostral bifurcation of the basilar artery or at the origin of the posterior inferior cerebellar artery. They cause a variety of cranial nerve and brain stem signs. Rupture of an aneurysm at any site may cause abrupt coma; the reason is uncertain but may be related to sudden increased intracranial pressure and functional disruption of vital pontomedullary structures.

Stroke Alters the Vascular Physiology of the Brain

After a stroke, cerebral blood flow and cerebrovascular responses to changes in blood pressure or arterial gases are altered. The term *luxury perfusion* refers to the frequent appearance of hyperemia relative to demand after brain infarction. Red venous blood may be seen draining infarcts (reflecting decreased oxygen extraction), and regional cerebral blood flow may or may not be absolutely increased. In addition, there may be vasomotor paralysis with loss of autoregulation to blood pressure changes and then blunted responses to alterations in arterial O_2 or CO_2. This kind of physiological abnormality occurs both within and around ischemic lesions.

In such patients CO_2 (or other cerebral vasodilators) may produce a paradoxical response, increasing cerebral blood flow in brain regions distant from the infarct without affecting the vessels around the lesion. Blood may therefore be shunted from ischemic to normal brain (*intracerebral steal*). In contrast, cerebral vasoconstrictors, by decreasing cerebral blood flow in normal brain without affecting the vessels of ischemic brain, may shunt blood into the area of ischemia or infarction (*inverse intracerebral steal*).

There is controversy about the frequency of these phenomena. Hyperperfusion is not invariable in infarcted brain, and it may coexist with adjacent hypoperfusion with increased oxygen extraction. Similarly, intracerebral steal, while probably most frequent with very large infarcts, is quite unpredictable (particularly in duration) in any single patient. It is also not clear whether increasing cerebral blood flow to infarcted or ischemic areas improves matters by increasing oxygen delivery and the removal of tissue-damaging metabolites or makes matters worse by increasing edema, mass effect, and anastomotic compromise.

Selected Readings

Adams RD, Victor M, Ropper AH. 1996. *Principles of Neurology*, 6th ed., pp. 777–873. New York: McGraw-Hill.

Brust JCM. 1995. Cerebral infarction. In: LP Rowland (ed). *Merritt's Textbook of Neurology*, 9th ed., pp. 246–256. Philadelphia, PA: Lea & Febiger.

Feinberg TE, Farah MJ (eds). 1997. *Behavioral Neurology and Neuropsychology*. NewYork: McGraw-Hill.

Pulsinelli WA. 1996. Cerebrovascular diseases. In: JC Bennett, F Plum (eds). *Cecil Textbook of Medicine*, 20th ed., pp. 2057–2080. Philadelphia, PA: Saunders.

Appendix D

Consciousness and the Neurobiology of the Twenty-First Century