CHAPTER 46

Sensory Receptors, Neuronal Circuits for Processing Information

Input to the nervous system is provided by sensory receptors that detect such sensory stimuli as touch, sound, light, pain, cold, and warmth. The purpose of this chapter is to discuss the basic mechanisms by which these receptors change sensory stimuli into nerve signals that are then conveyed to and processed in the central nervous system.

Types of Sensory Receptors and the Stimuli They Detect

Table 46-1 lists and classifies five basic types of sensory receptors: (1) *mechanoreceptors,* which detect mechanical compression or stretching of the receptor or of tissues adjacent to the receptor; (2) *thermoreceptors,* which detect changes in temperature, with some receptors detecting cold and others warmth; (3) *nociceptors* (pain receptors), which detect damage occurring in the tissues, whether physical damage or chemical damage; (4) *electromagnetic receptors,* which detect light on the retina of the eye; and (5) *chemoreceptors,* which detect taste in the mouth, smell in the nose, oxygen level in the arterial blood, osmolality of the body fluids, carbon dioxide concentration, and other factors that make up the chemistry of the body.

In this chapter, we discuss the function of a few specific types of receptors, primarily peripheral mechanoreceptors, to illustrate some of the principles by which receptors operate. Other receptors are discussed in other chapters in relation to the sensory systems that they subserve. Figure 46-1 shows some of the types of mechanoreceptors found in the skin or in deep tissues of the body.

Differential Sensitivity of Receptors

How do two types of sensory receptors detect different types of sensory stimuli? The answer is, by *"differential sensitivities."* That is, each type of receptor is highly sensitive to one type of stimulus for which it is designed and yet is almost nonresponsive to other types of sensory stimuli. Thus, the rods and cones of the eyes are highly responsive to light but are almost completely nonresponsive to normal ranges of heat, cold, pressure on the eyeballs, or chemical changes in the blood. The osmoreceptors of the supraoptic nuclei in the hypothalamus detect minute changes in the osmolality of the body fluids but have never been known to respond to sound. Finally, pain receptors in the skin are almost never stimulated by usual touch or pressure stimuli but do become highly active the moment tactile stimuli become severe enough to damage the tissues.

Modality of Sensation—The "Labeled Line" Principle

Each of the principal types of sensation that we can experience—pain, touch, sight, sound, and so forth—is called a *modality* of sensation. Yet despite the fact that we experience these different modalities of sensation, nerve fibers transmit only impulses. Therefore, how do different nerve fibers transmit different modalities of sensation?

The answer is that each nerve tract terminates at a specific point in the central nervous system, and the type of sensation felt when a nerve fiber is stimulated is determined by the point in the nervous system to which the fiber leads. For instance, if a pain fiber is stimulated, the person perceives pain regardless of what type of stimulus excites the fiber. The stimulus can be electricity, overheating of the fiber, crushing of the fiber, or stimulation of the pain nerve ending by damage to the tissue cells. In all these instances, the person perceives pain. Likewise, if a touch fiber is stimulated by electrical excitation of a touch receptor or in any other way, the person perceives touch because touch fibers lead to specific touch areas in the brain. Similarly, fibers from the retina of the eye terminate in the vision areas of the brain, fibers from the ear terminate in the auditory areas of the brain, and temperature fibers terminate in the temperature areas.

This specificity of nerve fibers for transmitting only one modality of sensation is called the *labeled line principle.*

Table 46-1 Classification of Sensory Receptors

- **I. Mechanoreceptors**
  - Skin tactile sensibilities (epidermis and dermis)
    - Free nerve endings
    - Expanded tip endings
      - Merkel's discs
      - Plus several other variants
    - Spray endings
    - Ruffini's endings
    - Encapsulated endings
      - Meissner's corpuscles
      - Krause's corpuscles
    - Hair end-organs
  - Deep tissue sensibilities
    - Free nerve endings
    - Expanded tip endings
    - Spray endings
      - Ruffini's endings
    - Encapsulated endings
      - Pacinian corpuscles
      - Plus a few other variants
    - Muscle endings
      - Muscle spindles
      - Golgi tendon receptors
  - Hearing
    - Sound receptors of cochlea
  - Equilibrium
    - Vestibular receptors
  - Arterial pressure
    - Baroreceptors of carotid sinuses and aorta
- **II. Thermoreceptors**
  - Cold
    - Cold receptors
  - Warmth
    - Warm receptors
- **III. Nociceptors**
  - Pain
    - Free nerve endings
- **IV. Electromagnetic receptors**
  - Vision
    - Rods
    - Cones
- **V. Chemoreceptors**
  - Taste
    - Receptors of taste buds
  - Smell
    - Receptors of olfactory epithelium
  - Arterial oxygen
    - Receptors of aortic and carotid bodies
  - Osmolality
    - Neurons in or near supraoptic nuclei
  - Blood CO_2
    - Receptors in or on surface of medulla and in aortic and carotid bodies
  - Blood glucose, amino acids, fatty acids
    - Receptors in hypothalamus

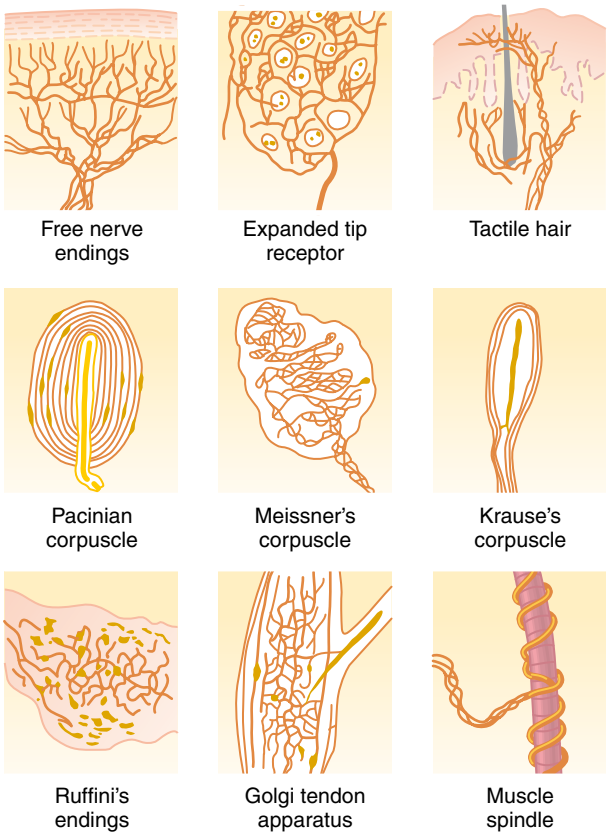


Figure 46-1 Several types of somatic sensory nerve endings.

Transduction of Sensory Stimuli into Nerve Impulses

Local Electrical Currents at Nerve Endings—Receptor Potentials

All sensory receptors have one feature in common. Whatever the type of stimulus that excites the receptor, its immediate effect is to change the membrane electrical potential of the receptor. This change in potential is called a *receptor potential.*

Mechanisms of Receptor Potentials. Different receptors can be excited in one of several ways to cause receptor potentials: (1) by mechanical deformation of the receptor, which stretches the receptor membrane and opens ion channels; (2) by application of a chemical to the membrane, which also opens ion channels; (3) by change of the temperature of the membrane, which alters the permeability of the membrane; or (4) by the effects of electromagnetic radiation, such as light on a retinal visual receptor, which either directly or indirectly changes the receptor membrane characteristics and allows ions to flow through membrane channels.

These four means of exciting receptors correspond in general with the different types of known sensory receptors. In all instances, the basic cause of the change in membrane potential is a change in membrane permeability of the receptor, which allows ions to diffuse more or

less readily through the membrane and thereby to change the *transmembrane potential.*

Maximum Receptor Potential Amplitude. The maximum amplitude of most sensory receptor potentials is about 100 millivolts, but this level occurs only at an extremely high intensity of sensory stimulus. This is about the same maximum voltage recorded in action potentials and is also the change in voltage when the membrane becomes maximally permeable to sodium ions.

Relation of the Receptor Potential to Action Potentials. When the receptor potential rises above the *threshold* for eliciting action potentials in the nerve fiber attached to the receptor, then action potentials occur, as illustrated in Figure 46-2. Note also that the more the receptor potential rises above the threshold level, the greater becomes the *action potential frequency.*

Receptor Potential of the Pacinian Corpuscle—An Example of Receptor Function

The student should at this point restudy the anatomical structure of the pacinian corpuscle shown in Figure 46-1. Note that the corpuscle has a central nerve fiber extending through its core. Surrounding this are multiple concentric capsule layers, so compression anywhere on the outside of the corpuscle will elongate, indent, or otherwise deform the central fiber.

Now study Figure 46-3, which shows only the central fiber of the pacinian corpuscle after all capsule layers but one have been removed. The tip of the central fiber inside the capsule is unmyelinated, but the fiber does become myelinated (the blue sheath shown in the figure) shortly before leaving the corpuscle to enter a peripheral sensory nerve.

The figure also shows the mechanism by which a receptor potential is produced in the pacinian corpuscle. Observe the small area of the terminal fiber that has been deformed by compression of the corpuscle, and note that ion channels have opened in the membrane, allowing

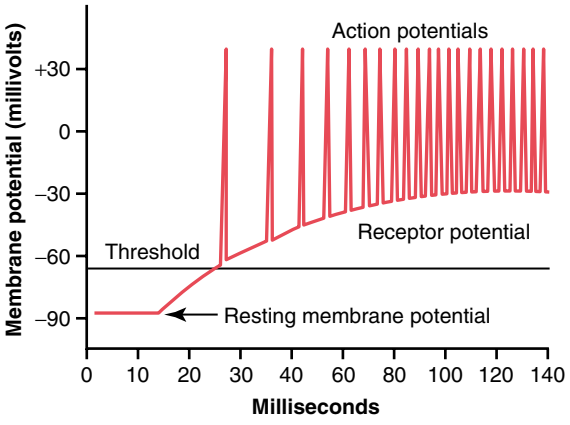


Figure 46-2 Typical relation between receptor potential and action potentials when the receptor potential rises above threshold level.

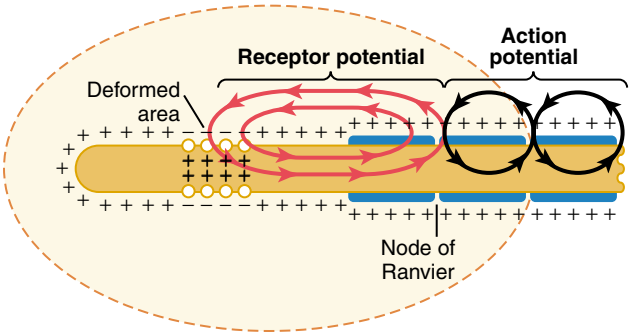


Figure 46-3 Excitation of a sensory nerve fiber by a receptor potential produced in a pacinian corpuscle. (Modified from Loëwenstein WR: Excitation and inactivation in a receptor membrane. Ann N Y Acad Sci 94:510, 1961.)

positively charged sodium ions to diffuse to the interior of the fiber. This creates increased positivity inside the fiber, which is the "receptor potential." The receptor potential in turn induces a *local circuit* of current flow, shown by the arrows, that spreads along the nerve fiber. At the first node of Ranvier, which itself lies inside the capsule of the pacinian corpuscle, the local current flow depolarizes the fiber membrane at this node, which then sets off typical action potentials that are transmitted along the nerve fiber toward the central nervous system.

Relation Between Stimulus Intensity and the Receptor Potential. Figure 46-4 shows the changing amplitude of the receptor potential caused by progressively stronger mechanical compression (increasing "stimulus strength") applied experimentally to the central core of a pacinian corpuscle. Note that the amplitude increases rapidly at first but then progressively less rapidly at high stimulus strength.

In turn, the *frequency of repetitive action potentials* transmitted from sensory receptors increases approximately in

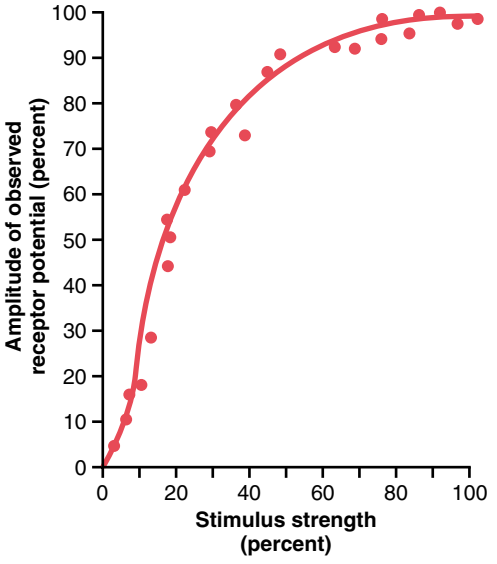


Figure 46-4 Relation of amplitude of receptor potential to strength of a mechanical stimulus applied to a pacinian corpuscle. (Data from Loëwenstein WR: Excitation and inactivation in a receptor membrane. Ann N Y Acad Sci 94:510, 1961.)

proportion to the increase in receptor potential. Putting this principle together with the data in Figure 46-4, one can see that very intense stimulation of the receptor causes progressively less and less additional increase in numbers of action potentials. This is an exceedingly important principle that is applicable to almost all sensory receptors. It allows the receptor to be sensitive to very weak sensory experience and yet not reach a maximum firing rate until the sensory experience is extreme. This allows the receptor to have an extreme range of response, from very weak to very intense.

Adaptation of Receptors

Another characteristic of all sensory receptors is that they *adapt* either partially or completely to any constant stimulus after a period of time. That is, when a continuous sensory stimulus is applied, the receptor responds at a high impulse rate at first and then at a progressively slower rate until finally the rate of action potentials decreases to very few or often to none at all.

Figure 46-5 shows typical adaptation of certain types of receptors. Note that the pacinian corpuscle adapts very rapidly, hair receptors adapt within a second or so, and some joint capsule and muscle spindle receptors adapt slowly.

Furthermore, some sensory receptors adapt to a far greater extent than others. For example, the pacinian corpuscles adapt to "extinction" within a few hundredths of a second, and the receptors at the bases of the hairs adapt to extinction within a second or more. It is probable that all other *mechanoreceptors* eventually adapt almost completely, but some require hours or days to do so, for which reason they are called "nonadapting" receptors. The longest measured time for almost complete adaptation of a mechanoreceptor is about 2 days, which is the adaptation time for many carotid and aortic baroreceptors. Conversely, some of the nonmechanoreceptors—the chemoreceptors and pain receptors, for instance—probably never adapt completely.

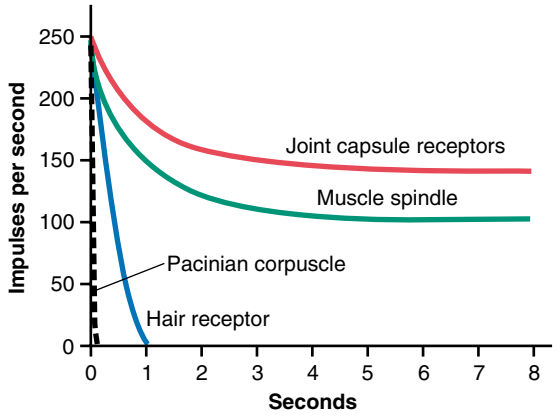


Figure 46-5 Adaptation of different types of receptors, showing rapid adaptation of some receptors and slow adaptation of others.

Mechanisms by Which Receptors Adapt. The mechanism of receptor adaptation is different for each type of receptor, in much the same way that development of a receptor potential is an individual property. For instance, in the eye, the rods and cones adapt by changing the concentrations of their light-sensitive chemicals (which is discussed in Chapter 50).

In the case of the mechanoreceptors, the receptor that has been studied in greatest detail is the pacinian corpuscle. Adaptation occurs in this receptor in two ways. First, the pacinian corpuscle is a viscoelastic structure, so that when a distorting force is suddenly applied to one side of the corpuscle, this force is instantly transmitted by the viscous component of the corpuscle directly to the same side of the central nerve fiber, thus eliciting a receptor potential. However, within a few hundredths of a second, the fluid within the corpuscle redistributes and the receptor potential is no longer elicited. Thus, the receptor potential appears at the onset of compression but disappears within a small fraction of a second even though the compression continues.

The second mechanism of adaptation of the pacinian corpuscle, but a much slower one, results from a process called *accommodation,* which occurs in the nerve fiber itself. That is, even if by chance the central core fiber should continue to be distorted, the tip of the nerve fiber itself gradually becomes "accommodated" to the stimulus. This probably results from progressive "inactivation" of the sodium channels in the nerve fiber membrane, which means that sodium current flow through the channels causes them gradually to close, an effect that seems to occur for all or most cell membrane sodium channels, as was explained in Chapter 5.

Presumably, these same two general mechanisms of adaptation apply also to the other types of mechanoreceptors. That is, part of the adaptation results from readjustments in the structure of the receptor itself, and part from an electrical type of accommodation in the terminal nerve fibril.

Slowly Adapting Receptors Detect Continuous Stimulus Strength—The "Tonic" Receptors. Slowly adapting receptors continue to transmit impulses to the brain as long as the stimulus is present (or at least for many minutes or hours). Therefore, they keep the brain constantly apprised of the status of the body and its relation to its surroundings. For instance, impulses from the muscle spindles and Golgi tendon apparatuses allow the nervous system to know the status of muscle contraction and load on the muscle tendon at each instant.

Other slowly adapting receptors include (1) receptors of the macula in the vestibular apparatus, (2) pain receptors, (3) baroreceptors of the arterial tree, and (4) chemoreceptors of the carotid and aortic bodies.

Because the slowly adapting receptors can continue to transmit information for many hours, they are called *tonic* receptors.

Rapidly Adapting Receptors Detect Change in Stimulus Strength—The "Rate Receptors," "Movement Receptors," or "Phasic Receptors." Receptors that adapt rapidly cannot be used to transmit a continuous signal because these receptors are stimulated only when the stimulus strength changes. Yet they react strongly *while a change is actually taking place*. Therefore, these receptors are called *rate* receptors, *movement* receptors, or *phasic* receptors. Thus, in the case of the pacinian corpuscle, sudden pressure applied to the tissue excites this receptor for a few milliseconds, and then its excitation is over even though the pressure continues. But later, it transmits a signal again when the pressure is released. In other words, the pacinian corpuscle is exceedingly important in apprising the nervous system of rapid tissue deformations, but it is useless for transmitting information about constant conditions in the body.

Importance of the Rate Receptors—Their Predictive Function. If one knows the rate at which some change in bodily status is taking place, one can predict in one's mind the state of the body a few seconds or even a few minutes later. For instance, the receptors of the semicircular canals in the vestibular apparatus of the ear detect the rate at which the head begins to turn when one runs around a curve. Using this information, a person can predict how much he or she will turn within the next 2 seconds and can adjust the motion of the legs *ahead of time* to keep from losing balance. Likewise, receptors located in or near the joints help detect the rates of movement of the different parts of the body. For instance, when one is running, information from the joint rate receptors allows the nervous system to predict where the feet will be during any precise fraction of the next second. Therefore, appropriate motor signals can be transmitted to the muscles of the legs to make any necessary anticipatory corrections in position so that the person will not fall. Loss of this predictive function makes it impossible for the person to run.

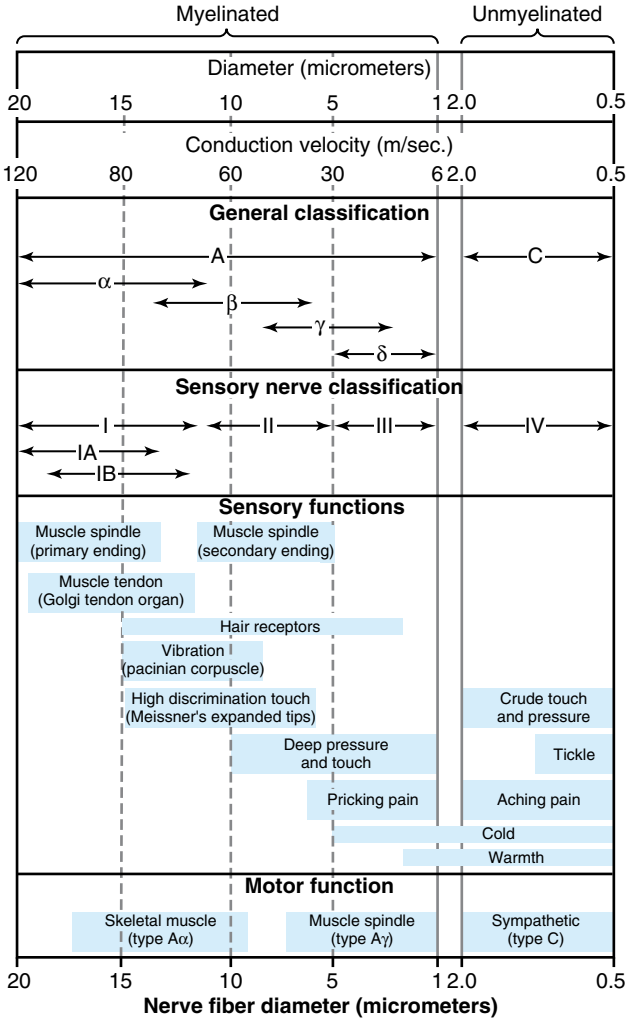


Figure 46-6 Physiologic classifications and functions of nerve fibers.

Nerve Fibers That Transmit Different Types of Signals and Their Physiologic Classification

Some signals need to be transmitted to or from the central nervous system extremely rapidly; otherwise, the information would be useless. An example of this is the sensory signals that apprise the brain of the momentary positions of the legs at each fraction of a second during running. At the other extreme, some types of sensory information, such as that depicting prolonged, aching pain, do not need to be transmitted rapidly, so slowly conducting fibers will suffice. As shown in Figure 46-6, nerve fibers come in all sizes between 0.5 and 20 micrometers in diameter—the larger the diameter, the greater the conducting velocity. The range of conducting velocities is between 0.5 and 120 m/sec.

General Classification of Nerve Fibers. Shown in Figure 46-6 is a "general classification" and a "sensory nerve classification" of the different types of nerve fibers. In the general classification, the fibers are divided into types A and C, and the type A fibers are further subdivided into α, β, γ, and δ fibers.

Type A fibers are the typical large and medium-sized *myelinated* fibers of spinal nerves. Type C fibers are the small *unmyelinated* nerve fibers that conduct impulses at low velocities. The C fibers constitute more than one half of the sensory fibers in most peripheral nerves, as well as all the postganglionic autonomic fibers.

The sizes, velocities of conduction, and functions of the different nerve fiber types are also given in Figure 46-6. Note that a few large myelinated fibers can transmit impulses at velocities as great as 120 m/sec, a distance in 1 second that is longer than a football field. Conversely, the smallest fibers transmit impulses as slowly as 0.5 m/sec, requiring about 2 seconds to go from the big toe to the spinal cord.

Alternative Classification Used by Sensory Physiologists. Certain recording techniques have made it possible to separate the type Aα fibers into two subgroups; yet these same recording techniques cannot distinguish easily between Aβ and Aγ fibers. Therefore, the following classification is frequently used by sensory physiologists:

Group Ia
Fibers from the annulospiral endings of muscle spindles (average about 17 microns in diameter; these are α-type A fibers in the general classification).

Group Ib
Fibers from the Golgi tendon organs (average about 16 micrometers in diameter; these also are α-type A fibers).

Group II
Fibers from most discrete cutaneous tactile receptors and from the flower-spray endings of the muscle spindles (average about 8 micrometers in diameter; these are β- and γ-type A fibers in the general classification).

Group III
Fibers carrying temperature, crude touch, and pricking pain sensations (average about 3 micrometers in diameter; they are δ-type A fibers in the general classification).

Group IV
Unmyelinated fibers carrying pain, itch, temperature, and crude touch sensations (0.5 to 2 micrometers in diameter; they are type C fibers in the general classification).

Transmission of Signals of Different Intensity in Nerve Tracts—Spatial and Temporal Summation

One of the characteristics of each signal that always must be conveyed is signal intensity—for instance, the intensity of pain. The different gradations of intensity can be transmitted either by using increasing numbers of parallel fibers or by sending more action potentials along a single fiber. These two mechanisms are called, respectively, *spatial summation* and *temporal summation*.

Spatial Summation. Figure 46-7 shows the phenomenon of *spatial summation,* whereby increasing signal strength is transmitted by using progressively greater numbers of fibers. This figure shows a section of skin innervated by a large number of parallel pain fibers. Each of these arborizes into hundreds of minute *free nerve endings* that serve as pain receptors. The entire cluster of fibers from one pain fiber frequently covers an area of skin as large as 5 centimeters in diameter. This area is called the *receptor field* of that fiber. The number of endings is large in the center of the field but diminishes toward the periphery. One can also see from the figure that the arborizing fibrils overlap those from other pain fibers. Therefore, a pinprick of the skin usually stimulates endings from many different pain fibers simultaneously. When the pinprick is in the center of the receptive field of a particular pain fiber, the degree of stimulation of that fiber is far greater than when it is in the periphery of the field because the number of free nerve endings in the middle of the field is much greater than at the periphery.

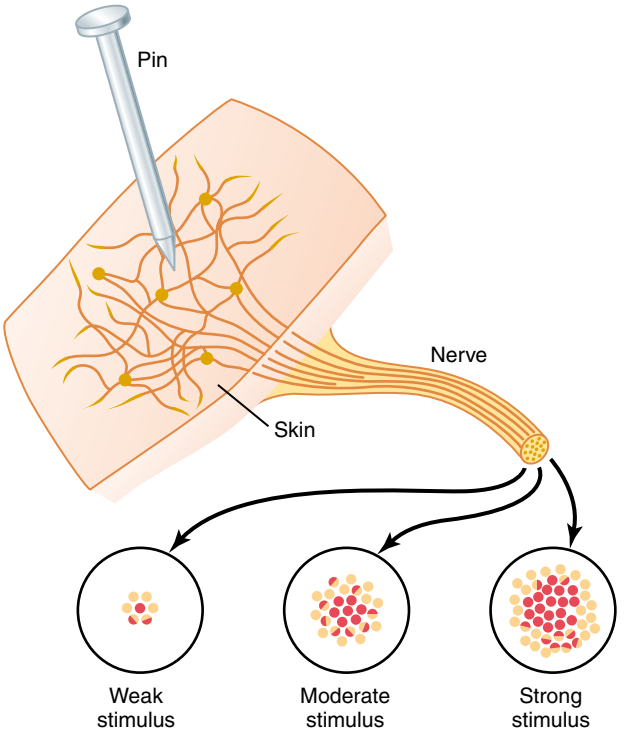


Figure 46-7 Pattern of stimulation of pain fibers in a nerve leading from an area of skin pricked by a pin. This is an example of *spatial summation*.

Thus, the lower part of Figure 46-7 shows three views of the cross section of the nerve bundle leading from the skin area. To the left is the effect of a weak stimulus, with only a single nerve fiber in the middle of the bundle stimulated strongly (represented by the red-colored fiber), whereas several adjacent fibers are stimulated weakly (half-red fibers). The other two views of the nerve cross section show the effect of a moderate stimulus and a strong stimulus, with progressively more fibers being stimulated. Thus, the stronger signals spread to more and more fibers. This is the phenomenon of *spatial summation*.

Temporal Summation. A second means for transmitting signals of increasing strength is by increasing the *frequency* of nerve impulses in each fiber, which is called *temporal summation*. Figure 46-8 demonstrates this, showing in the upper part a changing strength of signal and in the lower part the actual impulses transmitted by the nerve fiber.

Transmission and Processing of Signals in Neuronal Pools

The central nervous system is composed of thousands to millions of neuronal pools; some of these contain few neurons, whereas others have vast numbers. For instance, the entire cerebral cortex could be considered to be a single large neuronal pool. Other neuronal pools include the different basal ganglia and the specific nuclei in the thalamus,

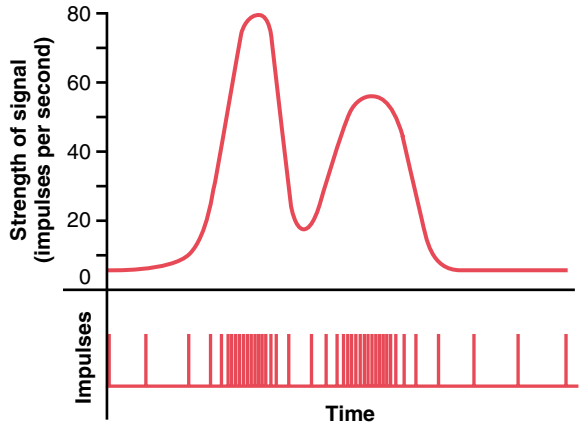


Figure 46-8 Translation of signal strength into a frequency-modulated series of nerve impulses, showing the strength of signal (*above*) and the separate nerve impulses (*below*). This is an example of *temporal summation*.

cerebellum, mesencephalon, pons, and medulla. Also, the entire dorsal gray matter of the spinal cord could be considered one long pool of neurons.

Each neuronal pool has its own special organization that causes it to process signals in its own unique way, thus allowing the total consortium of pools to achieve the multitude of functions of the nervous system. Yet despite their differences in function, the pools also have many similar principles of function, described in the following pages.

Relaying of Signals Through Neuronal Pools

Organization of Neurons for Relaying Signals. Figure 46-9 is a schematic diagram of several neurons in a neuronal pool, showing "input" fibers to the left and "output" fibers to the right. Each input fiber divides hundreds to thousands of times, providing a thousand or more terminal fibrils that spread into a large area in the pool to synapse with dendrites or cell bodies of the neurons in the pool. The dendrites usually also arborize and spread hundreds to thousands of micrometers in the pool.

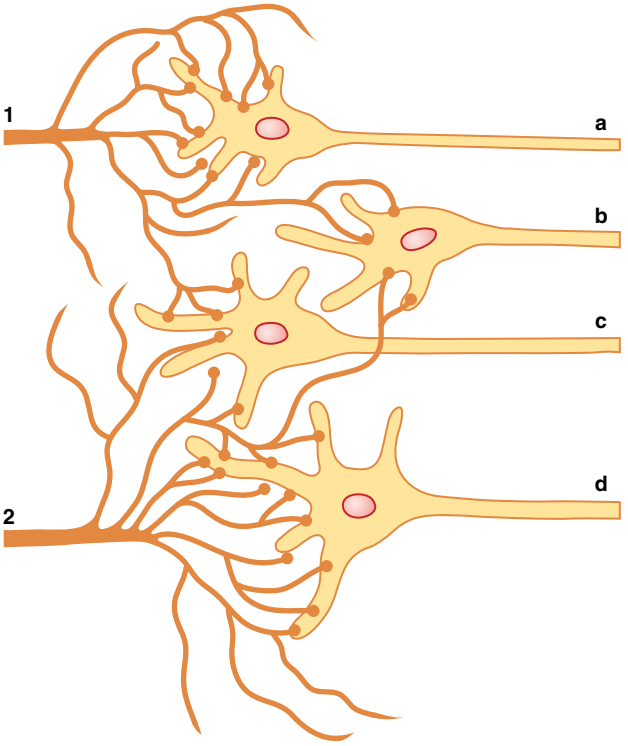


Figure 46-9 Basic organization of a neuronal pool.

The neuronal area stimulated by each incoming nerve fiber is called its *stimulatory field.* Note in Figure 46-9 that large numbers of the terminals from each input fiber lie on the nearest neuron in its "field," but progressively fewer terminals lie on the neurons farther away.

Threshold and Subthreshold Stimuli—Excitation or Facilitation. From the discussion of synaptic function in Chapter 45, it will be recalled that discharge of a single excitatory presynaptic terminal almost never causes an action potential in a postsynaptic neuron. Instead, large numbers of input terminals must discharge on the same neuron either simultaneously or in rapid succession to cause excitation. For instance, in Figure 46-9, let us assume that six terminals must discharge almost simultaneously to excite any one of the neurons. If the student counts the number of terminals on each one of the neurons from each input fiber, he or she will see that *input fiber 1* has more than enough terminals to cause *neuron a* to discharge. The stimulus from input fiber 1 to this neuron is said to be an *excitatory stimulus;* it is also called a *suprathreshold stimulus* because it is above the threshold required for excitation.

Input fiber 1 also contributes terminals to neurons b and c, but not enough to cause excitation. Nevertheless, discharge of these terminals makes both these neurons more likely to be excited by signals arriving through other incoming nerve fibers. Therefore, the stimuli to these neurons are said to be *subthreshold,* and the neurons are said to be *facilitated.*

Similarly, for *input fiber 2,* the stimulus to *neuron d* is a suprathreshold stimulus, and the stimuli to *neurons b* and *c* are subthreshold, but facilitating, stimuli.

Figure 46-9 represents a highly condensed version of a neuronal pool because each input nerve fiber usually provides massive numbers of branching terminals to hundreds or thousands of neurons in its distribution "field," as shown in Figure 46-10. In the central portion of the field in this figure, designated by the circled area, all the neurons are stimulated by the incoming fiber. Therefore, this is said to be the *discharge zone* of the incoming fiber, also called the *excited zone* or *liminal zone.* To each side,

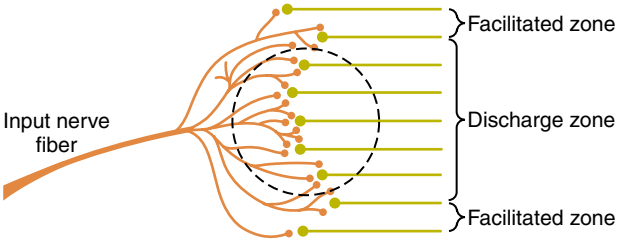


Figure 46-10 "Discharge" and "facilitated" zones of a neuronal pool.

the neurons are facilitated but not excited, and these areas are called the *facilitated zone,* also called the *subthreshold zone* or *subliminal zone.*

Inhibition of a Neuronal Pool. We must also remember that some incoming fibers inhibit neurons, rather than exciting them. This is the opposite of facilitation, and the entire field of the inhibitory branches is called the *inhibitory zone.* The degree of inhibition in the center of this zone is great because of large numbers of endings in the center; it becomes progressively less toward its edges.

Divergence of Signals Passing Through Neuronal Pools

Often it is important for weak signals entering a neuronal pool to excite far greater numbers of nerve fibers leaving the pool. This phenomenon is called *divergence.* Two major types of divergence occur and have entirely different purposes.

An *amplifying* type of divergence is shown in Figure 46-11*A.* This means simply that an input signal spreads to an increasing number of neurons as it passes through successive orders of neurons in its path. This type of divergence is characteristic of the corticospinal pathway in its control of skeletal muscles, with a single large pyramidal cell in the motor cortex capable, under highly facilitated conditions, of exciting as many as 10,000 muscle fibers.

The second type of divergence, shown in Figure 46-11*B,* is *divergence into multiple tracts.* In this case, the signal is transmitted in two directions from the pool. For instance, information transmitted up the dorsal columns of the spinal cord takes two courses in the lower part of the brain: (1) into the cerebellum and (2) on through the lower regions of the brain to the thalamus and cerebral cortex. Likewise, in the thalamus, almost all sensory information is relayed both into still deeper structures of the thalamus and at the same time to discrete regions of the cerebral cortex.

Convergence of Signals

Convergence means signals from multiple inputs uniting to excite a single neuron. Figure 46-12*A* shows *convergence from a single source.* That is, multiple terminals from a single incoming fiber tract terminate on the same neuron. The importance of this is that neurons are almost never excited by an action potential from a single input terminal. But action potentials converging on the neuron from multiple terminals provide enough spatial summation to bring the neuron to the threshold required for discharge.

Convergence can also result from input signals (excitatory or inhibitory) *from multiple sources,* as shown in Figure 46-12*B.* For instance, the interneurons of the spinal cord receive converging signals from (1) peripheral nerve fibers entering the cord, (2) propriospinal fibers passing from one segment of the cord to another, (3) corticospinal fibers from the cerebral cortex, and (4) several other long pathways descending from the brain into the spinal cord. Then the signals from the interneurons converge on the anterior motor neurons to control muscle function.

Such convergence allows *summation* of information from different sources, and the resulting response is a summated effect of all the different types of information. Convergence is one of the important means by which the central nervous system correlates, summates, and sorts different types of information.

Neuronal Circuit with both Excitatory and Inhibitory Output Signals

Sometimes an incoming signal to a neuronal pool causes an output excitatory signal going in one direction and at the same time an inhibitory signal going elsewhere. For instance, at the same time that an excitatory signal is transmitted by one set of neurons in the spinal cord to cause forward movement of a leg, an inhibitory signal is transmitted through a separate set of neurons to inhibit the muscles on the back of the leg so that they will not oppose the forward movement. This type of circuit is characteristic for controlling all antagonistic pairs of muscles, and it is called the *reciprocal inhibition circuit.*

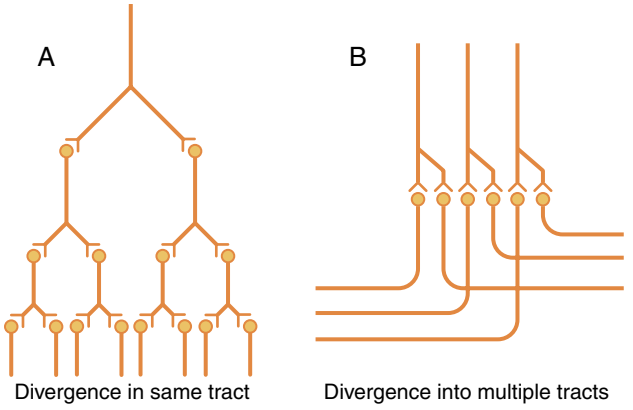


Figure 46-11 "Divergence" in neuronal pathways. *A,* Divergence within a pathway to cause "amplification" of the signal. *B,* Divergence into multiple tracts to transmit the signal to separate areas.

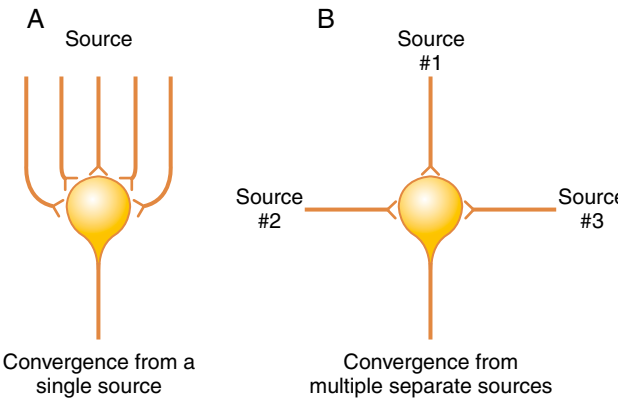


Figure 46-12 "Convergence" of multiple input fibers onto a single neuron. *A,* Multiple input fibers from a single source. *B,* Input fibers from multiple separate sources.

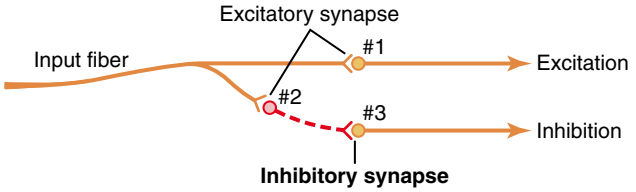


Figure 46-13 Inhibitory circuit. Neuron 2 is an inhibitory neuron.

Figure 46-13 shows the means by which the inhibition is achieved. The input fiber directly excites the excitatory output pathway, but it stimulates an intermediate *inhibitory neuron* (neuron 2), which secretes a different type of transmitter substance to inhibit the second output pathway from the pool. This type of circuit is also important in preventing overactivity in many parts of the brain.

Prolongation of a Signal by a Neuronal Pool—"Afterdischarge"

Thus far, we have considered signals that are merely relayed through neuronal pools. However, in many instances, a signal entering a pool causes a prolonged output discharge, called *afterdischarge,* lasting a few milliseconds to as long as many minutes after the incoming signal is over. The most important mechanisms by which afterdischarge occurs are the following.

Synaptic Afterdischarge. When excitatory synapses discharge on the surfaces of dendrites or soma of a neuron, a postsynaptic electrical potential develops in the neuron and lasts for many milliseconds, especially when some of the long-acting synaptic transmitter substances are involved. As long as this potential lasts, it can continue to excite the neuron, causing it to transmit a continuous train of output impulses, as was explained in Chapter 45. Thus, as a result of this synaptic "afterdischarge" mechanism alone, it is possible for a single instantaneous input signal to cause a sustained signal output (a series of repetitive discharges) lasting for many milliseconds.

Reverberatory (Oscillatory) Circuit as a Cause of Signal Prolongation. One of the most important of all circuits in the entire nervous system is the *reverberatory,* or *oscillatory, circuit.* Such circuits are caused by positive feedback within the neuronal circuit that feeds back to re-excite the input of the same circuit. Consequently, once stimulated, the circuit may discharge repetitively for a long time.

Several possible varieties of reverberatory circuits are shown in Figure 46-14. The simplest, shown in Figure 46-14*A,* involves only a single neuron. In this case, the output neuron simply sends a collateral nerve fiber back to its own dendrites or soma to restimulate itself. Although this type of circuit probably is not an important one, theoretically, once the neuron discharges, the feedback stimuli could keep the neuron discharging for a protracted time thereafter.

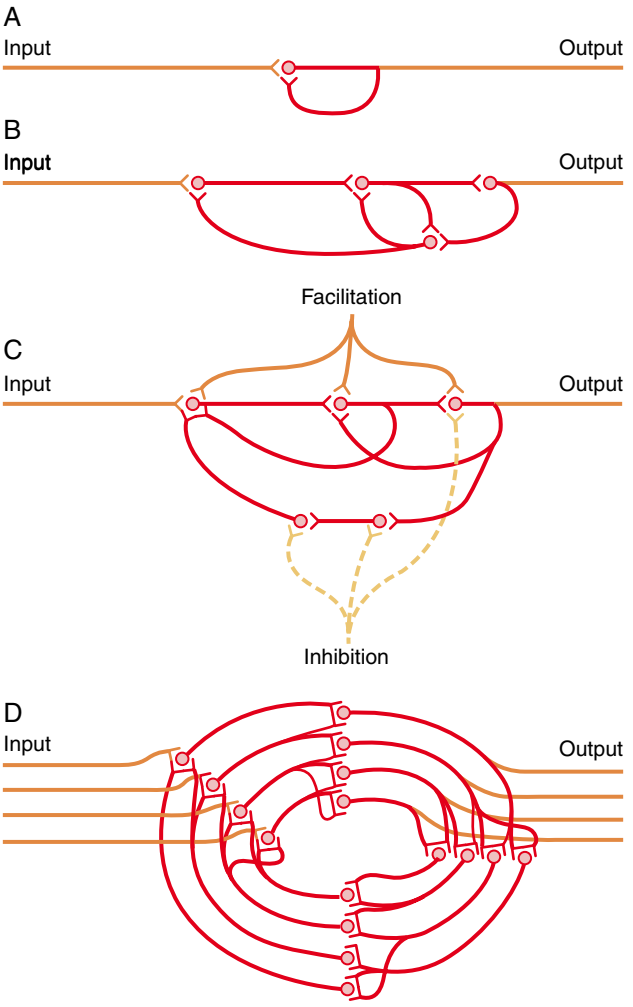


Figure 46-14 Reverberatory circuits of increasing complexity.

Figure 46-14*B* shows a few additional neurons in the feedback circuit, which causes a longer delay between initial discharge and the feedback signal. Figure 46-14*C* shows a still more complex system in which both facilitatory and inhibitory fibers impinge on the reverberating circuit. A facilitatory signal enhances the intensity and frequency of reverberation, whereas an inhibitory signal depresses or stops the reverberation.

Figure 46-14*D* shows that most reverberating pathways are constituted of many parallel fibers. At each cell station, the terminal fibrils spread widely. In such a system, the total reverberating signal can be either weak or strong, depending on how many parallel nerve fibers are momentarily involved in the reverberation.

Characteristics of Signal Prolongation from a Reverberatory Circuit. Figure 46-15 shows output signals from a typical reverberatory circuit. The input stimulus may last only 1 millisecond or so, and yet the output can last for many milliseconds or even minutes. The figure demonstrates that the intensity of the output signal usually increases to a high value early in reverberation and then decreases to a critical point, at which it suddenly ceases entirely. The cause of this sudden cessation of

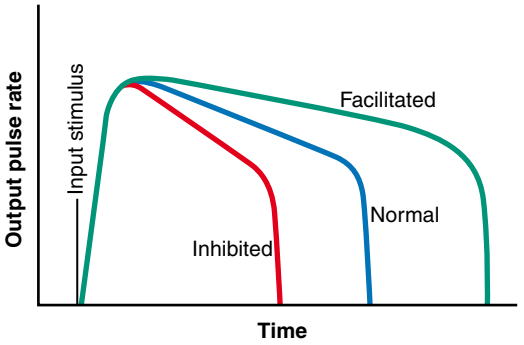


Figure 46-15 Typical pattern of the output signal from a reverberatory circuit after a single input stimulus, showing the effects of facilitation and inhibition.

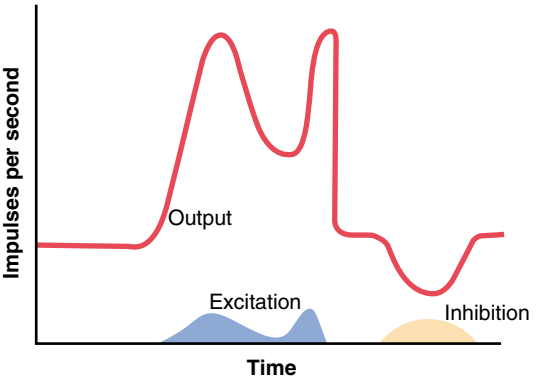


Figure 46-16 Continuous output from either a reverberating circuit or a pool of intrinsically discharging neurons. This figure also shows the effect of excitatory or inhibitory input signals.

reverberation is fatigue of synaptic junctions in the circuit. Fatigue beyond a certain critical level lowers the stimulation of the next neuron in the circuit below threshold level so that the circuit feedback is suddenly broken.

The duration of the total signal before cessation can also be controlled by signals from other parts of the brain that inhibit or facilitate the circuit. Almost these exact patterns of output signals are recorded from the motor nerves exciting a muscle involved in a flexor reflex after pain stimulation of the foot (as shown later in Figure 46-18).

Continuous Signal Output from Some Neuronal Circuits

Some neuronal circuits emit output signals continuously, even without excitatory input signals. At least two mechanisms can cause this effect: (1) continuous intrinsic neuronal discharge and (2) continuous reverberatory signals.

Continuous Discharge Caused by Intrinsic Neuronal Excitability. Neurons, like other excitable tissues, discharge repetitively if their level of excitatory membrane potential rises above a certain threshold level. The membrane potentials of many neurons even normally are high enough to cause them to emit impulses continually. This occurs especially in many of the neurons of the cerebellum, as well as in most of the interneurons of the spinal cord. The rates at which these cells emit impulses can be increased by excitatory signals or decreased by inhibitory signals; inhibitory signals often can decrease the rate of firing to zero.

Continuous Signals Emitted from Reverberating Circuits as a Means for Transmitting Information. A reverberating circuit that does not fatigue enough to stop reverberation is a source of continuous impulses. And excitatory impulses entering the reverberating pool can increase the output signal, whereas inhibition can decrease or even extinguish the signal.

Figure 46-16 shows a continuous output signal from a pool of neurons. The pool may be emitting impulses because of intrinsic neuronal excitability or as a result of reverberation. Note that an excitatory input signal greatly increases the output signal, whereas an inhibitory input signal greatly decreases the output. Those students who are familiar with radio transmitters will recognize this to be a *carrier wave* type of information transmission. That is, the excitatory and inhibitory control signals are not the *cause* of the output signal, but they do *control* its changing level of intensity. Note that this carrier wave system allows a *decrease* in signal intensity, as well as an increase, whereas up to this point, the types of information transmission we have discussed have been mainly positive information rather than negative information. This type of information transmission is used by the autonomic nervous system to control such functions as vascular tone, gut tone, degree of constriction of the iris in the eye, and heart rate. That is, the nerve excitatory signal to each of these can be either increased or decreased by accessory input signals into the reverberating neuronal pathway.

Rhythmical Signal Output

Many neuronal circuits emit rhythmical output signals—for instance, a rhythmical respiratory signal originates in the respiratory centers of the medulla and pons. This respiratory rhythmical signal continues throughout life. Other rhythmical signals, such as those that cause scratching movements by the hind leg of a dog or the walking movements of any animal, require input stimuli into the respective circuits to initiate the rhythmical signals.

All or almost all rhythmical signals that have been studied experimentally have been found to result from reverberating circuits or a succession of sequential reverberating circuits that feed excitatory or inhibitory signals in a circular pathway from one neuronal pool to the next.

Excitatory or inhibitory signals can also increase or decrease the amplitude of the rhythmical signal output. Figure 46-17, for instance, shows changes in the respiratory signal output in the phrenic nerve. When the carotid body is stimulated by arterial oxygen deficiency, both the frequency and the amplitude of the respiratory rhythmical output signal increase progressively.

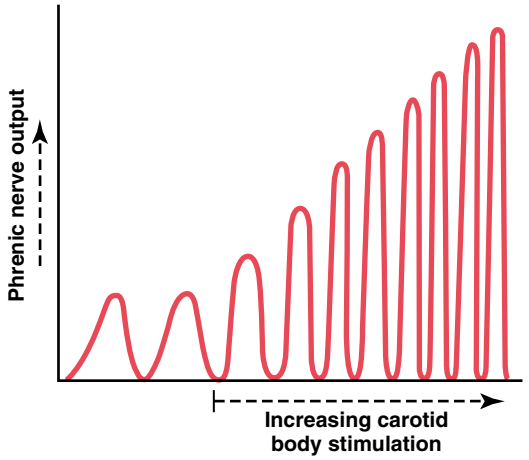


Figure 46-17 The rhythmical output of summated nerve impulses from the respiratory center, showing that progressively increasing stimulation of the carotid body increases both the intensity and the frequency of the phrenic nerve signal to the diaphragm to increase respiration.

Instability and Stability of Neuronal Circuits

Almost every part of the brain connects either directly or indirectly with every other part, and this creates a serious problem. If the first part excites the second, the second the third, the third the fourth, and so on until finally the signal re-excites the first part, it is clear that an excitatory signal entering any part of the brain would set off a continuous cycle of re-excitation of all parts. If this should occur, the brain would be inundated by a mass of uncontrolled reverberating signals—signals that would be transmitting no information but, nevertheless, would be consuming the circuits of the brain so that none of the informational signals could be transmitted. Such an effect occurs in widespread areas of the brain during *epileptic seizures.* How does the central nervous system prevent this from happening all the time? The answer lies mainly in two basic mechanisms that function throughout the central nervous system: (1) inhibitory circuits and (2) fatigue of synapses.

Inhibitory Circuits as a Mechanism for Stabilizing Nervous System Function

Two types of inhibitory circuits in widespread areas of the brain help prevent excessive spread of signals: (1) inhibitory feedback circuits that return from the termini of pathways back to the initial excitatory neurons of the same pathways—these circuits occur in virtually all sensory nervous pathways and inhibit either the input neurons or the intermediate neurons in the sensory pathway when the termini become overly excited; and (2) some neuronal pools that exert gross inhibitory control over widespread areas of the brain—for instance, many of the basal ganglia exert inhibitory influences throughout the muscle control system.

Synaptic Fatigue as a Means of Stabilizing the Nervous System

Synaptic fatigue means simply that synaptic transmission becomes progressively weaker the more prolonged and more intense the period of excitation. Figure 46-18 shows three successive records of a flexor reflex elicited in an animal caused by inflicting pain in the footpad of the paw. Note in each record that the strength of contraction progressively "decrements"—that is, its strength diminishes; much of this effect is caused by *fatigue* of synapses in the flexor reflex circuit. Furthermore, the shorter the interval between successive flexor reflexes, the less the intensity of the subsequent reflex response.

Automatic Short-Term Adjustment of Pathway Sensitivity by the Fatigue Mechanism. Now let us apply this phenomenon of fatigue to other pathways in the brain. Those that are overused usually become fatigued, so their sensitivities decrease. Conversely, those that are underused become rested and their sensitivities increase. Thus, fatigue and recovery from fatigue constitute an important short-term means of moderating the sensitivities of the different nervous system circuits. These help to keep the circuits operating in a range of sensitivity that allows effective function.

Long-Term Changes in Synaptic Sensitivity Caused by Automatic Down-regulation or Up-regulation of Synaptic Receptors. The long-term sensitivities of synapses can be changed tremendously by up-regulating the number of receptor proteins at the synaptic sites when there is underactivity and down-regulating the receptors when there is overactivity. The mechanism for this is the following: Receptor proteins are being formed constantly by the endoplasmic reticular–Golgi apparatus system and are constantly being inserted into the receptor neuron synaptic membrane. However, when the synapses are overused so that excesses of transmitter substance combine with the receptor proteins, many of these receptors are inactivated and removed from the synaptic membrane.

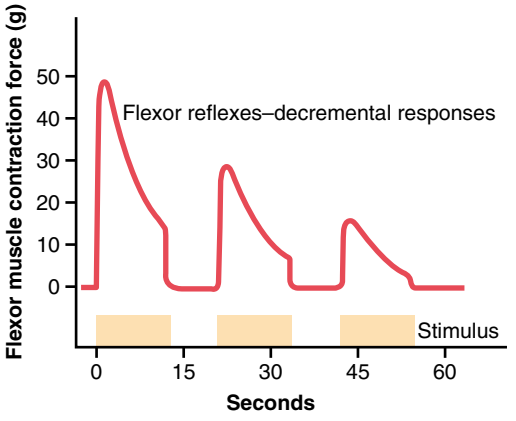


Figure 46-18 Successive flexor reflexes showing fatigue of conduction through the reflex pathway.

It is indeed fortunate that up-regulation and down-regulation of receptors, as well as other control mechanisms for adjusting synaptic sensitivity, continually adjust the sensitivity in each circuit to almost the exact level required for proper function. Think for a moment how serious it would be if the sensitivities of only a few of these circuits were abnormally high; one might then expect almost continual muscle cramps, seizures, psychotic disturbances, hallucinations, mental tension, or other nervous disorders. But fortunately, the automatic controls normally readjust the sensitivities of the circuits back to controllable ranges of reactivity any time the circuits begin to be too active or too depressed.

Bibliography

Bensmaia SJ: Tactile intensity and population codes, *Behav Brain Res* 190:165, 2008.

Buzsaki G: Large-scale recording of neuronal ensembles, *Nat Neurosci* 7:446, 2004.

Faisal AA, Selen LP, Wolpert DM: Noise in the nervous system, *Nat Rev Neurosci* 9:292, 2008.

Fontanini A, Katz DB: Behavioral states, network states, and sensory response variability, *J Neurophysiol* 100:1160, 2008.

Gandevia SC: Spinal and supraspinal factors in human muscle fatigue, *Physiol Rev* 81:1725, 2001.

Gebhart GF: Descending modulation of pain, *Neurosci Biobehav Rev* 27:729, 2004.

Hamill OP, Martinac B: Molecular basis of mechanotransduction in living cells, *Physiol Rev* 81:685, 2001.

Housley GD, Bringmann A: Reichenbach A Purinergic signaling in special senses, *Trends Neurosci* 32:128, 2009.

Kandel ER, Schwartz JH, Jessell TM: *Principles of Neural Science*, ed 4, New York, 2000, McGraw-Hill.

Katz DB, Matsunami H, Rinberg D, et al: Receptors, circuits, and behaviors: new directions in chemical senses, *J Neurosci* 28:11802, 2008.

Lumpkin EA, Caterina MJ: Mechanisms of sensory transduction in the skin, *Nature* 445:858, 2007.

Pearson KG: Neural adaptation in the generation of rhythmic behavior, *Annu Rev Physiol* 62:723, 2000.

Pugh JR, Raman IM: Nothing can be coincidence: synaptic inhibition and plasticity in the cerebellar nuclei, *Trends Neurosci* 32:170, 2009.

Ramocki MB, Zoghbi HY: Failure of neuronal homeostasis results in common neuropsychiatric phenotypes, *Nature* 455:912, 2008.

Richerson GB, Wu Y: Dynamic equilibrium of neurotransmitter transporters: not just for reuptake anymore, *J Neurophysiol* 90:1363, 2003.

Schepers RJ, Ringkamp M: Thermoreceptors and thermosensitive afferents, *Neurosci Biobehav Rev* 33:205, 2009.

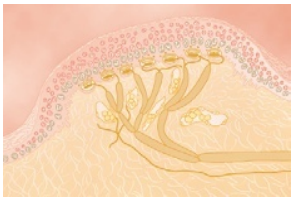
Schoppa NE: Making scents out of how olfactory neurons are ordered in space, *Nat Neurosci* 12:103, 2009.

Sjöström PJ, Rancz EA, Roth A, et al: Dendritic excitability and synaptic plasticity, *Physiol Rev* 88:769, 2008.

Stein BE, Stanford TR: Multisensory integration: current issues from the perspective of the single neuron, *Nat Rev Neurosci* 9:255, 2008.

CHAPTER 47

Somatic Sensations: I. General Organization, the Tactile and Position Senses



The *somatic senses* are the nervous mechanisms that collect sensory information from all over the body. These senses are in contradistinction to the *special senses,* which mean specifically vision, hearing, smell, taste, and equilibrium.

Classification of Somatic Senses

The somatic senses can be classified into three physiologic types: (1) the *mechanoreceptive somatic senses,* which include both *tactile* and *position* sensations that are stimulated by mechanical displacement of some tissue of the body; (2) the *thermoreceptive senses,* which detect heat and cold; and (3) the *pain sense,* which is activated by factors that damage the tissues.

This chapter deals with the mechanoreceptive tactile and position senses. Chapter 48 discusses the thermoreceptive and pain senses. The tactile senses include *touch, pressure, vibration,* and *tickle* senses, and the position senses include *static position* and *rate of movement* senses.

Other Classifications of Somatic Sensations. Somatic sensations are also often grouped together in other classes, as follows.

Exteroreceptive sensations are those from the surface of the body. *Proprioceptive sensations* are those relating to the physical state of the body, including position sensations, tendon and muscle sensations, pressure sensations from the bottom of the feet, and even the sensation of equilibrium (which is often considered a "special" sensation rather than a somatic sensation).

Visceral sensations are those from the viscera of the body; in using this term, one usually refers specifically to sensations from the internal organs.

Deep sensations are those that come from deep tissues, such as from fasciae, muscles, and bone. These include mainly "deep" pressure, pain, and vibration.

Detection and Transmission of Tactile Sensations

Interrelations Among the Tactile Sensations of Touch, Pressure, and Vibration. Although touch, pressure, and vibration are frequently classified as separate sensations, they are all detected by the same types of receptors. There are three principal differences among them: (1) touch sensation generally results from stimulation of tactile receptors in the skin or in tissues immediately beneath the skin; (2) pressure sensation generally results from deformation of deeper tissues; and (3) vibration sensation results from rapidly repetitive sensory signals, but some of the same types of receptors as those for touch and pressure are used.

Tactile Receptors. There are at least six entirely different types of tactile receptors, but many more similar to these also exist. Some were shown in Figure 46-1 of the previous chapter; their special characteristics are the following.

First, some *free nerve endings,* which are found everywhere in the skin and in many other tissues, can detect touch and pressure. For instance, even light contact with the cornea of the eye, which contains no other type of nerve ending besides free nerve endings, can nevertheless elicit touch and pressure sensations.

Second, a touch receptor with great sensitivity is the *Meissner's corpuscle* (illustrated in Figure 46-1), an elongated encapsulated nerve ending of a large (type Aβ) myelinated sensory nerve fiber. Inside the capsulation are many branching terminal nerve filaments. These corpuscles are present in the nonhairy parts of the skin and are particularly abundant in the fingertips, lips, and other areas of the skin where one's ability to discern spatial locations of touch sensations is highly developed. Meissner's corpuscles adapt in a fraction of a second after they are stimulated, which means that they are particularly sensitive to movement of objects over the surface of the skin, as well as to low-frequency vibration.

Third, the fingertips and other areas that contain large numbers of Meissner's corpuscles usually also contain large numbers of *expanded tip tactile receptors,* one type of which is *Merkel's discs,* shown in Figure 47-1. The hairy parts of the skin also contain moderate numbers of expanded tip receptors, even though they have almost no Meissner's corpuscles. These receptors differ from Meissner's corpuscles in that they transmit an initially strong but partially adapting signal and then a continuing weaker signal that adapts only slowly. Therefore, they are responsible for giving steady-state signals that allow one to determine continuous touch of objects against the skin.

Merkel's discs are often grouped together in a receptor organ called the *Iggo dome receptor,* which projects upward against the underside of the epithelium of the skin, as also shown in Figure 47-1. This causes the epithelium at this point to protrude outward, thus creating a dome and constituting an extremely sensitive receptor. Also note that the entire group of Merkel's discs is innervated by a single large myelinated nerve fiber (type Aβ). These receptors, along with the Meissner's corpuscles discussed earlier, play extremely important roles in localizing touch sensations to specific surface areas of the body and in determining the texture of what is felt.

Fourth, slight movement of any hair on the body stimulates a nerve fiber entwining its base. Thus, each hair and its basal nerve fiber, called the *hair end-organ,* are also touch receptors. A receptor adapts readily and, like Meissner's corpuscles, detects mainly (a) movement of objects on the surface of the body or (b) initial contact with the body.

Fifth, located in the deeper layers of the skin and also in still deeper internal tissues are many *Ruffini's endings,* which are multibranched, encapsulated endings, as shown in Figure 46-1. These endings adapt very slowly and, therefore, are important for signaling continuous states of deformation of the tissues, such as heavy prolonged touch and pressure signals. They are also found in joint capsules and help to signal the degree of joint rotation.

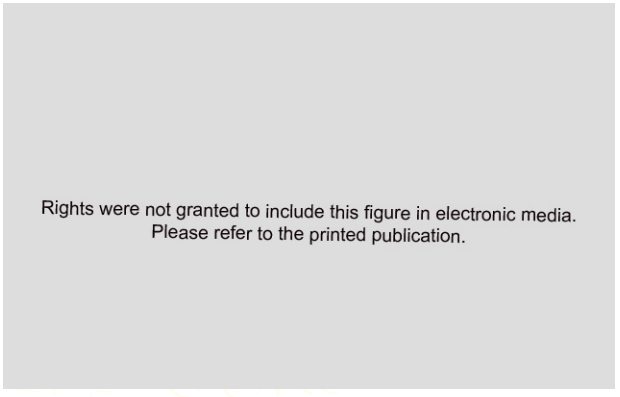


Figure 47-1 Iggo dome receptor. Note the multiple numbers of Merkel's discs connecting to a single large myelinated fiber and abutting tightly the undersurface of the epithelium. (From Iggo A, Muir AR: The structure and function of a slowly adapting touch corpuscle in hairy skin. J Physiol 200:763, 1969.)

Sixth, pacinian corpuscles, which were discussed in detail in Chapter 46, lie both immediately beneath the skin and deep in the fascial tissues of the body. They are stimulated only by rapid local compression of the tissues because they adapt in a few hundredths of a second. Therefore, they are particularly important for detecting tissue vibration or other rapid changes in the mechanical state of the tissues.

Transmission of Tactile Signals in Peripheral Nerve Fibers. Almost all specialized sensory receptors, such as Meissner's corpuscles, Iggo dome receptors, hair receptors, pacinian corpuscles, and Ruffini's endings, transmit their signals in type Aβ nerve fibers that have transmission velocities ranging from 30 to 70 m/sec. Conversely, free nerve ending tactile receptors transmit signals mainly by way of the small type Aδ myelinated fibers that conduct at velocities of only 5 to 30 m/sec.

Some tactile free nerve endings transmit by way of type C unmyelinated fibers at velocities from a fraction of a meter up to 2 m/sec; these send signals into the spinal cord and lower brain stem, probably subserving mainly the sensation of tickle.

Thus, the more critical types of sensory signals—those that help to determine precise localization on the skin, minute gradations of intensity, or rapid changes in sensory signal intensity—are all transmitted in more rapidly conducting types of sensory nerve fibers. Conversely, the cruder types of signals, such as pressure, poorly localized touch, and especially tickle, are transmitted by way of much slower, very small nerve fibers that require much less space in the nerve bundle than the fast fibers.

Detection of Vibration

All tactile receptors are involved in detection of vibration, although different receptors detect different frequencies of vibration. Pacinian corpuscles can detect signal vibrations from 30 to 800 cycles per second because they respond extremely rapidly to minute and rapid deformations of the tissues, and they also transmit their signals over type Aβ nerve fibers, which can transmit as many as 1000 impulses per second. Low-frequency vibrations from 2 up to 80 cycles per second, in contrast, stimulate other tactile receptors, especially Meissner's corpuscles, which are less rapidly adapting than pacinian corpuscles.

Detection of Tickle and Itch by Mechanoreceptive Free Nerve Endings

Neurophysiologic studies have demonstrated the existence of very sensitive, rapidly adapting mechanoreceptive free nerve endings that elicit only the tickle and itch sensations. Furthermore, these endings are found almost exclusively in superficial layers of the skin, which is also the only tissue from which the tickle and itch sensations usually can be elicited. These sensations are transmitted by very small type C, unmyelinated fibers similar to those that transmit the aching, slow type of pain.

The purpose of the itch sensation is presumably to call attention to mild surface stimuli such as a flea crawling on the skin or a fly about to bite, and the elicited signals then activate the scratch reflex or other maneuvers that rid the host of the irritant. Itch can be relieved by scratching if this removes the irritant or if the scratch is strong enough to elicit pain. The pain signals are believed to suppress the itch signals in the cord by lateral inhibition, as described in Chapter 48.

Sensory Pathways for Transmitting Somatic Signals into the Central Nervous System

Almost all sensory information from the somatic segments of the body enters the spinal cord through the dorsal roots of the spinal nerves. However, from the entry point into the cord and then to the brain, the sensory signals are carried through one of two alternative sensory pathways: (1) the *dorsal column*–medial lemniscal system or (2) the *anterolateral system*. These two systems come back together partially at the level of the thalamus.

The dorsal column–medial lemniscal system, as its name implies, carries signals upward to the medulla of the brain mainly in the *dorsal columns* of the cord. Then, after the signals synapse and cross to the opposite side in the medulla, they continue upward through the brain stem to the thalamus by way of the *medial lemniscus.*

Conversely, signals in the anterolateral system, immediately after entering the spinal cord from the dorsal spinal nerve roots, synapse in the dorsal horns of the spinal gray matter, then cross to the opposite side of the cord and ascend through the anterior and lateral white columns of the cord. They terminate at all levels of the lower brain stem and in the thalamus.

The dorsal column–medial lemniscal system is composed of large, myelinated nerve fibers that transmit signals to the brain at velocities of 30 to 110 m/sec, whereas the anterolateral system is composed of smaller myelinated fibers that transmit signals at velocities ranging from a few meters per second up to 40 m/sec.

Another difference between the two systems is that the dorsal column–medial lemniscal system has a high degree of spatial orientation of the nerve fibers with respect to their origin, while the anterolateral system has much less spatial orientation. These differences immediately characterize the types of sensory information that can be transmitted by the two systems. That is, sensory information that must be transmitted rapidly and with temporal and spatial fidelity is transmitted mainly in the dorsal column–medial lemniscal system; that which does not need to be transmitted rapidly or with great spatial fidelity is transmitted mainly in the anterolateral system.

The anterolateral system has a special capability that the dorsal system does not have: the ability to transmit a broad spectrum of sensory modalities—pain, warmth, cold, and crude tactile sensations; most of these are discussed in detail in Chapter 48. The dorsal system is limited to discrete types of mechanoreceptive sensations.

With this differentiation in mind, we can now list the types of sensations transmitted in the two systems.

Dorsal Column—Medial Lemniscal System

1. Touch sensations requiring a high degree of localization of the stimulus
2. Touch sensations requiring transmission of fine gradations of intensity
3. Phasic sensations, such as vibratory sensations
4. Sensations that signal movement against the skin
5. Position sensations from the joints
6. Pressure sensations related to fine degrees of judgment of pressure intensity

Anterolateral System

1. Pain
2. Thermal sensations, including both warmth and cold sensations
3. Crude touch and pressure sensations capable only of crude localizing ability on the surface of the body
4. Tickle and itch sensations
5. Sexual sensations

Transmission in the Dorsal Column–Medial Lemniscal System

Anatomy of the Dorsal Column–Medial Lemniscal System

On entering the spinal cord through the spinal nerve dorsal roots, the large myelinated fibers from the specialized mechanoreceptors divide almost immediately to form a *medial branch* and a *lateral branch,* shown by the right-hand fiber entering through the spinal root in Figure 47-2.

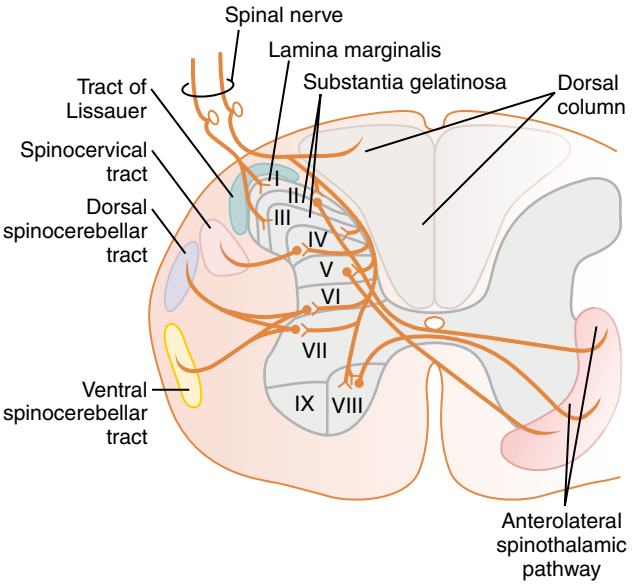


Figure 47-2 Cross section of the spinal cord, showing the anatomy of the cord gray matter and of ascending sensory tracts in the white columns of the spinal cord.

The medial branch turns medially first and then upward in the dorsal column, proceeding by way of the dorsal column pathway all the way to the brain.

The lateral branch enters the dorsal horn of the cord gray matter, then divides many times to provide terminals that synapse with local neurons in the intermediate and anterior portions of the cord gray matter. These local neurons in turn serve three functions: (1) A major share of them give off fibers that enter the dorsal columns of the cord and then travel upward to the brain. (2) Many of the fibers are very short and terminate locally in the spinal cord gray matter to elicit local spinal cord reflexes, which are discussed in Chapter 54. (3) Others give rise to the spinocerebellar tracts, which we discuss in Chapter 56 in relation to the function of the cerebellum.

Dorsal Column—Medial Lemniscal Pathway. Note in Figure 47-3 that nerve fibers entering the dorsal columns pass uninterrupted up to the dorsal medulla, where they synapse in the *dorsal column nuclei* (the *cuneate* and *gracile nuclei*). From there, *second-order neurons* decussate immediately to the opposite side of the brain stem and continue upward through the *medial lemnisci* to the thalamus. In this pathway through the brain stem, each medial lemniscus is joined by additional fibers from the *sensory nuclei of the trigeminal nerve;* these fibers subserve the same sensory functions for the head that the dorsal column fibers subserve for the body.

In the thalamus, the medial lemniscal fibers terminate in the thalamic sensory relay area, called the *ventrobasal complex.* From the ventrobasal complex, *third-order nerve fibers* project, as shown in Figure 47-4, mainly to the *postcentral gyrus* of the *cerebral cortex,* which is called *somatic sensory area I* (as shown in Figure 47-6, these fibers also project to a smaller area in the lateral parietal cortex called *somatic sensory area II*).

Spatial Orientation of the Nerve Fibers in the Dorsal Column–Medial Lemniscal System

One of the distinguishing features of the dorsal column–medial lemniscal system is a distinct spatial orientation of nerve fibers from the individual parts of the body that is maintained throughout. For instance, in the dorsal columns of the spinal cord, the fibers from the lower parts of the body lie toward the center of the cord, whereas those that enter the cord at progressively higher segmental levels form successive layers laterally.

In the thalamus, distinct spatial orientation is still maintained, with the tail end of the body represented by the most lateral portions of the ventrobasal complex and the head and face represented by the medial areas of the complex. Because of the crossing of the medial lemnisci in the medulla, the left side of the body is represented in the right side of the thalamus, and the right side of the body in the left side of the thalamus.

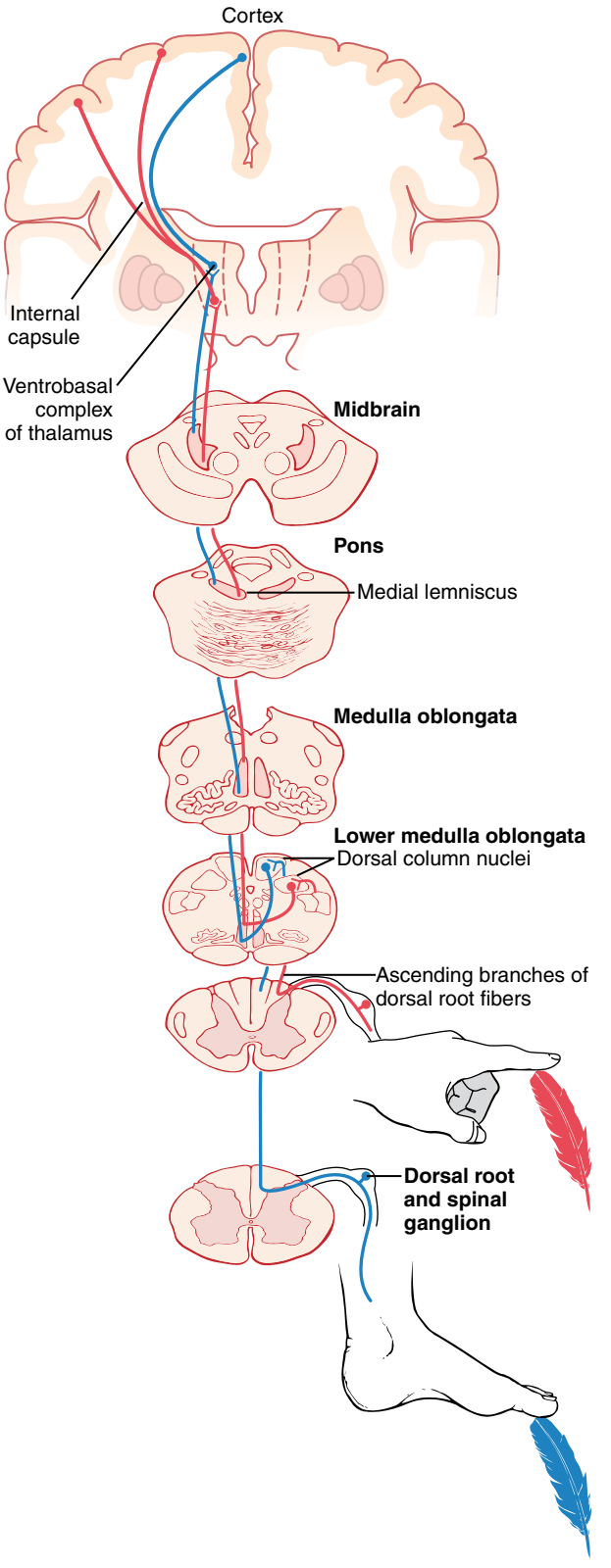


Figure 47-3 The dorsal column–medial lemniscal pathway for transmitting critical types of tactile signals.

Somatosensory Cortex

Before discussing the role of the cerebral cortex in somatic sensation, we need to give an orientation to the various areas of the cortex. Figure 47-5 is a map of the human cerebral cortex, showing that it is divided into about 50 distinct areas called *Brodmann's areas* based on histological structural differences. This map is important because virtually all neurophysiologists and neurologists use it to

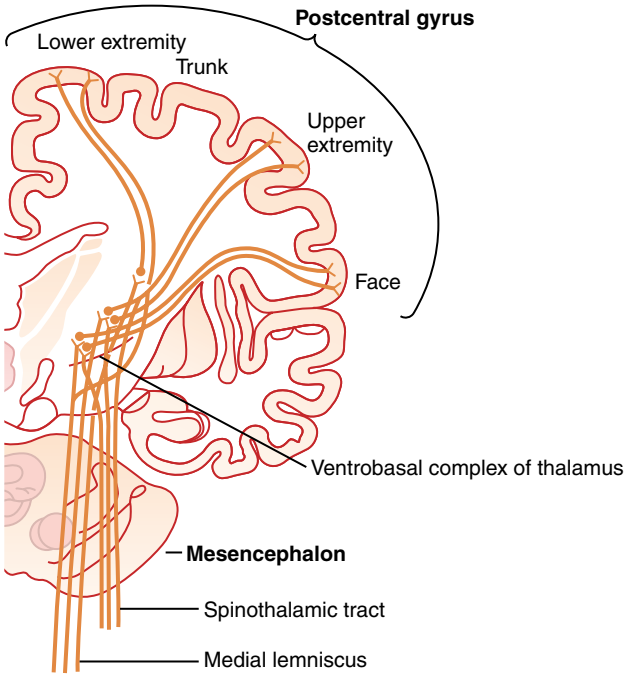


Figure 47-4 Projection of the dorsal column–medial lemniscal system through the thalamus to the somatosensory cortex. (Modified from Brodal A: Neurological Anatomy in Relation to Clinical Medicine. New York: Oxford University Press, 1969, by permission of Oxford University Press.)

refer by number to many of the different functional areas of the human cortex.

Note in the figure the large *central fissure* (also called *central sulcus*) that extends horizontally across the brain. In general, sensory signals from all modalities of sensation terminate in the cerebral cortex immediately posterior to the central fissure. And, generally, the anterior half of the *parietal lobe* is concerned almost entirely with reception and interpretation of *somatosensory signals.* But the posterior half of the parietal lobe provides still higher levels of interpretation.

Visual signals terminate in the *occipital lobe,* and *auditory signals* terminate in the *temporal lobe.*

Conversely, that portion of the cerebral cortex anterior to the central fissure and constituting the posterior half of the frontal lobe is called the *motor cortex* and is devoted almost entirely to control of muscle contractions and body movements. A major share of this motor control is in response to somatosensory signals received from the sensory portions of the cortex, which keep the motor cortex informed at each instant about the positions and motions of the different body parts.

Somatosensory Areas I and II. Figure 47-6 shows two separate sensory areas in the anterior parietal lobe called *somatosensory area I* and *somatosensory area II.* The reason for this division into two areas is that a distinct and separate spatial orientation of the different parts of the body is found in each of these two areas. However, somatosensory area I is so much more extensive and so much more important than somatosensory area II that in popular usage, the term "somatosensory cortex" almost always means area I.

Somatosensory area I has a high degree of localization of the different parts of the body, as shown by the names of virtually all parts of the body in Figure 47-6. By contrast, localization is poor in somatosensory area II, although roughly, the face is represented anteriorly, the arms centrally, and the legs posteriorly.

Little is known about the function of somatosensory area II. It is known that signals enter this area from the brain stem, transmitted upward from both sides of the body. In addition, many signals come secondarily from somatosensory area I, as well as from other sensory areas of the brain, even from the visual and auditory areas. Projections from somatosensory area I are required for function of somatosensory area II. However, removal of parts of somatosensory area II has no apparent effect on the response of neurons in somatosensory area I. Thus, much of what we know about somatic sensation appears to be explained by the functions of somatosensory area I.

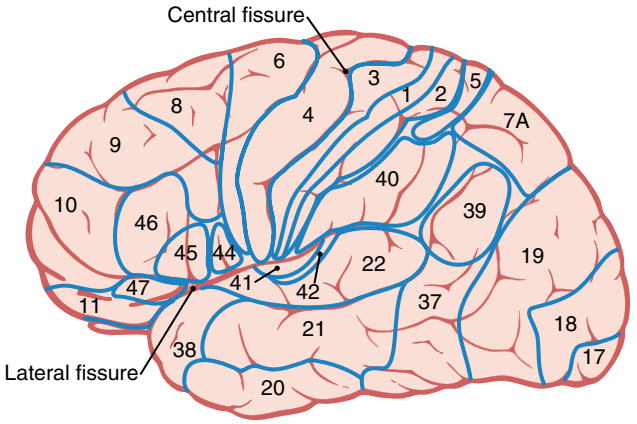


Figure 47-5 Structurally distinct areas, called *Brodmann's areas,* of the human cerebral cortex. Note specifically areas *1, 2,* and *3,* which constitute *primary somatosensory area I,* and areas *5* and *7,* which constitute the *somatosensory association area.*

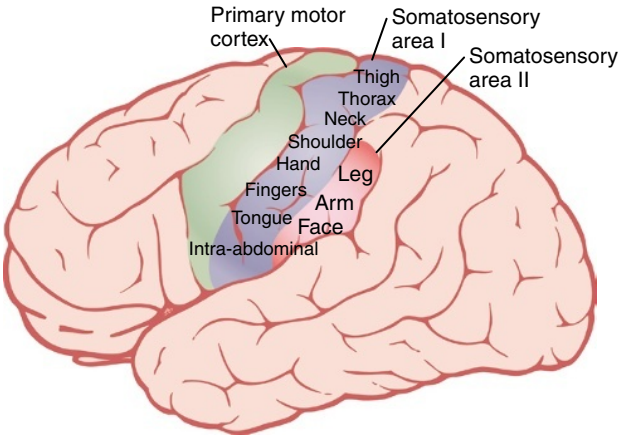


Figure 47-6 Two somatosensory cortical areas, somatosensory areas I and II.

Spatial Orientation of Signals from Different Parts of the Body in Somatosensory Area I. Somatosensory area I lies immediately behind the central fissure, located in the postcentral gyrus of the human cerebral cortex (in Brodmann's areas 3, 1, and 2).

Figure 47-7 shows a cross section through the brain at the level of the *postcentral gyrus,* demonstrating representations of the different parts of the body in separate regions of somatosensory area I. Note, however, that each lateral side of the cortex receives sensory information almost exclusively from the opposite side of the body.

Some areas of the body are represented by large areas in the somatic cortex—the lips the greatest of all, followed by the face and thumb—whereas the trunk and lower part of the body are represented by relatively small areas. The sizes of these areas are directly proportional to the number of specialized sensory receptors in each respective peripheral area of the body. For instance, a great number of specialized nerve endings are found in the lips and thumb, whereas only a few are present in the skin of the body trunk.

Note also that the head is represented in the most lateral portion of somatosensory area I, and the lower part of the body is represented medially.

Layers of the Somatosensory Cortex and Their Function

The cerebral cortex contains *six* layers of neurons, beginning with layer I next to the brain surface and extending progressively deeper to layer VI, shown in Figure 47-8.

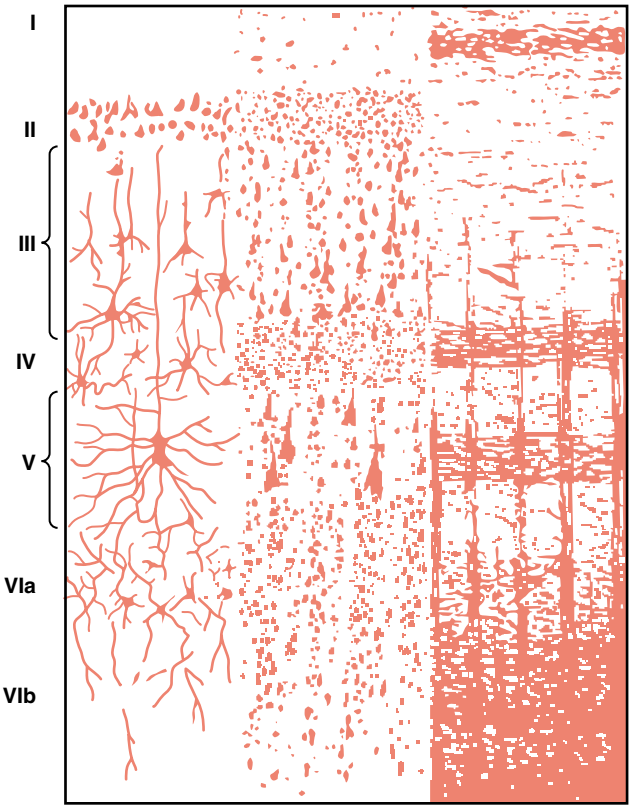


Figure 47-8 Structure of the cerebral cortex, showing I, molecular layer; II, external granular layer; III, layer of small pyramidal cells; IV, internal granular layer; V, large pyramidal cell layer; and VI, layer of fusiform or polymorphic cells. (From Ranson SW, Clark SL [after Brodmann]: Anatomy of the Nervous System. Philadelphia: WB Saunders, 1959.)

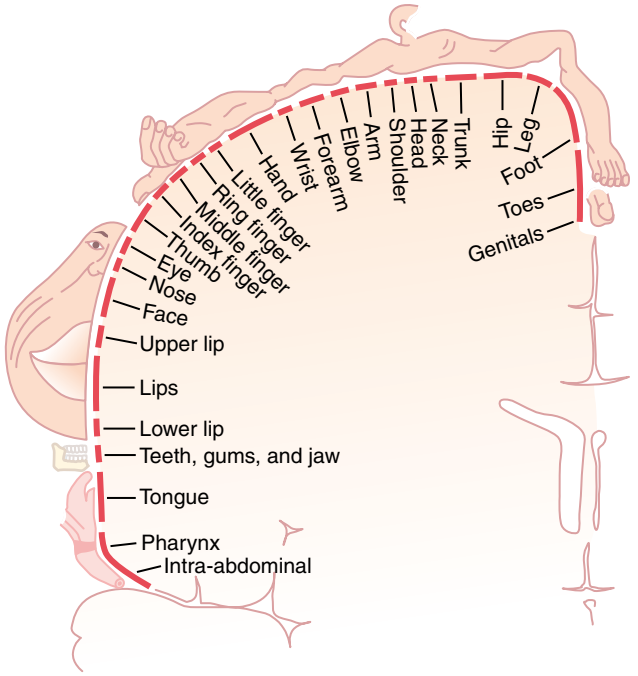


Figure 47-7 Representation of the different areas of the body in somatosensory area I of the cortex. (From Penfield W, Rasmussen T: Cerebral Cortex of Man: A Clinical Study of Localization of Function. New York: Hafner, 1968.)

As would be expected, the neurons in each layer perform functions different from those in other layers. Some of these functions are:

1. The incoming sensory signal excites neuronal layer IV first; then the signal spreads toward the surface of the cortex and also toward deeper layers.
2. Layers I and II receive diffuse, nonspecific input signals from lower brain centers that facilitate specific regions of the cortex; this system is described in Chapter 57. This input mainly controls the overall level of excitability of the respective regions stimulated.
3. The neurons in layers II and III send axons to related portions of the cerebral cortex on the opposite side of the brain through the *corpus callosum.*
4. The neurons in layers V and VI send axons to the deeper parts of the nervous system. Those in layer V are generally larger and project to more distant areas, such as to the basal ganglia, brain stem, and spinal cord, where they control signal transmission. From layer VI, especially large numbers of axons extend to the thalamus, providing signals from the cerebral cortex that interact with and help to control the excitatory levels of incoming sensory signals entering the thalamus.

The Sensory Cortex Is Organized in Vertical Columns of Neurons; Each Column Detects a Different Sensory Spot on the Body with a Specific Sensory Modality

Functionally, the neurons of the somatosensory cortex are arranged in vertical columns extending all the way through the six layers of the cortex, each column having a diameter of 0.3 to 0.5 millimeter and containing perhaps 10,000 neuronal cell bodies. Each of these columns serves a single specific sensory modality, some columns responding to stretch receptors around joints, some to stimulation of tactile hairs, others to discrete localized pressure points on the skin, and so forth. At layer IV, where the input sensory signals first enter the cortex, the columns of neurons function almost entirely separately from one another. At other levels of the columns, interactions occur that initiate analysis of the meanings of the sensory signals.

In the most anterior 5 to 10 millimeters of the postcentral gyrus, located deep in the central fissure in Brodmann's area 3a, an especially large share of the vertical columns respond to muscle, tendon, and joint stretch receptors. Many of the signals from these sensory columns then spread anteriorly, directly to the motor cortex located immediately forward of the central fissure. These signals play a major role in controlling the effluent motor signals that activate sequences of muscle contraction.

As one moves posteriorly in somatosensory area I, more and more of the vertical columns respond to slowly adapting cutaneous receptors, and then still farther posteriorly, greater numbers of the columns are sensitive to deep pressure.

In the most posterior portion of somatosensory area I, about 6 percent of the vertical columns respond only when a stimulus moves across the skin in a particular direction. Thus, this is a still higher order of interpretation of sensory signals; the process becomes even more complex as the signals spread farther backward from somatosensory area I into the parietal cortex, an area called the *somatosensory association area,* as we discuss subsequently.

Functions of Somatosensory Area I

Widespread bilateral excision of somatosensory area I causes loss of the following types of sensory judgment:

1. The person is unable to localize discretely the different sensations in the different parts of the body. However, he or she can localize these sensations crudely, such as to a particular hand, to a major level of the body trunk, or to one of the legs. Thus, it is clear that the brain stem, thalamus, or parts of the cerebral cortex not normally considered to be concerned with somatic sensations can perform some degree of localization.
2. The person is unable to judge critical degrees of pressure against the body.
3. The person is unable to judge the weights of objects.
4. The person is unable to judge shapes or forms of objects. This is called *astereognosis.*
5. The person is unable to judge texture of materials because this type of judgment depends on highly critical sensations caused by movement of the fingers over the surface to be judged.

Note that in the list nothing has been said about loss of pain and temperature sense. In specific absence of only somatosensory area I, appreciation of these sensory modalities is still preserved both in quality and intensity. But the sensations are poorly localized, indicating that pain and temperature *localization* depend greatly on the topographical map of the body in somatosensory area I to localize the source.

Somatosensory Association Areas

Brodmann's areas 5 and 7 of the cerebral cortex, located in the parietal cortex behind somatosensory area I (see Figure 47-5), play important roles in deciphering deeper meanings of the sensory information in the somatosensory areas. Therefore, these areas are called *somatosensory association areas.*

Electrical stimulation in a somatosensory association area can occasionally cause an awake person to experience a complex body sensation, sometimes even the "feeling" of an object such as a knife or a ball. Therefore, it seems clear that the somatosensory association area combines information arriving from multiple points in the primary somatosensory area to decipher its meaning. This also fits with the anatomical arrangement of the neuronal tracts that enter the somatosensory association area because it receives signals from (1) somatosensory area I, (2) the ventrobasal nuclei of the thalamus, (3) other areas of the thalamus, (4) the visual cortex, and (5) the auditory cortex.

Effect of Removing the Somatosensory Association Area—Amorphosynthesis. When the somatosensory association area is removed on one side of the brain, the person loses ability to recognize complex objects and complex forms felt on the opposite side of the body. In addition, he or she loses most of the sense of form of his or her own body or body parts on the opposite side. In fact, the person is mainly oblivious to the opposite side of the body—that is, forgets that it is there. Therefore, he or she also often forgets to use the other side for motor functions as well. Likewise, when feeling objects, the person tends to recognize only one side of the object and forgets that the other side even exists. This complex sensory deficit is called *amorphosynthesis.*

Overall Characteristics of Signal Transmission and Analysis in the Dorsal Column–Medial Lemniscal System

Basic Neuronal Circuit in the Dorsal Column–Medial Lemniscal System. The lower part of Figure 47-9 shows the basic organization of the neuronal circuit of the spinal cord dorsal column pathway, demonstrating that at each synaptic stage, divergence occurs. The upper

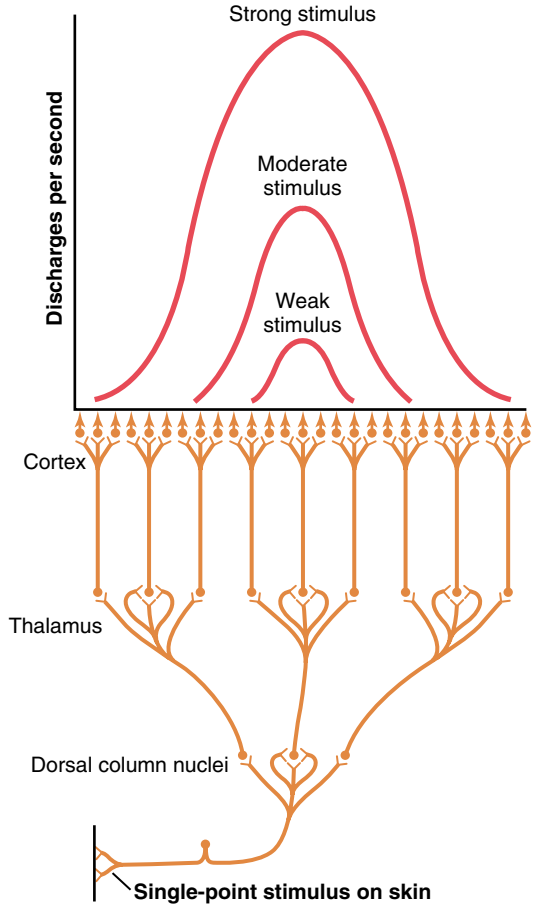


Figure 47-9 Transmission of a pinpoint stimulus signal to the cerebral cortex.

curves of the figure show that the cortical neurons that discharge to the greatest extent are those in a central part of the cortical "field" for each respective receptor. Thus, a weak stimulus causes only the centralmost neurons to fire. A stronger stimulus causes still more neurons to fire, but those in the center discharge at a considerably more rapid rate than do those farther away from the center.

Two-Point Discrimination. A method frequently used to test tactile discrimination is to determine a person's so-called "two-point" discriminatory ability. In this test, two needles are pressed lightly against the skin at the same time, and the person determines whether two points of stimulus are felt or one point. On the tips of the fingers, a person can normally distinguish two separate points even when the needles are as close together as 1 to 2 millimeters. However, on the person's back, the needles must usually be as far apart as 30 to 70 millimeters before two separate points can be detected. The reason for this difference is the different numbers of specialized tactile receptors in the two areas.

Figure 47-10 shows the mechanism by which the dorsal column pathway (as well as all other sensory pathways) transmits two-point discriminatory information. This figure shows two adjacent points on the skin that are strongly stimulated, as well as the areas of the somatosensory

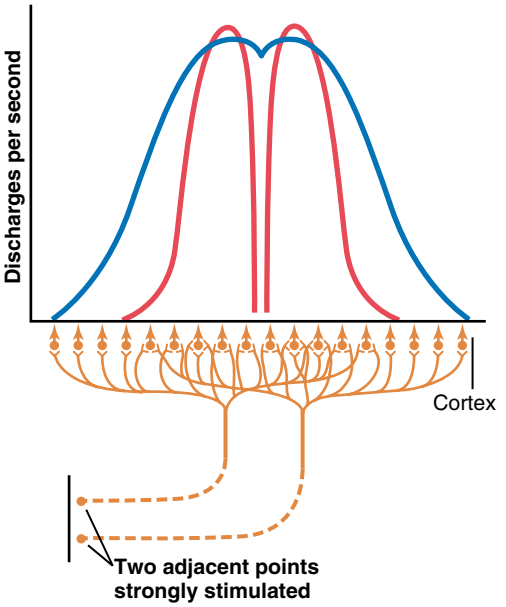


Figure 47-10 Transmission of signals to the cortex from two adjacent pinpoint stimuli. The blue curve represents the pattern of cortical stimulation without "surround" inhibition, and the two red curves represent the pattern when "surround" inhibition does occur.

cortex (greatly enlarged) that are excited by signals from the two stimulated points. The blue curve shows the spatial pattern of cortical excitation when both skin points are stimulated simultaneously. Note that the resultant zone of excitation has two separate peaks. These two peaks, separated by a valley, allow the sensory cortex to detect the presence of two stimulatory points, rather than a single point. The capability of the sensorium to distinguish this presence of two points of stimulation is strongly influenced by another mechanism, *lateral inhibition,* as explained in the next section.

Effect of Lateral Inhibition (Also Called *Surround Inhibition*) to Increase the Degree of Contrast in the Perceived Spatial Pattern. As pointed out in Chapter 46, virtually every sensory pathway, when excited, gives rise simultaneously to lateral *inhibitory* signals; these spread to the sides of the excitatory signal and inhibit adjacent neurons. For instance, consider an excited neuron in a dorsal column nucleus. Aside from the central excitatory signal, short lateral pathways transmit inhibitory signals to the surrounding neurons. That is, these signals pass through additional interneurons that secrete an inhibitory transmitter.

The importance of *lateral inhibition* is that it blocks lateral spread of the excitatory signals and, therefore, increases the degree of contrast in the sensory pattern perceived in the cerebral cortex.

In the case of the dorsal column system, lateral inhibitory signals occur at each synaptic level—for instance, in (1) the dorsal column nuclei of the medulla, (2) the ventrobasal nuclei of the thalamus, and (3) the cortex itself. At each of these levels, the lateral inhibition helps to block lateral spread of the excitatory signal. As a result, the peaks

of excitation stand out, and much of the surrounding diffuse stimulation is blocked. This effect is demonstrated by the two red curves in Figure 47-10, showing complete separation of the peaks when the intensity of lateral inhibition is great.

Transmission of Rapidly Changing and Repetitive Sensations. The dorsal column system is also of particular importance in apprising the sensorium of rapidly changing peripheral conditions. Based on recorded action potentials, this system can recognize changing stimuli that occur in as little as 1/400 of a second.

Vibratory Sensation. Vibratory signals are rapidly repetitive and can be detected as vibration up to 700 cycles per second. The higher-frequency vibratory signals originate from the pacinian corpuscles in the skin and deeper tissues, but lower-frequency signals (below about 200 per second) can originate from Meissner's corpuscles as well. These signals are transmitted only in the dorsal column pathway. For this reason, application of vibration (e.g., from a "tuning fork") to different peripheral parts of the body is an important tool used by neurologists for testing functional integrity of the dorsal columns.

Interpretation of Sensory Stimulus Intensity

The ultimate goal of most sensory stimulation is to apprise the psyche of the state of the body and its surroundings. Therefore, it is important that we discuss briefly some of the principles related to transmission of sensory *stimulus intensity* to the higher levels of the nervous system.

One question that comes to mind is, how is it possible for the sensory system to transmit sensory experiences of tremendously varying intensities? For instance, the auditory system can detect the weakest possible whisper but can also discern the meanings of an explosive sound, even though the sound intensities of these two experiences can vary more than 10 billion times; the eyes can see visual images with light intensities that vary as much as a half million times; and the skin can detect pressure differences of 10,000 to 100,000 times.

As a partial explanation of these effects, Figure 46-4 in the previous chapter shows the relation of the receptor potential produced by the pacinian corpuscle to the intensity of the sensory stimulus. At low stimulus intensity, slight changes in intensity increase the potential markedly, whereas at high levels of stimulus intensity, further increases in receptor potential are slight. Thus, the pacinian corpuscle is capable of accurately measuring extremely minute *changes* in stimulus at low-intensity levels, but at high-intensity levels, the change in stimulus must be much greater to cause the same amount of *change* in receptor potential.

The transduction mechanism for detecting sound by the cochlea of the ear demonstrates still another method for separating gradations of stimulus intensity. When sound stimulates a specific point on the basilar membrane, weak sound stimulates only those hair cells at the point of maximum sound vibration. But as the sound intensity increases, many more hair cells in each direction farther away from the maximum vibratory point also become stimulated. Thus, signals are transmitted over progressively increasing numbers of nerve fibers, which is another mechanism by which stimulus intensity is transmitted to the central nervous system. This mechanism, plus the direct effect of stimulus intensity on impulse rate in each nerve fiber, as well as several other mechanisms, makes it possible for some sensory systems to operate reasonably faithfully at stimulus intensity levels changing as much as millions of times.

Importance of the Tremendous Intensity Range of Sensory Reception. Were it not for the tremendous intensity range of sensory reception that we can experience, the various sensory systems would more often than not be operating in the wrong range. This is demonstrated by the attempts of most people, when taking photographs with a camera, to adjust the light exposure without using a light meter. Left to intuitive judgment of light intensity, a person almost always overexposes the film on bright days and greatly underexposes the film at twilight. Yet that person's own eyes are capable of discriminating with great detail visual objects in bright sunlight or at twilight; the camera cannot do this without very special manipulation because of the narrow critical range of light intensity required for proper exposure of film.

Judgment of Stimulus Intensity

Weber-Fechner Principle—Detection of "Ratio" of Stimulus Strength. In the mid-1800s, Weber first and Fechner later proposed the principle that *gradations of stimulus strength are discriminated approximately in proportion to the logarithm of stimulus strength.* That is, a person already holding 30 grams weight in his or her hand can barely detect an additional 1-gram increase in weight. And, when already holding 300 grams, he or she can barely detect a 10-gram increase in weight. Thus, in this instance, the *ratio* of the change in stimulus strength required for detection remains essentially constant, about 1 to 30, which is what the logarithmic principle means. To express this mathematically.

$$\text{Interpreted signal strength} = \text{Log (Stimulus)} + \text{Constant}$$

More recently, it has become evident that the Weber-Fechner principle is quantitatively accurate only for higher intensities of visual, auditory, and cutaneous sensory experience and applies only poorly to most other types of sensory experience. Yet the Weber-Fechner principle is still a good one to remember because it emphasizes that the greater the background sensory intensity, the greater an additional change must be for the psyche to detect the change.

Power Law. Another attempt by physiopsychologists to find a good mathematical relation is the following formula, known as the power law.

$$\text{Interpreted signal strength} = K \times (\text{Stimulus} - k)^{y}$$

In this formula, the exponent y and the constants K and k are different for each type of sensation.

When this power law relation is plotted on a graph using double logarithmic coordinates, as shown in Figure 47-11, and when appropriate quantitative values for the constants y, K, and k are found, a linear relation can be attained between interpreted stimulus strength and actual stimulus strength over a large range for almost any type of sensory perception.

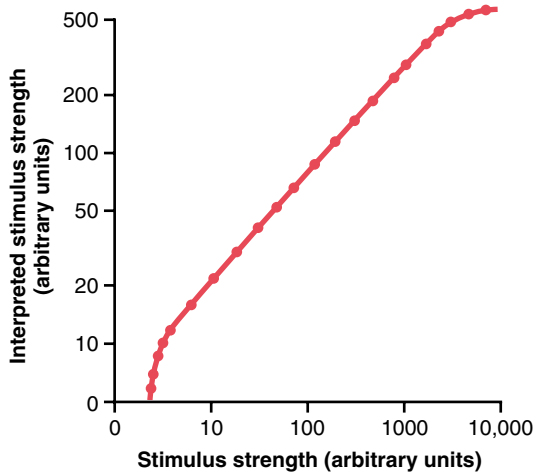


Figure 47-11 Graphical demonstration of the "power law" relation between actual stimulus strength and strength that the psyche interprets it to be. Note that the power law does not hold at either very weak or very strong stimulus strengths.

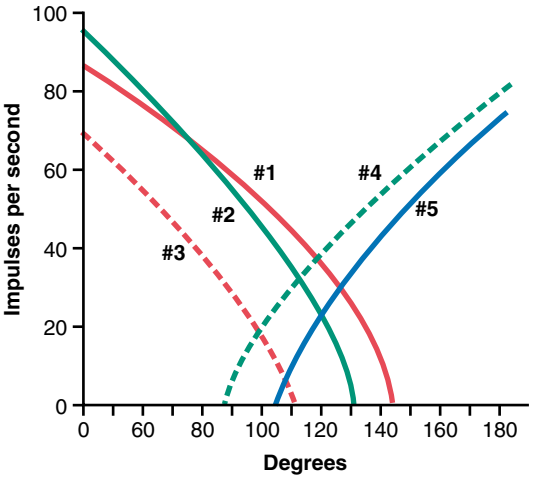


Figure 47-12 Typical responses of five different thalamic neurons in the thalamic ventrobasal complex when the knee joint is moved through its range of motion. (Data from Mountcastle VB, Poggie GF, Werner G: The relation of thalamic cell response to peripheral stimuli varied over an intensive continuum. J Neurophysiol 26:807, 1963.)

Position Senses

The *position senses* are frequently also called *proprioceptive senses*. They can be divided into two subtypes: (1) *static position sense,* which means conscious perception of the orientation of the different parts of the body with respect to one another, and (2) *rate of movement sense,* also called *kinesthesia* or *dynamic proprioception.*

Position Sensory Receptors. Knowledge of position, both static and dynamic, depends on knowing the degrees of angulation of all joints in all planes and their rates of change. Therefore, multiple different types of receptors help to determine joint angulation and are used together for position sense. Both skin tactile receptors and deep receptors near the joints are used. In the case of the fingers, where skin receptors are in great abundance, as much as half of position recognition is believed to be detected through the skin receptors. Conversely, for most of the larger joints of the body, deep receptors are more important.

For determining joint angulation in midranges of motion, among the most important receptors are the *muscle spindles.* They are also exceedingly important in helping to control muscle movement, as we shall see in Chapter 54. When the angle of a joint is changing, some muscles are being stretched while others are loosened, and the net stretch information from the spindles is transmitted into the computational system of the spinal cord and higher regions of the dorsal column system for deciphering joint angulations.

At the extremes of joint angulation, stretch of the ligaments and deep tissues around the joints is an additional important factor in determining position. Types of sensory endings used for this are the pacinian corpuscles, Ruffini's endings, and receptors similar to the Golgi tendon receptors found in muscle tendons.

The pacinian corpuscles and muscle spindles are especially adapted for detecting rapid rates of change. It is likely that these are the receptors most responsible for detecting rate of movement.

Processing of Position Sense Information in the Dorsal Column–Medial Lemniscal Pathway. Referring to Figure 47-12, one sees that *thalamic neurons* responding to joint rotation are of two categories: (1) those maximally stimulated when the joint is at full rotation and (2) those maximally stimulated when the joint is at minimal rotation. Thus, the signals from the individual joint receptors are used to tell the psyche how much each joint is rotated.

Transmission of Less Critical Sensory Signals in the Anterolateral Pathway

The anterolateral pathway for transmitting sensory signals up the spinal cord and into the brain, in contrast to the dorsal column pathway, transmits sensory signals that do not require highly discrete localization of the signal source and do not require discrimination of fine gradations of intensity. These types of signals include pain, heat, cold, crude tactile, tickle, itch, and sexual sensations. In Chapter 48, pain and temperature sensations are discussed specifically.

Anatomy of the Anterolateral Pathway

The *spinal cord anterolateral fibers* originate mainly in dorsal horn laminae I, IV, V, and VI (see Figure 47-2). These laminae are where many of the dorsal root sensory nerve fibers terminate after entering the cord.

As shown in Figure 47-13, the anterolateral fibers cross immediately in the *anterior commissure* of the cord to the opposite *anterior* and *lateral white columns,* where they turn upward toward the brain by way of the *anterior spinothalamic* and *lateral spinothalamic tracts.*

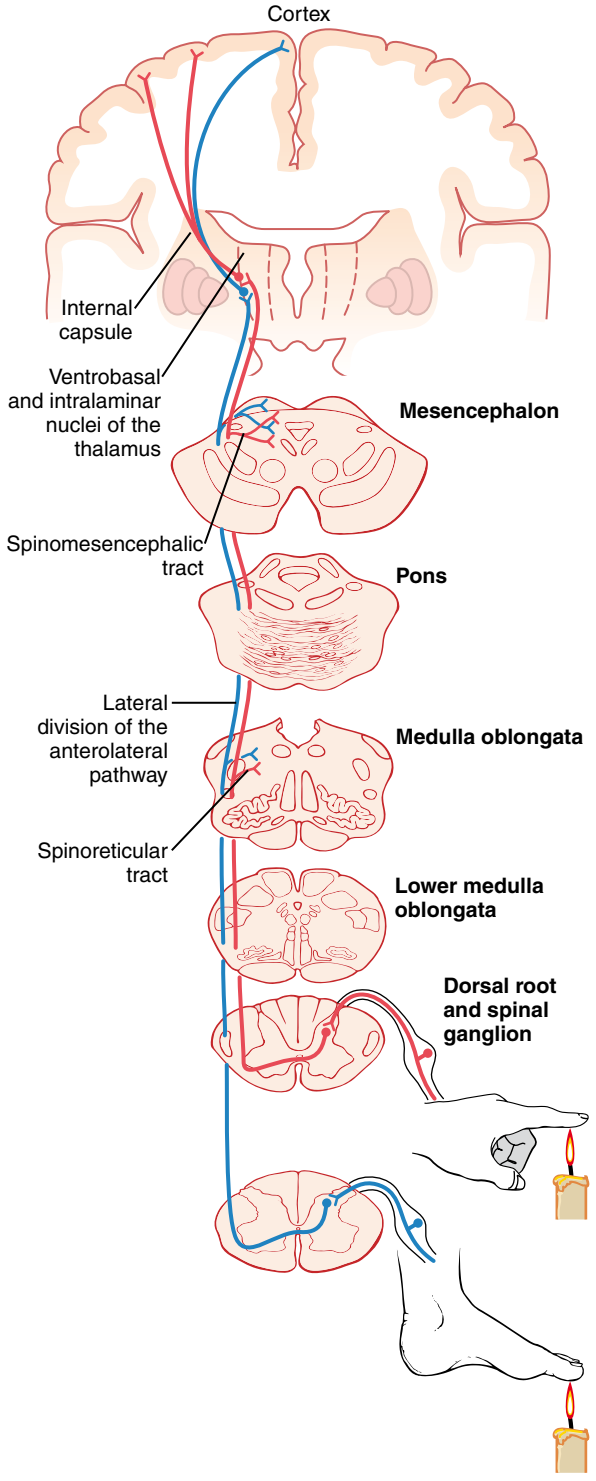


Figure 47-13 Anterior and lateral divisions of the anterolateral sensory pathway.

The upper terminus of the two spinothalamic tracts is mainly twofold: (1) throughout the *reticular nuclei of the brain stem* and (2) in two different nuclear complexes of the thalamus, the *ventrobasal complex* and the *intralaminar nuclei.* In general, the tactile signals are transmitted mainly into the ventrobasal complex, terminating in some of the same thalamic nuclei where the dorsal column tactile signals terminate. From here, the signals are transmitted to the somatosensory cortex along with the signals from the dorsal columns.

Conversely, only a small fraction of the pain signals project directly to the ventrobasal complex of the thalamus. Instead, most pain signals terminate in the reticular nuclei of the brain stem and from there are relayed to the intralaminar nuclei of the thalamus where the pain signals are further processed, as discussed in greater detail in Chapter 48.

Characteristics of Transmission in the Anterolateral Pathway. In general, the same principles apply to transmission in the anterolateral pathway as in the dorsal column–medial lemniscal system, except for the following differences: (1) the velocities of transmission are only one-third to one-half those in the dorsal column–medial lemniscal system, ranging between 8 and 40 m/sec; (2) the degree of spatial localization of signals is poor; (3) the gradations of intensities are also far less accurate, most of the sensations being recognized in 10 to 20 gradations of strength, rather than as many as 100 gradations for the dorsal column system; and (4) the ability to transmit rapidly changing or rapidly repetitive signals is poor.

Thus, it is evident that the anterolateral system is a cruder type of transmission system than the dorsal column–medial lemniscal system. Even so, certain modalities of sensation are transmitted only in this system and not at all in the dorsal column–medial lemniscal system. They are pain, temperature, tickle, itch, and sexual sensations, in addition to crude touch and pressure.

Some Special Aspects of Somatosensory Function

Function of the Thalamus in Somatic Sensation

When the somatosensory cortex of a human being is destroyed, that person loses most critical tactile sensibilities, but a slight degree of crude tactile sensibility does return. Therefore, it must be assumed that the thalamus (as well as other lower centers) has a slight ability to discriminate tactile sensation, even though the thalamus normally functions mainly to relay this type of information to the cortex.

Conversely, loss of the somatosensory cortex has little effect on one's perception of pain sensation and only a moderate effect on the perception of temperature. Therefore, there is much reason to believe that the lower brain stem, the thalamus, and other associated basal regions of the brain play dominant roles in discrimination of these sensibilities. It is interesting that these sensibilities appeared very early in the phylogenetic development of animals, whereas the critical tactile sensibilities and the somatosensory cortex were late developments.

Cortical Control of Sensory Sensitivity—"Corticofugal" Signals

In addition to somatosensory signals transmitted from the periphery to the brain, *corticofugal* signals are transmitted in the backward direction from the cerebral cortex to the lower sensory relay stations of the thalamus, medulla, and spinal cord; they control the intensity of sensitivity of the sensory input.

Corticofugal signals are almost entirely inhibitory, so when sensory input intensity becomes too great, the corticofugal signals automatically decrease transmission in

the relay nuclei. This does two things: First, it decreases lateral spread of the sensory signals into adjacent neurons and, therefore, increases the degree of sharpness in the signal pattern. Second, it keeps the sensory system operating in a range of sensitivity that is not so low that the signals are ineffectual nor so high that the system is swamped beyond its capacity to differentiate sensory patterns. This principle of corticofugal sensory control is used by all sensory systems, not only the somatic system, as explained in subsequent chapters.

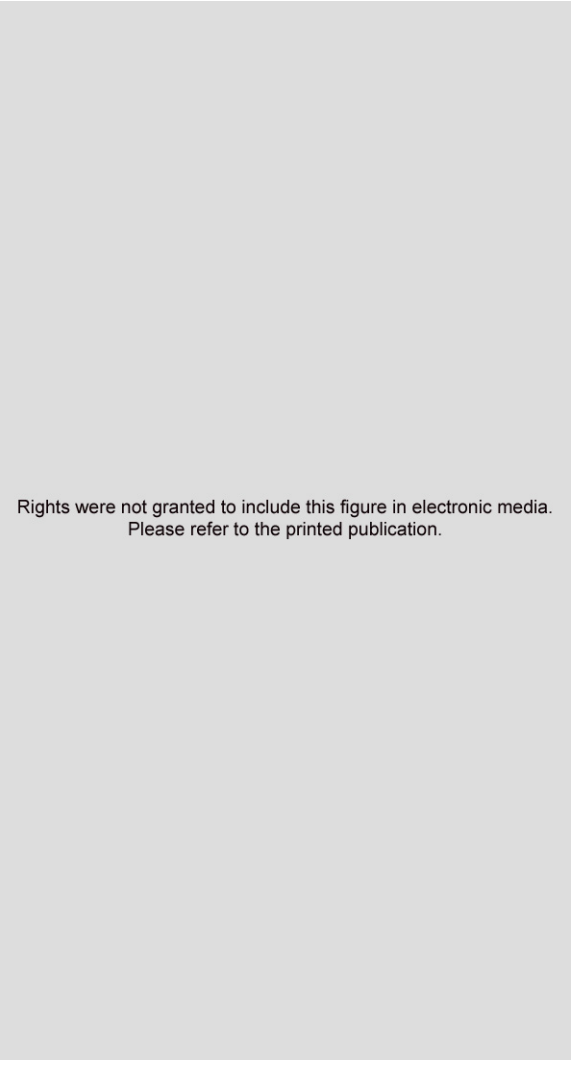


Figure 47-14 Dermatomes. (Modified from Grinker RR, Sahs AL: Neurology, 6th ed. Springfield, Ill: Charles C Thomas, 1966. Courtesy Charles C Thomas, Publisher, Ltd., Springfield, Ill.)

Segmental Fields of Sensation—Dermatomes

Each spinal nerve innervates a "segmental field" of the skin called a *dermatome.* The different dermatomes are shown in Figure 47-14. They are shown in the figure as if there were distinct borders between the adjacent dermatomes, which is far from true because much overlap exists from segment to segment.

The figure shows that the anal region of the body lies in the dermatome of the most distal cord segment, dermatome S5. In the embryo, this is the tail region and the most distal portion of the body. The legs originate embryologically from the lumbar and upper sacral segments (L2 to S3), rather than from the distal sacral segments, which is evident from the dermatomal map. One can use a dermatomal map as shown in Figure 47-14 to determine the level in the spinal cord at which a cord injury has occurred when the peripheral sensations are disturbed by the injury.

Bibliography

Alonso JM, Swadlow HA: Thalamocortical specificity and the synthesis of sensory cortical receptive fields, *J Neurophysiol* 94:26, 2005.

Baker SN: Oscillatory interactions between sensorimotor cortex and the periphery, *Curr Opin Neurobiol* 17:649, 2007.

Bosco G, Poppele RE: Proprioception from a spinocerebellar perspective, *Physiol Rev* 81:539, 2001.

Chalfie M: Neurosensory mechanotransduction, *Nat Rev Mol Cell Biol* 10:44, 2009.

Cohen YE, Andersen RA: A common reference frame for movement plans in the posterior parietal cortex, *Nat Rev Neurosci* 3:553, 2002.

Craig AD: Pain mechanisms: labeled lines versus convergence in central processing, *Annu Rev Neurosci* 26:1, 2003.

Fontanini A, Katz DB: Behavioral states, network states, and sensory response variability, *J Neurophysiol* 100:1160, 2008.

Fox K: Experience-dependent plasticity mechanisms for neural rehabilitation in somatosensory cortex, *Philos Trans R Soc Lond B Biol Sci* 364:369, 2009.

Haines DE: *Fundamental Neuroscience for Basic and Clinical Applications*, ed 3, Philadelphia, 2006, Churchill Livingstone, Elsevier.

Hsiao S: Central mechanisms of tactile shape perception, *Curr Opin Neurobiol* 18:418, 2008.

Johansson RS, Flanagan JR: Coding and use of tactile signals from the fingertips in object manipulation tasks, *Nat Rev Neurosci* 10:345, 2009.

Kaas JH: The evolution of the complex sensory and motor systems of the human brain, *Brain Res Bull* 75:384, 2008.

Kaas JH, Qi HX, Burish MJ, et al: Cortical and subcortical plasticity in the brains of humans, primates, and rats after damage to sensory afferents in the dorsal columns of the spinal cord, *Exp Neurol* 209:407, 2008.

Kandel ER, Schwartz JH, Jessell TM: *Principles of Neural Science*, ed 4, New York, 2000, McGraw-Hill.

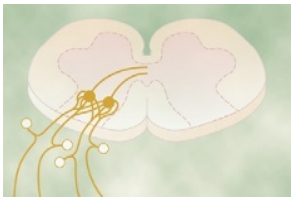
Knutsen PM, Ahissar E: Orthogonal coding of object location, *Trends Neurosci* 32:101, 2009.

Pelli DG, Tillman KA: The uncrowded window of object recognition, *Nat Neurosci* 11:1129, 2008.

Suga N, Ma X: Multiparametric corticofugal modulation and plasticity in the auditory system, *Nat Rev Neurosci* 4:783, 2003.

CHAPTER 48

Somatic Sensations: II. Pain, Headache, and Thermal Sensations



Many, if not most, ailments of the body cause pain. Furthermore, the ability to diagnose different diseases depends to a great extent on a physician's knowledge of the different qualities of pain. For these reasons, the first part of this chapter is devoted mainly to pain and to the physiologic bases of some associated clinical phenomena.

Pain Is a Protective Mechanism. Pain occurs whenever tissues are being damaged, and it causes the individual to react to remove the pain stimulus. Even such simple activities as sitting for a long time on the ischia can cause tissue destruction because of lack of blood flow to the skin where it is compressed by the weight of the body. When the skin becomes painful as a result of the ischemia, the person normally shifts weight subconsciously. But a person who has lost the pain sense, as after spinal cord injury, fails to feel the pain and, therefore, fails to shift. This soon results in total breakdown and desquamation of the skin at the areas of pressure.

Types of Pain and Their Qualities—Fast Pain and Slow Pain

Pain has been classified into two major types: *fast pain* and *slow pain.* Fast pain is felt within about 0.1 second after a pain stimulus is applied, whereas slow pain begins only after 1 second or more and then increases slowly over many seconds and sometimes even minutes. During the course of this chapter, we shall see that the conduction pathways for these two types of pain are different and that each of them has specific qualities.

Fast pain is also described by many alternative names, such as *sharp pain, pricking pain, acute pain,* and *electric pain.* This type of pain is felt when a needle is stuck into the skin, when the skin is cut with a knife, or when the skin is acutely burned. It is also felt when the skin is subjected to electric shock. Fast-sharp pain is not felt in most deeper tissues of the body.

Slow pain also goes by many names, such as *slow burning pain, aching pain, throbbing pain, nauseous pain,* and *chronic pain.* This type of pain is usually associated with *tissue destruction.* It can lead to prolonged, almost unbearable suffering. It can occur both in the skin and in almost any deep tissue or organ.

Pain Receptors and Their Stimulation

Pain Receptors Are Free Nerve Endings. The pain receptors in the skin and other tissues are all free nerve endings. They are widespread in the superficial layers of the *skin,* as well as in certain internal tissues, such as the *periosteum,* the *arterial walls,* the *joint surfaces,* and the *falx* and *tentorium* in the cranial vault. Most other deep tissues are only sparsely supplied with pain endings; nevertheless, any widespread tissue damage can summate to cause the slow-chronic-aching type of pain in most of these areas.

Three Types of Stimuli Excite Pain Receptors—Mechanical, Thermal, and Chemical. Pain can be elicited by multiple types of stimuli. They are classified as *mechanical, thermal,* and *chemical pain stimuli.* In general, fast pain is elicited by the mechanical and thermal types of stimuli, whereas slow pain can be elicited by all three types.

Some of the chemicals that excite the chemical type of pain are *bradykinin, serotonin, histamine, potassium ions, acids, acetylcholine,* and *proteolytic enzymes.* In addition, *prostaglandins* and *substance P* enhance the sensitivity of pain endings but do not directly excite them. The chemical substances are especially important in stimulating the slow, suffering type of pain that occurs after tissue injury.

Nonadapting Nature of Pain Receptors. In contrast to most other sensory receptors of the body, pain receptors adapt very little and sometimes not at all. In fact, under some conditions, excitation of pain fibers becomes progressively greater, especially so for slow-aching-nauseous pain, as the pain stimulus continues. This increase in sensitivity of the pain receptors is called *hyperalgesia.*

One can readily understand the importance of this failure of pain receptors to adapt because it allows the pain to keep the person apprised of a tissue-damaging stimulus as long as it persists.

Rate of Tissue Damage as a Stimulus for Pain

The average person begins to perceive pain when the skin is heated above 45°C, as shown in Figure 48-1. This is also the temperature at which the tissues begin to be damaged by heat; indeed, the tissues are eventually destroyed if the temperature remains above this level indefinitely. Therefore, it is immediately apparent that pain resulting from heat is closely correlated with the *rate at which damage to the tissues is occurring* and not with the total damage that has already occurred.

The intensity of pain is also closely correlated with the *rate of tissue damage* from causes other than heat, such as bacterial infection, tissue ischemia, tissue contusion, and so forth.

Special Importance of Chemical Pain Stimuli During Tissue Damage. Extracts from damaged tissue cause intense pain when injected beneath the normal skin. Most of the chemicals listed earlier that excite the chemical pain receptors can be found in these extracts. One chemical that seems to be more painful than others is *bradykinin.* Many researchers have suggested that bradykinin might be the agent most responsible for causing pain following tissue damage. Also, the intensity of the pain felt correlates with the local increase in potassium ion concentration or the increase in proteolytic enzymes that directly attack the nerve endings and excite pain by making the nerve membranes more permeable to ions.

Tissue Ischemia as a Cause of Pain. When blood flow to a tissue is blocked, the tissue often becomes very painful within a few minutes. The greater the rate of metabolism of the tissue, the more rapidly the pain appears. For instance, if a blood pressure cuff is placed around the upper arm and inflated until the arterial blood flow ceases, exercise of the forearm muscles sometimes can cause muscle pain within 15 to 20 seconds. In the absence of muscle exercise, the pain may not appear for 3 to 4 minutes even though the muscle blood flow remains zero.

One of the suggested causes of pain during ischemia is accumulation of large amounts of lactic acid in the tissues, formed as a consequence of anaerobic metabolism (metabolism without oxygen). It is also probable that other chemical agents, such as bradykinin and proteolytic enzymes, are formed in the tissues because of cell damage and that these, in addition to lactic acid, stimulate the pain nerve endings.

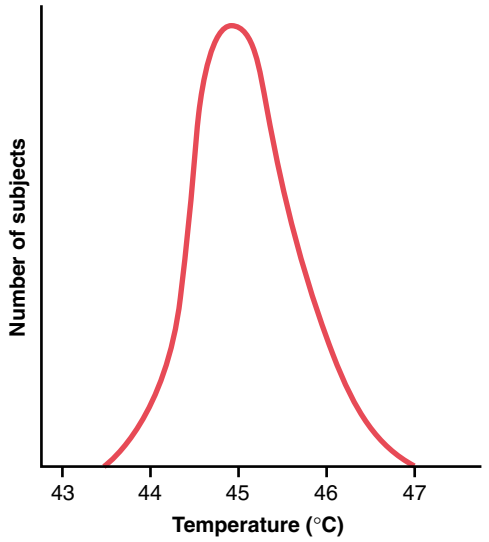


Figure 48-1 Distribution curve obtained from a large number of persons showing the minimal skin temperature that will cause pain. (Modified from Hardy DJ: Nature of pain. J Clin Epidemiol 4:22, 1956.)

Muscle Spasm as a Cause of Pain. Muscle spasm is also a common cause of pain, and it is the basis of many clinical pain syndromes. This pain probably results partially from the direct effect of muscle spasm in stimulating mechanosensitive pain receptors, but it might also result from the indirect effect of muscle spasm to compress the blood vessels and cause ischemia. Also, the spasm increases the rate of metabolism in the muscle tissue, thus making the relative ischemia even greater, creating ideal conditions for the release of chemical pain-inducing substances.

Dual Pathways for Transmission of Pain Signals into the Central Nervous System

Even though all pain receptors are free nerve endings, these endings use two separate pathways for transmitting pain signals into the central nervous system. The two pathways mainly correspond to the two types of pain—a *fast-sharp pain pathway* and a *slow-chronic pain pathway.*

Peripheral Pain Fibers—"Fast" and "Slow" Fibers. The fast-sharp pain signals are elicited by either mechanical or thermal pain stimuli; they are transmitted in the peripheral nerves to the spinal cord by small type Aδ fibers at velocities between 6 and 30 m/sec. Conversely, the slow-chronic type of pain is elicited mostly by chemical types of pain stimuli but sometimes by persisting mechanical or thermal stimuli. This slow-chronic pain is transmitted to the spinal cord by type C fibers at velocities between 0.5 and 2 m/sec.

Because of this double system of pain innervation, a sudden painful stimulus often gives a "double" pain sensation: a fast-sharp pain that is transmitted to the brain by the Aδ fiber pathway, followed a second or so later by a slow pain that is transmitted by the C fiber pathway. The sharp pain apprises the person rapidly of a damaging influence and, therefore, plays an important role in making the person react immediately to remove himself or herself from the stimulus. The slow pain tends to become greater over time. This sensation eventually produces intolerable

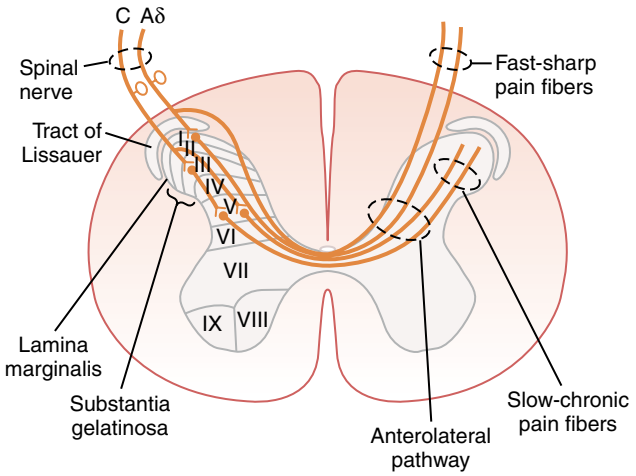


Figure 48-2 Transmission of both "fast-sharp" and "slow-chronic" pain signals into and through the spinal cord on their way to the brain.

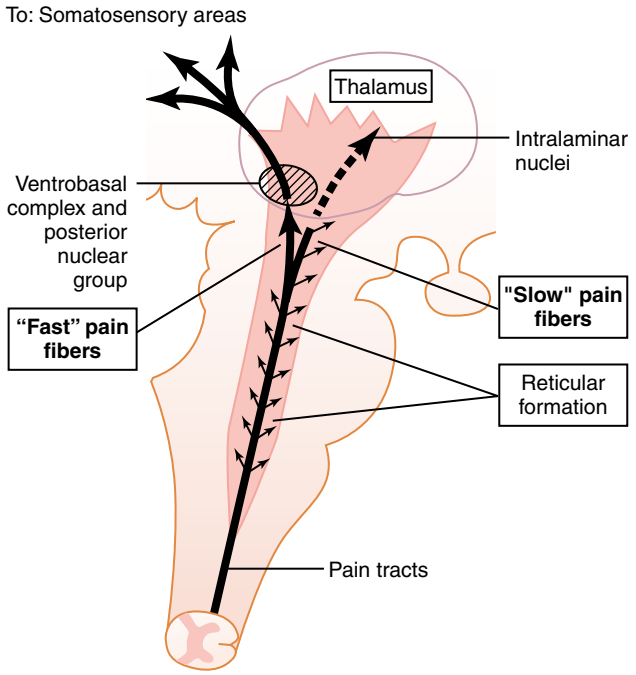


Figure 48-3 Transmission of pain signals into the brain stem, thalamus, and cerebral cortex by way of the *fast pricking pain pathway* and the *slow burning pain pathway.*

pain and makes the person keep trying to relieve the cause of the pain.

On entering the spinal cord from the dorsal spinal roots, the pain fibers terminate on relay neurons in the dorsal horns. Here again, there are two systems for processing the pain signals on their way to the brain, as shown in Figures 48-2 and 48-3.

Dual Pain Pathways in the Cord and Brain Stem—The Neospinothalamic Tract and the Paleospinothalamic Tract

On entering the spinal cord, the pain signals take two pathways to the brain, through (1) the *neospinothalamic tract* and (2) the *paleospinothalamic tract.*

Neospinothalamic Tract for Fast Pain. The fast type Aδ pain fibers transmit mainly mechanical and acute thermal pain. They terminate mainly in lamina I (lamina marginalis) of the dorsal horns, as shown in Figure 48-2, and there excite second-order neurons of the neospinothalamic tract. These give rise to long fibers that cross immediately to the opposite side of the cord through the anterior commissure and then turn upward, passing to the brain in the anterolateral columns.

Termination of the Neospinothalamic Tract in the Brain Stem and Thalamus. A few fibers of the neospinothalamic tract terminate in the reticular areas of the brain stem, but most pass all the way to the thalamus without interruption, terminating in the *ventrobasal complex* along with the dorsal column–medial lemniscal tract for tactile sensations, as was discussed in Chapter 47. A few fibers also terminate in the posterior nuclear group of the thalamus. From these thalamic areas, the signals are transmitted to other basal areas of the brain, as well as to the somatosensory cortex.

Capability of the Nervous System to Localize Fast Pain in the Body. The fast-sharp type of pain can be localized much more exactly in the different parts of the body than can slow-chronic pain. However, when only pain receptors are stimulated, without the simultaneous stimulation of tactile receptors, even fast pain may be poorly localized, often only within 10 centimeters or so of the stimulated area. Yet when tactile receptors that excite the dorsal column–medial lemniscal system are simultaneously stimulated, the localization can be nearly exact.

Glutamate, the Probable Neurotransmitter of the Type Aδ Fast Pain Fibers. It is believed that *glutamate* is the neurotransmitter substance secreted in the spinal cord at the type Aδ pain nerve fiber endings. This is one of the most widely used excitatory transmitters in the central nervous system, usually having a duration of action lasting for only a few milliseconds.

Paleospinothalamic Pathway for Transmitting Slow-Chronic Pain. The paleospinothalamic pathway is a much older system and transmits pain mainly from the peripheral slow-chronic type C pain fibers, although it does transmit some signals from type Aδ fibers as well. In this pathway, the peripheral fibers terminate in the spinal cord almost entirely in laminae II and III of the dorsal horns, which together are called the *substantia gelatinosa,* as shown by the lateral most dorsal root type C fiber in Figure 48-2. Most of the signals then pass through one or more additional short fiber neurons within the dorsal horns themselves before entering mainly lamina V, also in the dorsal horn. Here the last neurons in the series give rise to long axons that mostly join the fibers from the fast pain pathway, passing first through the anterior commissure to the opposite side of the cord, then upward to the brain in the anterolateral pathway.

Substance P, the Probable Slow-Chronic Neurotransmitter of Type C Nerve Endings. Research suggests that type C pain fiber terminals entering the spinal cord release both glutamate transmitter and substance P transmitter. The glutamate transmitter acts instantaneously and lasts for only a few milliseconds. Substance P is released much more slowly, building up in concentration over a period of seconds or even minutes. In fact, it has been suggested that the "double" pain sensation one feels after a pinprick might result partly from the fact that the glutamate transmitter gives a faster pain sensation, whereas the substance P transmitter gives a more lagging sensation. Regardless of the yet unknown details, it seems clear that glutamate is the neurotransmitter most involved in transmitting fast pain into the central nervous system, and substance P is concerned with slow-chronic pain.

Projection of the Paleospinothalamic Pathway (Slow-Chronic Pain Signals) into the Brain Stem and Thalamus. The slow-chronic paleospinothalamic pathway terminates widely in the brain stem, in the large shaded area shown in Figure 48-3. Only one tenth to one fourth of the fibers pass all the way to the thalamus. Instead, most terminate in one of three areas: (1) the *reticular nuclei* of the medulla, pons, and mesencephalon; (2) the *tectal area* of the mesencephalon deep to the superior and inferior colliculi; or (3) the *periaqueductal gray region* surrounding the aqueduct of Sylvius. These lower regions of the brain appear to be important for feeling the suffering types of pain, because animals whose brains have been sectioned above the mesencephalon to block pain signals from reaching the cerebrum still evince undeniable evidence of suffering when any part of the body is traumatized. From the brain stem pain areas, multiple short-fiber neurons relay the pain signals upward into the intralaminar and ventrolateral nuclei of the thalamus and into certain portions of the hypothalamus and other basal regions of the brain.

Very Poor Capability of the Nervous System to Localize Precisely the Source of Pain Transmitted in the Slow-Chronic Pathway. Localization of pain transmitted by way of the paleospinothalamic pathway is imprecise. For instance, slow-chronic pain can usually be localized only to a major part of the body, such as to one arm or leg but not to a specific point on the arm or leg. This is in keeping with the multisynaptic, diffuse connectivity of this pathway. It explains why patients often have serious difficulty in localizing the source of some chronic types of pain.

Function of the Reticular Formation, Thalamus, and Cerebral Cortex in the Appreciation of Pain. Complete removal of the somatic sensory areas of the cerebral cortex does not destroy an animal's ability to perceive pain. Therefore, it is likely that pain impulses entering the brain stem reticular formation, the thalamus, and other lower brain centers cause conscious perception of pain. This does not mean that the cerebral cortex has nothing to do with normal pain appreciation; electrical stimulation of cortical somatosensory areas does cause a human being to perceive mild pain from about 3 percent of the points stimulated. However, it is believed that the cortex plays an especially important role in interpreting pain quality, even though pain perception might be principally the function of lower centers.

Special Capability of Pain Signals to Arouse Overall Brain Excitability. Electrical stimulation in the *reticular areas of the brain stem* and in the *intralaminar nuclei of the thalamus,* the areas where the slow-suffering type of pain terminates, has a strong arousal effect on nervous activity throughout the entire brain. In fact, these two areas constitute part of the brain's principal "arousal system," which is discussed in Chapter 59. This explains why it is almost impossible for a person to sleep when he or she is in severe pain.

Surgical Interruption of Pain Pathways. When a person has severe and intractable pain (sometimes resulting from rapidly spreading cancer), it is necessary to relieve the pain. To do this, the pain nervous pathways can be cut at any one of several points. If the pain is in the lower part of the body, a *cordotomy* in the thoracic region of the spinal cord often relieves the pain for a few weeks to a few months. To do this, the spinal cord on the side opposite to the pain is partially cut in its *anterolateral quadrant* to interrupt the anterolateral sensory pathway.

A cordotomy, however, is not always successful in relieving pain, for two reasons. First, many pain fibers from the upper part of the body do not cross to the opposite side of the spinal cord until they have reached the brain, so the cordotomy does not transect these fibers. Second, pain frequently returns several months later, partly as a result of sensitization of other pathways that normally are too weak to be effectual (e.g., sparse pathways in the dorsolateral cord). Another experimental operative procedure to relieve pain has been to cauterize specific pain areas in the intralaminar nuclei in the thalamus, which often relieves suffering types of pain while leaving intact one's appreciation of "acute" pain, an important protective mechanism.

Pain Suppression ("Analgesia") System in the Brain and Spinal Cord

The degree to which a person reacts to pain varies tremendously. This results partly from a capability of the brain itself to suppress input of pain signals to the nervous system by activating a pain control system, called an *analgesia system.*

The analgesia system is shown in Figure 48-4. It consists of three major components: (1) The *periaqueductal gray* and *periventricular areas* of the mesencephalon and upper pons surround the aqueduct of Sylvius and

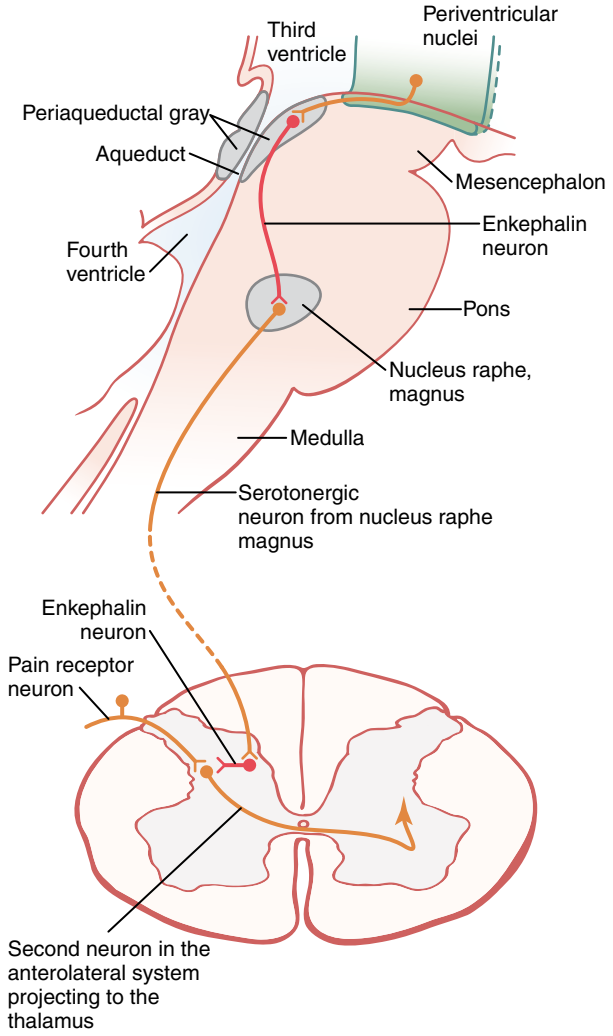


Figure 48-4 Analgesia system of the brain and spinal cord, showing (1) inhibition of incoming pain signals at the cord level and (2) presence of *enkephalin-secreting neurons* that suppress pain signals in both the cord and the brain stem.

portions of the third and fourth ventricles. Neurons from these areas send signals to (2) the *raphe magnus nucleus,* a thin midline nucleus located in the lower pons and upper medulla, and the *nucleus reticularis paragigantocellularis,* located laterally in the medulla. From these nuclei, second-order signals are transmitted down the dorsolateral columns in the spinal cord to (3) a *pain inhibitory complex located in the dorsal horns of the spinal cord.* At this point, the analgesia signals can block the pain before it is relayed to the brain.

Electrical stimulation either in the periaqueductal gray area or in the raphe magnus nucleus can suppress many strong pain signals entering by way of the dorsal spinal roots. Also, stimulation of areas at still higher levels of the brain that excite the periaqueductal gray area can also suppress pain. Some of these areas are (1) the *periventricular nuclei in the hypothalamus,* lying adjacent to the third ventricle, and (2) to a lesser extent, the *medial forebrain bundle,* also in the hypothalamus.

Several transmitter substances are involved in the analgesia system; especially involved are *enkephalin* and *serotonin.* Many nerve fibers derived from the periventricular nuclei and from the periaqueductal gray area secrete enkephalin at their endings. Thus, as shown in Figure 48-4, the endings of many fibers in the raphe magnus nucleus release enkephalin when stimulated.

Fibers originating in this area send signals to the dorsal horns of the spinal cord to secrete serotonin at their endings. The serotonin causes local cord neurons to secrete enkephalin as well. The enkephalin is believed to cause both *presynaptic* and *postsynaptic inhibition* of incoming type C and type Aδ pain fibers where they synapse in the dorsal horns.

Thus, the analgesia system can block pain signals at the initial entry point to the spinal cord. In fact, it can also block many local cord reflexes that result from pain signals, especially withdrawal reflexes described in Chapter 54.

Brain's Opiate System—Endorphins and Enkephalins

More than 40 years ago it was discovered that injection of minute quantities of morphine either into the periventricular nucleus around the third ventricle or into the periaqueductal gray area of the brain stem causes an extreme degree of analgesia. In subsequent studies, it has been found that morphine-like agents, mainly the opiates, also act at many other points in the analgesia system, including the dorsal horns of the spinal cord. Because most drugs that alter excitability of neurons do so by acting on synaptic receptors, it was assumed that the "morphine receptors" of the analgesia system must be receptors for some morphine-like neurotransmitter that is naturally secreted in the brain. Therefore, an extensive search was undertaken for the natural opiate of the brain. About a dozen such opiate-like substances have now been found at different points of the nervous system; all are breakdown products of three large protein molecules: *pro-opiomelanocortin, proenkephalin,* and *prodynorphin.* Among the more important of these opiate-like substances are *β-endorphin, met-enkephalin, leu-enkephalin*, and *dynorphin.*

The two enkephalins are found in the brain stem and spinal cord, in the portions of the analgesia system described earlier, and β-endorphin is present in both the hypothalamus and the pituitary gland. Dynorphin is found mainly in the same areas as the enkephalins, but in much lower quantities.

Thus, although the fine details of the brain's opiate system are not understood, *activation of the analgesia system* by nervous signals entering the periaqueductal gray and periventricular areas, or *inactivation of pain pathways* by morphine-like drugs, can almost totally suppress many pain signals entering through the peripheral nerves.

Inhibition of Pain Transmission by Simultaneous Tactile Sensory Signals

Another important event in the saga of pain control was the discovery that stimulation of large-type Aβ sensory fibers from peripheral tactile receptors can depress transmission

of pain signals from the same body area. This presumably results from local lateral inhibition in the spinal cord. It explains why such simple maneuvers as rubbing the skin near painful areas is often effective in relieving pain. And it probably also explains why liniments are often useful for pain relief.

This mechanism and the simultaneous psychogenic excitation of the central analgesia system are probably also the basis of pain relief by *acupuncture.*

Treatment of Pain by Electrical Stimulation

Several clinical procedures have been developed for suppressing pain by electrical stimulation. Stimulating electrodes are placed on selected areas of the skin or, on occasion, implanted over the spinal cord, supposedly to stimulate the dorsal sensory columns.

In some patients, electrodes have been placed stereotaxically in appropriate intralaminar nuclei of the thalamus or in the periventricular or periaqueductal area of the diencephalon. The patient can then personally control the degree of stimulation. Dramatic relief has been reported in some instances. Also, pain relief has been reported to last for as long as 24 hours after only a few minutes of stimulation.

Referred Pain

Often a person feels pain in a part of the body that is fairly remote from the tissue causing the pain. This is called *referred pain.* For instance, pain in one of the visceral organs often is referred to an area on the body surface. Knowledge of the different types of referred pain is important in clinical diagnosis because in many visceral ailments the only clinical sign is referred pain.

Mechanism of Referred Pain. Figure 48-5 shows the probable mechanism by which most pain is referred. In the figure, branches of visceral pain fibers are shown to synapse in the spinal cord on the same second-order neurons (1 and 2) that receive pain signals from the skin. When the visceral pain fibers are stimulated, pain signals from the viscera are conducted through at least some of the same neurons that conduct pain signals from the skin, and the person has the feeling that the sensations originate in the skin itself.

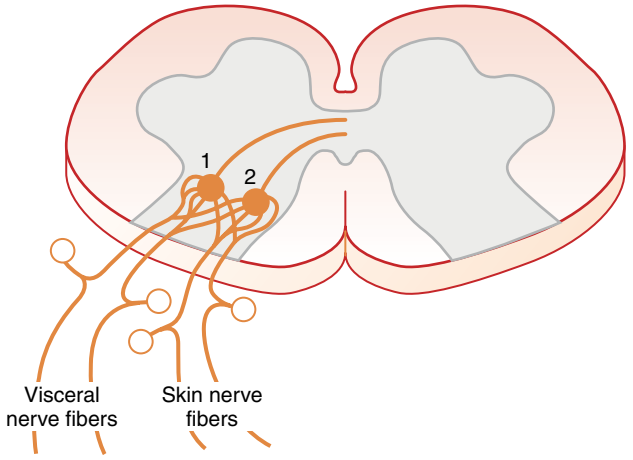


Figure 48-5 Mechanism of referred pain and referred hyperalgesia.

Visceral Pain

Pain from the different viscera of the abdomen and chest is one of the few criteria that can be used for diagnosing visceral inflammation, visceral infectious disease, and other visceral ailments. Often, the viscera have sensory receptors for no other modalities of sensation besides pain. Also, visceral pain differs from surface pain in several important aspects.

One of the most important differences between surface pain and visceral pain is that highly localized types of damage to the viscera seldom cause severe pain. For instance, a surgeon can cut the gut entirely in two in a patient who is awake without causing significant pain. Conversely, any stimulus that causes *diffuse stimulation of pain nerve endings* throughout a viscus causes pain that can be severe. For instance, ischemia caused by occluding the blood supply to a large area of gut stimulates many diffuse pain fibers at the same time and can result in extreme pain.

Causes of True Visceral Pain

Any stimulus that excites pain nerve endings in diffuse areas of the viscera can cause visceral pain. Such stimuli include ischemia of visceral tissue, chemical damage to the surfaces of the viscera, spasm of the smooth muscle of a hollow viscus, excess distention of a hollow viscus, and stretching of the connective tissue surrounding or within the viscus. Essentially all visceral pain that originates in the thoracic and abdominal cavities is transmitted through small type C pain fibers and, therefore, can transmit only the chronic-aching-suffering type of pain.

Ischemia. Ischemia causes visceral pain in the same way that it does in other tissues, presumably because of the formation of acidic metabolic end products or tissue-degenerative products such as bradykinin, proteolytic enzymes, or others that stimulate pain nerve endings.

Chemical Stimuli. On occasion, damaging substances leak from the gastrointestinal tract into the peritoneal cavity. For instance, proteolytic acidic gastric juice may leak through a ruptured gastric or duodenal ulcer. This juice causes widespread digestion of the visceral peritoneum, thus stimulating broad areas of pain fibers. The pain is usually excruciatingly severe.

Spasm of a Hollow Viscus. Spasm of a portion of the gut, the gallbladder, a bile duct, a ureter, or any other hollow viscus can cause pain, possibly by mechanical stimulation of the pain nerve endings. Or the spasm might cause diminished blood flow to the muscle, combined with the muscle's increased metabolic need for nutrients, thus causing severe pain.

Often pain from a spastic viscus occurs in the form of *cramps,* with the pain increasing to a high degree of severity and then subsiding. This process continues intermittently, once every few minutes. The intermittent cycles result from periods of contraction of smooth muscle. For instance, each time a peristaltic wave travels along an overly excitable spastic gut, a cramp occurs. The cramping type of pain frequently occurs in appendicitis, gastroenteritis, constipation,

menstruation, parturition, gallbladder disease, or ureteral obstruction.

Overdistention of a Hollow Viscus. Extreme overfilling of a hollow viscus also can result in pain, presumably because of overstretch of the tissues themselves. Overdistention can also collapse the blood vessels that encircle the viscus or that pass into its wall, thus perhaps promoting ischemic pain.

Insensitive Viscera. A few visceral areas are almost completely insensitive to pain of any type. These include the parenchyma of the liver and the alveoli of the lungs. Yet the liver *capsule* is extremely sensitive to both direct trauma and stretch, and the *bile ducts* are also sensitive to pain. In the lungs, even though the alveoli are insensitive, both the *bronchi* and the *parietal pleura* are very sensitive to pain.

"Parietal Pain" Caused by Visceral Disease

When a disease affects a viscus, the disease process often spreads to the parietal peritoneum, pleura, or pericardium. These parietal surfaces, like the skin, are supplied with extensive pain innervation from the peripheral spinal nerves. Therefore, pain from the parietal wall overlying a viscus is frequently sharp. An example can emphasize the difference between this pain and true visceral pain: a knife incision through the *parietal* peritoneum is very painful, whereas a similar cut through the visceral peritoneum or through a gut wall is not very painful, if painful at all.

Localization of Visceral Pain—"Visceral" and the "Parietal" Pain Transmission Pathways

Pain from the different viscera is frequently difficult to localize, for a number of reasons. First, the patient's brain does not know from firsthand experience that the different internal organs exist; therefore, any pain that originates internally can be localized only generally. Second, sensations from the abdomen and thorax are transmitted through two pathways to the central nervous system—the *true visceral pathway* and the *parietal pathway.* True visceral pain is transmitted via pain sensory fibers within the autonomic nerve bundles, and the sensations are *referred* to surface areas of the body often far from the painful organ. Conversely, parietal sensations are conducted *directly* into local spinal nerves from the parietal peritoneum, pleura, or pericardium, and these sensations are usually *localized directly over the painful area.*

Localization of Referred Pain Transmitted via Visceral Pathways. When visceral pain is referred to the surface of the body, the person generally localizes it in the dermatomal segment from which the visceral organ originated in the embryo, not necessarily where the visceral organ now lies. For instance, the heart originated in the neck and upper thorax, so the heart's visceral pain fibers pass upward along the sympathetic sensory nerves and enter the spinal cord between segments C-3 and T-5. Therefore, as shown in Figure 48-6, pain from the heart is referred to the side of the neck, over the shoulder, over the pectoral muscles, down the arm, and into the substernal area of the upper chest. These are the areas of the body surface that send their own somatosensory nerve fibers into the C-3 to T-5 cord segments. Most frequently, the pain is on the left side rather than on the right because the left side of the heart is much more frequently involved in coronary disease than the right.

The stomach originated approximately from the seventh to ninth thoracic segments of the embryo. Therefore, stomach pain is referred to the anterior epigastrium above the umbilicus, which is the surface area of the body subserved by the seventh through ninth thoracic segments. Figure 48-6 shows several other surface areas to which visceral pain is referred from other organs, representing in general the areas in the embryo from which the respective organs originated.

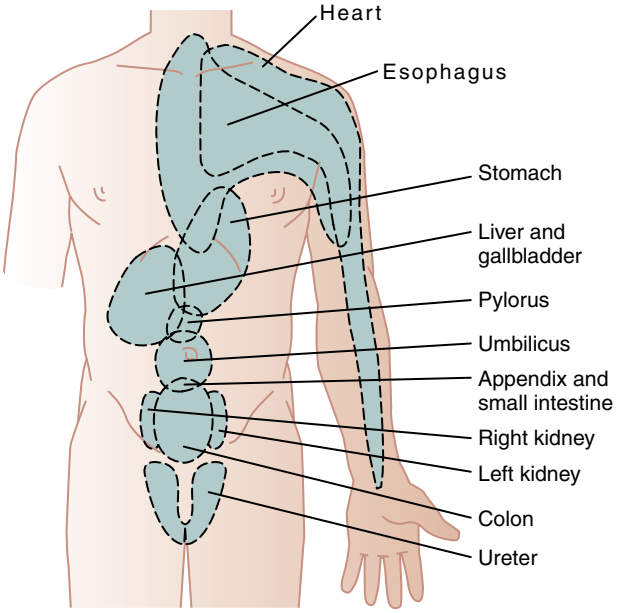


Figure 48-6 Surface areas of referred pain from different visceral organs.

Parietal Pathway for Transmission of Abdominal and Thoracic Pain. Pain from the viscera is frequently localized to two surface areas of the body at the same time because of the dual transmission of pain through the referred visceral pathway and the direct parietal pathway. Thus, Figure 48-7 shows dual transmission from an inflamed appendix. Pain impulses pass first from the appendix through visceral

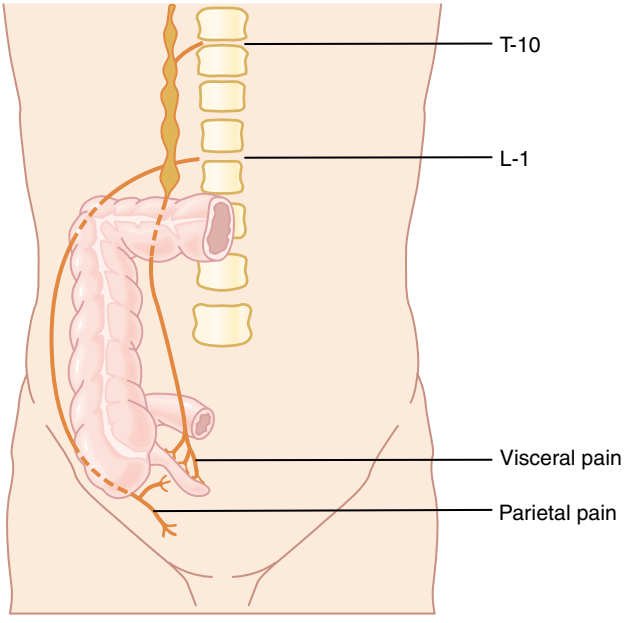


Figure 48-7 Visceral and parietal transmission of pain signals from the appendix.

pain fibers located within sympathetic nerve bundles, and then into the spinal cord at about T-10 or T-11; this pain is referred to an area around the umbilicus and is of the aching, cramping type. Pain impulses also often originate in the parietal peritoneum where the inflamed appendix touches or is adherent to the abdominal wall. These cause pain of the sharp type directly over the irritated peritoneum in the right lower quadrant of the abdomen.

Some Clinical Abnormalities of Pain and Other Somatic Sensations

Hyperalgesia

A pain nervous pathway sometimes becomes excessively excitable; this gives rise to *hyperalgesia,* which means hypersensitivity to pain. Possible causes of hyperalgesia are (1) excessive sensitivity of the pain receptors themselves, which is called *primary hyperalgesia,* and (2) facilitation of sensory transmission, which is called *secondary hyperalgesia.*

An example of primary hyperalgesia is the extreme sensitivity of sunburned skin, which results from sensitization of the skin pain endings by local tissue products from the burn—perhaps histamine, and prostaglandins, and others. Secondary hyperalgesia frequently results from lesions in the spinal cord or the thalamus. Several of these lesions are discussed in subsequent sections.

Herpes Zoster (Shingles)

Occasionally *herpesvirus* infects a dorsal root ganglion. This causes severe pain in the dermatomal segment subserved by the ganglion, thus eliciting a segmental type of pain that circles halfway around the body. The disease is called *herpes zoster,* or "shingles," because of a skin eruption that often ensues.

The cause of the pain is presumably infection of the pain neuronal cells in the dorsal root ganglion by the virus. In addition to causing pain, the virus is carried by neuronal cytoplasmic flow outward through the neuronal peripheral axons to their cutaneous origins. Here the virus causes a rash that vesiculates within a few days and then crusts over within another few days, all of this occurring within the dermatomal area served by the infected dorsal root.

Tic Douloureux

Lancinating pain occasionally occurs in some people over one side of the face in the sensory distribution area (or part of the area) of the fifth or ninth nerves; this phenomenon is called *tic douloureux* (or *trigeminal neuralgia* or *glossopharyngeal neuralgia*). The pain feels like sudden electrical shocks, and it may appear for only a few seconds at a time or may be almost continuous. Often it is set off by exceedingly sensitive trigger areas on the surface of the face, in the mouth, or inside the throat—almost always by a mechanoreceptive stimulus rather than a pain stimulus. For instance, when the patient swallows a bolus of food, as the food touches a tonsil, it might set off a severe lancinating pain in the mandibular portion of the fifth nerve.

The pain of tic douloureux can usually be blocked by surgically cutting the peripheral nerve from the hypersensitive area. The sensory portion of the fifth nerve is often sectioned immediately inside the cranium, where the motor and sensory roots of the fifth nerve separate from each other, so that the motor portions, which are necessary for many jaw movements, can be spared while the sensory elements are destroyed. This operation leaves the side of the face anesthetic, which in itself may be annoying. Furthermore, sometimes the operation is unsuccessful, indicating that the lesion that causes the pain might be in the sensory nucleus in the brain stem and not in the peripheral nerves.

Brown-Séquard Syndrome

If the spinal cord is transected entirely, all sensations and motor functions distal to the segment of transection are blocked, but if the spinal cord is transected on only one side, the *Brown-Séquard syndrome* occurs. The effects of such transection can be predicted from knowledge of the cord fiber tracts shown in Figure 48-8. All motor functions are blocked on the side of the transection in all segments below the level of the transection. Yet only some of the modalities of sensation are lost on the transected side, and others are lost on the opposite side. The sensations of pain, heat, and cold—sensations served by the spinothalamic pathway—are lost *on the opposite side of the body* in all dermatomes two to six segments below the level of the transection. By contrast, the sensations that are transmitted only in the dorsal and dorsolateral columns—kinesthetic and position sensations, vibration sensation, discrete localization, and two-point discrimination—are lost *on the side of the transection* in all dermatomes below the level of the transection. Discrete "light touch" is impaired on the side of the transection because the principal pathway for the transmission of light touch, the dorsal column, is transected. That is, the fibers in this column do not cross to the opposite side until they reach the medulla of the brain. "Crude touch," which is poorly localized, still persists because of partial transmission in the opposite spinothalamic tract.

Headache

Headaches are a type of pain referred to the surface of the head from deep head structures. Some headaches result from pain stimuli arising inside the cranium, but others result from pain arising outside the cranium, such as from the nasal sinuses.

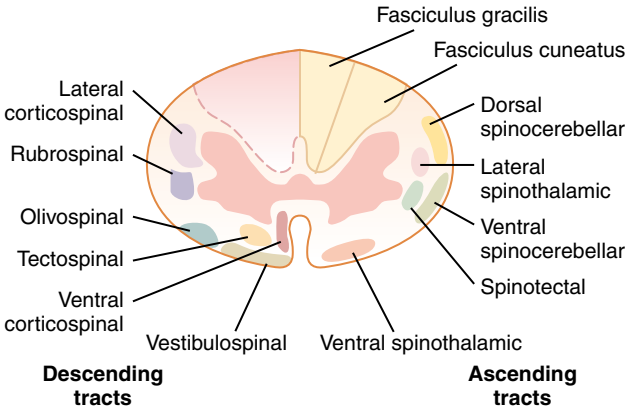


Figure 48-8 Cross section of the spinal cord, showing principal ascending tracts on the right and principal descending tracts on the left.

Headache of Intracranial Origin

Pain-Sensitive Areas in the Cranial Vault. The brain tissues themselves are almost totally insensitive to pain. Even cutting or electrically stimulating the sensory areas of the cerebral cortex only occasionally causes pain; instead, it causes prickly types of paresthesias on the area of the body represented by the portion of the sensory cortex stimulated. Therefore, it is likely that much or most of the pain of headache is not caused by damage within the brain itself.

Conversely, *tugging on the venous sinuses around the brain, damaging the tentorium,* or *stretching the dura at the base of the brain* can cause intense pain that is recognized as headache. Also, almost any type of traumatizing, crushing, or stretching stimulus to the *blood vessels of the meninges* can cause headache. An especially sensitive structure is the middle meningeal artery, and neurosurgeons are careful to anesthetize this artery specifically when performing brain operations under local anesthesia.

Areas of the Head to Which Intracranial Headache Is Referred. Stimulation of pain receptors in the cerebral vault above the tentorium, including the upper surface of the tentorium itself, initiates pain impulses in the cerebral portion of the fifth nerve and, therefore, causes referred headache to the front half of the head in the surface areas supplied by this somatosensory portion of the fifth cranial nerve, as shown in Figure 48-9.

Conversely, pain impulses from beneath the tentorium enter the central nervous system mainly through the glossopharyngeal, vagal, and second cervical nerves, which also supply the scalp above, behind, and slightly below the ear. Subtentorial pain stimuli cause "occipital headache" referred to the posterior part of the head.

Types of Intracranial Headache

Headache of Meningitis. One of the most severe headaches of all is that resulting from meningitis, which causes inflammation of all the meninges, including the sensitive areas of the dura and the sensitive areas around the venous sinuses. Such intense damage can cause extreme headache pain referred over the entire head.

Headache Caused by Low Cerebrospinal Fluid Pressure. Removing as little as 20 milliliters of fluid from the spinal canal, particularly if the person remains in an upright position, often causes intense intracranial headache. Removing this quantity of fluid removes part of the flotation for the brain that is normally provided by the cerebrospinal fluid. The weight of the brain stretches and otherwise distorts the various dural surfaces and thereby elicits the pain that causes the headache.

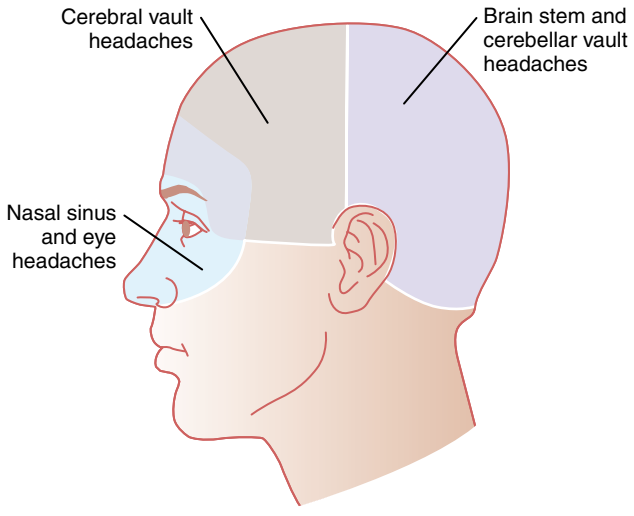


Figure 48-9 Areas of headache resulting from different causes.

Migraine Headache. Migraine headache is a special type of headache that may result from abnormal vascular phenomena, although the exact mechanism is unknown. Migraine headaches often begin with various prodromal sensations, such as nausea, loss of vision in part of the field of vision, visual aura, and other types of sensory hallucinations. Ordinarily, the prodromal symptoms begin 30 minutes to 1 hour before the beginning of the headache. Any theory that explains migraine headache must also explain the prodromal symptoms.

One theory of migraine headaches is that prolonged emotion or tension causes reflex vasospasm of some of the arteries of the head, including arteries that supply the brain. The vasospasm theoretically produces ischemia of portions of the brain, and this is responsible for the prodromal symptoms. Then, as a result of the intense ischemia, something happens to the vascular walls, perhaps exhaustion of smooth muscle contraction, to allow the blood vessels to become flaccid and incapable of maintaining normal vascular tone for 24 to 48 hours. The blood pressure in the vessels causes them to dilate and pulsate intensely, and it is postulated that the excessive stretching of the walls of the arteries—including some extracranial arteries, such as the temporal artery—causes the actual pain of migraine headaches. Other theories of the cause of migraine headaches include spreading cortical depression, psychological abnormalities, and vasospasm caused by excess local potassium in the cerebral extracellular fluid.

There may be a genetic predisposition to migraine headaches because a positive family history for migraine has been reported in 65 to 90 percent of cases. Migraine headaches also occur about twice as frequently in women as in men.

Alcoholic Headache.. As many people have experienced, a headache often follows excessive alcohol consumption. It is likely that alcohol, because it is toxic to tissues, directly irritates the meninges and causes the intracranial pain. Dehydration may also play a role in the "hangover" that follows an alcoholic binge; hydration usually attenuates but does not abolish headache and other symptoms of hangover.

Extracranial Types of Headache

Headache Resulting from Muscle Spasm. Emotional tension often causes many of the muscles of the head, especially those muscles attached to the scalp and the neck muscles attached to the occiput, to become spastic, and it is postulated that this is one of the common causes of headache. The pain of the spastic head muscles supposedly is referred to the overlying areas of the head and gives one the same type of headache as intracranial lesions do.

Headache Caused by Irritation of Nasal and Accessory Nasal Structures. The mucous membranes of the nose and nasal sinuses are sensitive to pain, but not intensely so. Nevertheless, infection or other irritative processes in widespread areas of the nasal structures often summate and cause headache that is referred behind the eyes or, in the case of

frontal sinus infection, to the frontal surfaces of the forehead and scalp, as shown in Figure 48-9. Also, pain from the lower sinuses, such as from the maxillary sinuses, can be felt in the face.

Headache Caused by Eye Disorders. Difficulty in focusing one's eyes clearly may cause excessive contraction of the eye ciliary muscles in an attempt to gain clear vision. Even though these muscles are extremely small, it is believed that tonic contraction of them can cause retro-orbital headache. Also, excessive attempts to focus the eyes can result in reflex spasm in various facial and extraocular muscles, which is a possible cause of headache.

A second type of headache that originates in the eyes occurs when the eyes are exposed to excessive irradiation by light rays, especially ultraviolet light. Looking at the sun or the arc of an arc-welder for even a few seconds may result in headache that lasts from 24 to 48 hours. The headache sometimes results from "actinic" irritation of the conjunctivae, and the pain is referred to the surface of the head or retro-orbitally. However, focusing intense light from an arc or the sun on the retina can also burn the retina, and this could be the cause of the headache.

Thermal Sensations

Thermal Receptors and Their Excitation

The human being can perceive different gradations of cold and heat, from *freezing cold* to *cold* to *cool* to *indifferent* to *warm* to *hot* to *burning hot.*

Thermal gradations are discriminated by at least three types of sensory receptors: cold receptors, warmth receptors, and pain receptors. The pain receptors are stimulated only by extreme degrees of heat or cold and, therefore, are responsible, along with the cold and warmth receptors, for "freezing cold" and "burning hot" sensations.

The cold and warmth receptors are located immediately under the skin at discrete separated *spots.* In most areas of the body, there are 3 to 10 times as many cold spots as warmth spots, and the number in different areas of the body varies from 15 to 25 cold spots per square centimeter in the lips to 3 to 5 cold spots per square centimeter in the finger to less than 1 cold spot per square centimeter in some broad surface areas of the trunk.

Although the existence of distinctive warmth nerve endings is quite certain, on the basis of psychological tests, they have not been identified histologically. They are presumed to be free nerve endings because warmth signals are transmitted mainly over type C nerve fibers at transmission velocities of only 0.4 to 2 m/sec.

A definitive cold receptor, however, has been identified. It is a special, small type Aδ myelinated nerve ending that branches several times, the tips of which protrude into the bottom surfaces of basal epidermal cells. Signals are transmitted from these receptors via type Aδ nerve fibers at velocities of about 20 m/sec. Some cold sensations are believed to be transmitted in type C nerve fibers as well, which suggests that some free nerve endings also might function as cold receptors.

Stimulation of Thermal Receptors—Sensations of Cold, Cool, Indifferent, Warm, and Hot. Figure 48-10 shows the effects of different temperatures on the responses of four types of nerve fibers: (1) a pain fiber stimulated by cold, (2) a cold fiber, (3) a warmth fiber, and (4) a pain fiber stimulated by heat. Note especially that these fibers respond differently at different levels of temperature. For instance, in the *very* cold region, only the cold-pain fibers are stimulated (if the skin becomes even colder so that it nearly freezes or actually does freeze, these fibers cannot be stimulated). As the temperature rises to +10° to 15°C, the cold-pain impulses cease, but the cold receptors begin to be stimulated, reaching peak stimulation at about 24°C and fading out slightly above 40°C. Above about 30°C, the warmth receptors begin to be stimulated, but these also fade out at about 49°C. Finally, at around 45°C, the heat-pain fibers begin to be stimulated by heat and, paradoxically, some of the cold fibers begin to be stimulated again, possibly because of damage to the cold endings caused by the excessive heat.

One can understand from Figure 48-10 that a person determines the different gradations of thermal sensations by the relative degrees of stimulation of the different types of endings. One can also understand why extreme degrees of both cold and heat can be painful and why both these sensations, when intense enough, may give almost the same quality of sensation—that is, freezing cold and burning hot sensations feel almost alike.

Stimulatory Effects of Rising and Falling Temperature—Adaptation of Thermal Receptors. When a cold receptor is suddenly subjected to an abrupt fall in temperature, it becomes strongly stimulated at first, but this stimulation fades rapidly during the first few seconds and progressively more slowly during the next 30 minutes or more. In other words, the receptor "adapts" to a great extent, but never 100 percent.

Thus, it is evident that the thermal senses respond markedly to *changes in temperature,* in addition to being able to respond to steady states of temperature. This means that when the temperature of the skin is actively

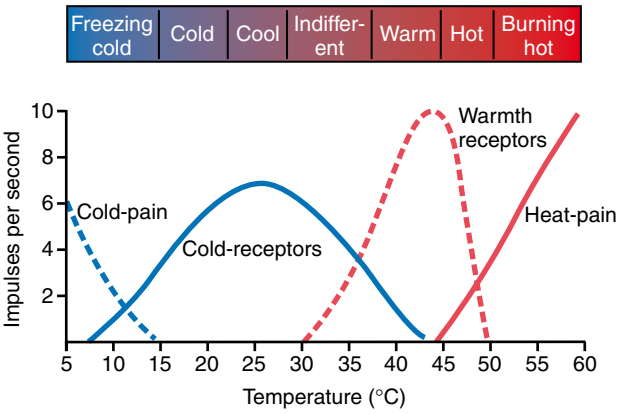


Figure 48-10 Discharge frequencies at different skin temperatures of a *cold-pain fiber,* a *cold fiber,* a *warmth fiber,* and a *heat-pain fiber.*

falling, a person feels much colder than when the temperature remains cold at the same level. Conversely, if the temperature is actively rising, the person feels much warmer than he or she would at the same temperature if it were constant. The response to changes in temperature explains the extreme degree of heat one feels on first entering a tub of hot water and the extreme degree of cold felt on going from a heated room to the out-of-doors on a cold day.

Mechanism of Stimulation of Thermal Receptors

It is believed that the cold and warmth receptors are stimulated by changes in their metabolic rates, and that these changes result from the fact that temperature alters the rate of intracellular chemical reactions more than twofold for each 10°C change. In other words, thermal detection probably results not from direct physical effects of heat or cold on the nerve endings but from chemical stimulation of the endings as modified by temperature.

Spatial Summation of Thermal Sensations. Because the number of cold or warm endings in any one surface area of the body is slight, it is difficult to judge gradations of temperature when small skin areas are stimulated. However, when a large skin area is stimulated all at once, the thermal signals from the entire area summate. For instance, rapid changes in temperature as little as 0.01°C can be detected if this change affects the entire surface of the body simultaneously. Conversely, temperature changes 100 times as great often will not be detected when the affected skin area is only 1 square centimeter in size.

Transmission of Thermal Signals in the Nervous System

In general, thermal signals are transmitted in pathways parallel to those for pain signals. On entering the spinal cord, the signals travel for a few segments upward or downward in the *tract of Lissauer* and then terminate mainly in laminae I, II, and III of the dorsal horns—the same as for pain. After a small amount of processing by one or more cord neurons, the signals enter long, ascending thermal fibers that cross to the opposite anterolateral sensory tract and terminate in both (1) the reticular areas of the brain stem and (2) the ventrobasal complex of the thalamus.

A few thermal signals are also relayed to the cerebral somatic sensory cortex from the ventrobasal complex. Occasionally a neuron in cortical somatic sensory area I has been found by microelectrode studies to be directly responsive to either cold or warm stimuli on a specific area of the skin. However, removal of the entire cortical postcentral gyrus in the human being reduces but does not abolish the ability to distinguish gradations of temperature.

Bibliography

Almeida TF, Roizenblatt S, Tufik S: Afferent pain pathways: a neuroanatomical review, *Brain Res* 1000:40, 2004.

Ballantyne JC, Mao J: Opioid therapy for chronic pain, *N Engl J Med* 349:1943, 2003.

Bandell M, Macpherson LJ, Patapoutian A: From chills to chilis: mechanisms for thermosensation and chemesthesis via thermoTRPs, *Curr Opin Neurobiol* 17:490, 2007.

Benarroch EE: Descending monoaminergic pain modulation: bidirectional control and clinical relevance, *Neurology* 71:217, 2008.

Bingel U, Tracey I: Imaging CNS modulation of pain in humans, *Physiology (Bethesda)* 23:371, 2008.

Borsook D, Becerra L: Pain imaging: future applications to integrative clinical and basic neurobiology, *Adv Drug Deliv Rev* 55:967, 2003.

Bromm B: Brain images of pain, *News Physiol Sci* 16:244, 2001.

Franks NP: General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal, *Nat Rev Neurosci* 9:370, 2008.

Gebhart GF: Descending modulation of pain, *Neurosci Biobehav Rev* 27:729, 2004.

Kandel ER, Schwartz JH, Jessell TM: *Principles of Neural Science*, ed 4, New York, 2000, McGraw-Hill.

Lumpkin EA, Caterina MJ: Mechanisms of sensory transduction in the skin, *Nature* 445:858, 2007.

McKemy DD: Temperature sensing across species, *Pflugers Arch* 454:777, 2007.

Mendell JR, Sahenk Z: Clinical practice: painful sensory neuropathy, *N Engl J Med* 348:1243, 2003.

Milligan ED, Watkins LR: Pathological and protective roles of glia in chronic pain, *Nat Rev Neurosci* 10:23, 2009.

Montell C: Thermosensation: hot findings make TRPNs very cool, *Curr Biol* 13:R476, 2003.

Sanchez-del-Rio M, Reuter U: Migraine aura: new information on underlying mechanisms, *Curr Opin Neurol* 17:289, 2004.

Sandkühler J: Models and mechanisms of hyperalgesia and allodynia, *Physiol Rev* 89:707, 2009.

Schaible HG, Ebersberger A, Von Banchet GS: Mechanisms of pain in arthritis, *Ann N Y Acad Sci* 966:343, 2002.

Schepers RJ, Ringkamp M: Thermoreceptors and thermosensitive afferents, *Neurosci Biobehav Rev* 33:205, 2009.

Silberstein SD: Recent developments in migraine, *Lancet* 372:1369, 2008.

Stein BE, Stanford TR: Multisensory integration: current issues from the perspective of the single neuron, *Nat Rev Neurosci* 9:255, 2008.

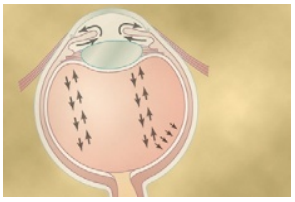
Watkins LR, Maier SF: Beyond neurons: evidence that immune and glial cells contribute to pathological pain states, *Physiol Rev* 82:981, 2002.

White FA, Jung H, Miller RJ: Chemokines and the pathophysiology of neuropathic pain, *Proc Natl Acad Sci U S A* 104:20151, 2007.

Zubrzycka M, Janecka A: Substance P: transmitter of nociception (minireview), *Endocr Regul* 34:195, 2000.

CHAPTER 49

The Eye: I. Optics of Vision



Physical Principles of Optics

Before it is possible to understand the optical system of the eye, the student must first be thoroughly familiar with the basic principles of optics, including the physics of light refraction, focusing, depth of focus, and so forth. A brief review of these physical principles is presented; then the optics of the eye is discussed.

Refraction of Light

Refractive Index of a Transparent Substance. Light rays travel through air at a velocity of about 300,000 km/sec, but they travel much slower through transparent solids and liquids. The refractive index of a transparent substance is the *ratio* of the velocity of light in air to the velocity in the substance. The refractive index of air itself is 1.00. Thus, if light travels through a particular type of glass at a velocity of 200,000 km/sec, the refractive index of this glass is 300,000 divided by 200,000, or 1.50.

Refraction of Light Rays at an Interface Between Two Media with Different Refractive Indices. When light rays traveling forward in a beam (as shown in Figure 49-1*A*) strike an interface that is *perpendicular* to the beam, the rays enter the second medium without deviating from their course. The only effect that occurs is decreased velocity of transmission and shorter wavelength, as shown in the figure by the shorter distances between wave fronts.

If the light rays pass through an angulated interface as shown in Figure 49-1*B*, the rays bend if the refractive indices of the two media are different from each other. In this particular figure, the light rays are leaving air, which has a refractive index of 1.00, and are entering a block of glass having a refractive index of 1.50. When the beam first strikes the angulated interface, the lower edge of the beam enters the glass ahead of the upper edge. The wave front in the upper portion of the beam continues to travel at a velocity of 300,000 km/sec, while that which entered the glass travels at a velocity of 200,000 km/sec. This causes the upper portion of the wave front to move ahead of the lower portion so that the wave front is no longer vertical but angulated to the right. Because *the direction in which light travels is always perpendicular to the plane of the wave front,* the direction of travel of the light beam bends downward.

This bending of light rays at an angulated interface is known as *refraction.* Note particularly that the degree of refraction increases as a function of (1) the ratio of the two refractive indices of the two transparent media and (2) the degree of angulation between the interface and the entering wave front.

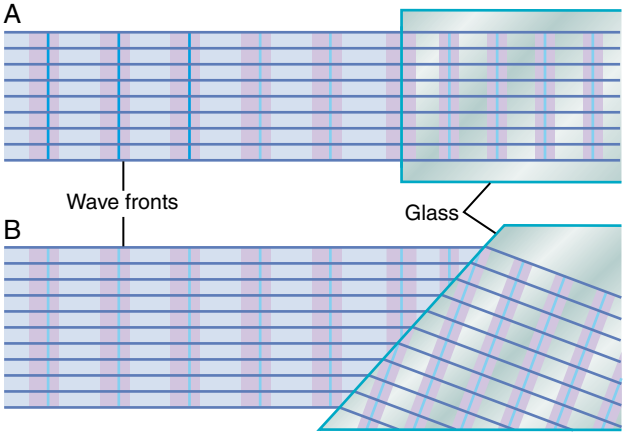


Figure 49-1 Light rays entering a glass surface perpendicular to the light rays (*A*) and a glass surface angulated to the light rays (*B*). This figure demonstrates that the distance between waves after they enter the glass is shortened to about two-thirds that in air. It also shows that light rays striking an angulated glass surface are bent.

Application of Refractive Principles to Lenses

Convex Lens Focuses Light Rays. Figure 49-2 shows parallel light rays entering a convex lens. The light rays passing through the center of the lens strike the lens exactly perpendicular to the lens surface and, therefore, pass through the lens without being refracted. Toward either edge of the lens,

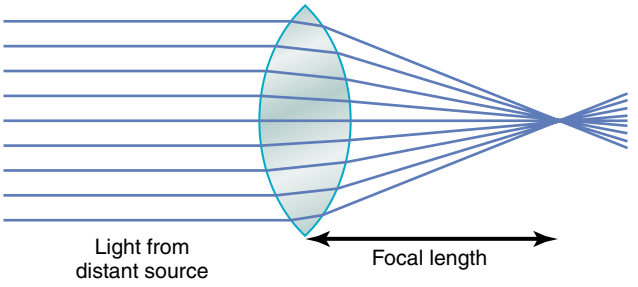


Figure 49-2 Bending of light rays at each surface of a convex spherical lens, showing that parallel light rays are focused to a *focal point.*

however, the light rays strike a progressively more angulated interface. The outer rays bend more and more toward the center, which is called *convergence* of the rays. Half the bending occurs when the rays enter the lens, and half as they exit from the opposite side. If the lens has exactly the proper curvature, parallel light rays passing through each part of the lens will be bent exactly enough so that all the rays will pass through a single point, which is called the *focal point.*

Concave Lens Diverges Light Rays. Figure 49-3 shows the effect of a concave lens on parallel light rays. The rays that enter the center of the lens strike an interface that is perpendicular to the beam and, therefore, do not refract. The rays at the edge of the lens enter the lens ahead of the rays in the center. This is opposite to the effect in the convex lens, and it causes the peripheral light rays to *diverge* from the light rays that pass through the center of the lens. Thus, the concave lens *diverges* light rays, but the convex lens *converges* light rays.

Cylindrical Lens Bends Light Rays in Only One Plane—Comparison with Spherical Lenses. Figure 49-4 shows both a convex *spherical* lens and a convex *cylindrical* lens. Note that the cylindrical lens bends light rays from the two sides of the lens but not from the top or the bottom. That is, bending occurs in one plane but not the other. Thus, parallel light rays are bent to a *focal line.* Conversely, light rays that pass through the spherical lens are refracted at all edges of the lens (in both planes) toward the central ray, and all the rays come to a *focal point.*

The cylindrical lens is well demonstrated by a test tube full of water. If the test tube is placed in a beam of sunlight and a piece of paper is brought progressively closer to the opposite side of the tube, a certain distance will be found at which the light rays come to a *focal line.* The spherical lens is demonstrated by an ordinary magnifying glass. If such a lens is placed in a beam of sunlight and a piece of paper is brought progressively closer to the lens, the light rays will impinge on a common focal point at an appropriate distance.

Concave cylindrical lenses *diverge* light rays in only one plane in the same manner that *convex* cylindrical lenses *converge* light rays in one plane.

Combination of Two Cylindrical Lenses at Right Angles Equals a Spherical Lens. Figure 49-5*B* shows two convex cylindrical lenses at right angles to each other. The vertical cylindrical lens converges the light rays that pass through the two sides of the lens, and the horizontal lens converges the top and bottom rays. Thus, all the light rays come to a single-point focus. In other words, *two cylindrical lenses crossed at right angles to each other perform the same function as one spherical lens of the same refractive power.*

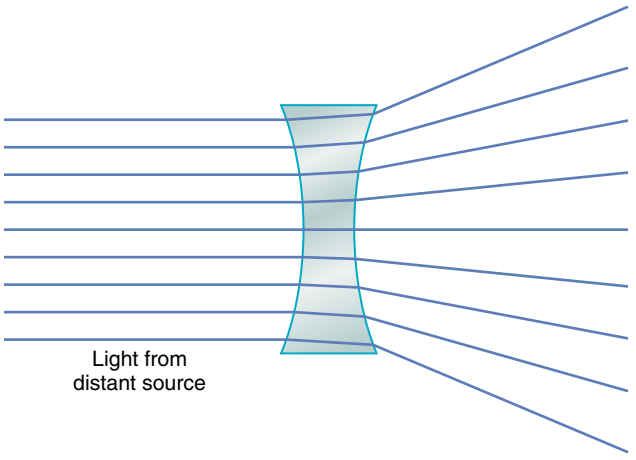


Figure 49-3 Bending of light rays at each surface of a concave spherical lens, showing that parallel light rays are *diverged.*

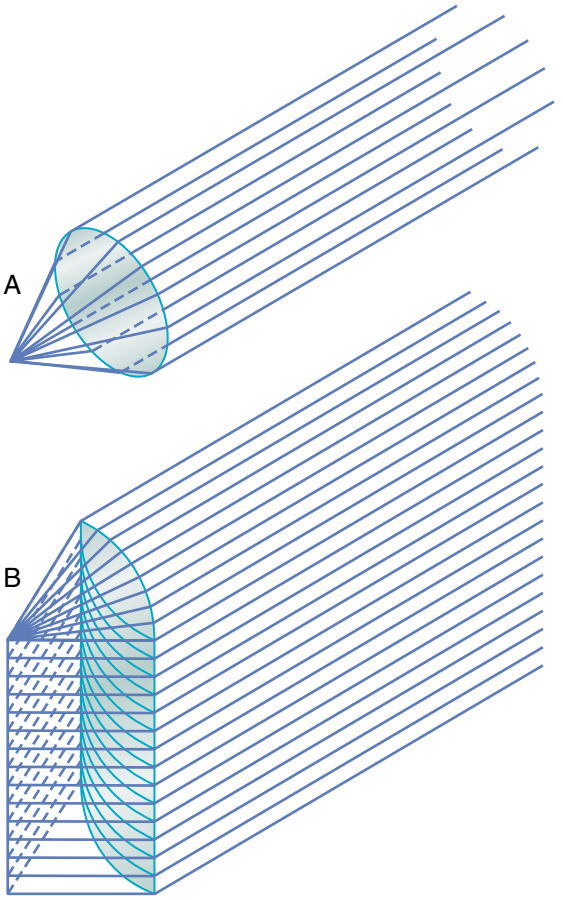


Figure 49-4 *A, Point focus* of parallel light rays by a spherical convex lens. *B, Line focus* of parallel light rays by a cylindrical convex lens.

Focal Length of a Lens

The distance beyond a convex lens at which *parallel* rays converge to a common focal point is called the *focal length* of the lens. The diagram at the top of Figure 49-6 demonstrates this focusing of parallel light rays.

In the middle diagram, the light rays that enter the convex lens are not parallel but are *diverging* because the origin of the light is a point source not far away from the lens itself. Because these rays are diverging outward from the point source, it can be seen from the diagram that they do not focus at the same distance away from the lens as do parallel rays. In other words, when rays of light that are already diverging enter a convex lens, the distance of focus on the other side of the lens is farther from the lens than is the focal length of the lens for parallel rays.

The bottom diagram of Figure 49-6 shows light rays that are diverging toward a convex lens that has far greater curvature than that of the other two lenses in the figure. In this diagram, the distance from the lens at which the light rays come to focus is exactly the same as that from the lens in the first diagram, in which the lens is less convex but the rays entering it are parallel. This demonstrates that both parallel rays and

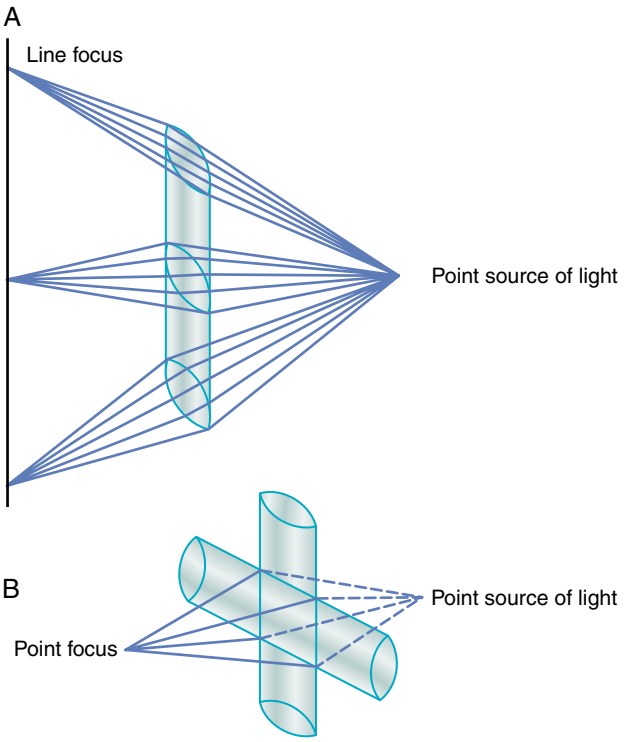


Figure 49-5 *A,* Focusing of light from a point source to a line focus by a cylindrical lens. *B,* Two cylindrical convex lenses at right angles to each other, demonstrating that one lens converges light rays in one plane and the other lens converges light rays in the plane at a right angle. The two lenses combined give the same point focus as that obtained with a single spherical convex lens.

diverging rays can be focused at the same distance beyond a lens, provided the lens changes its convexity.

The relation of focal length of the lens, distance of the point source of light, and distance of focus is expressed by the following formula:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b}$$

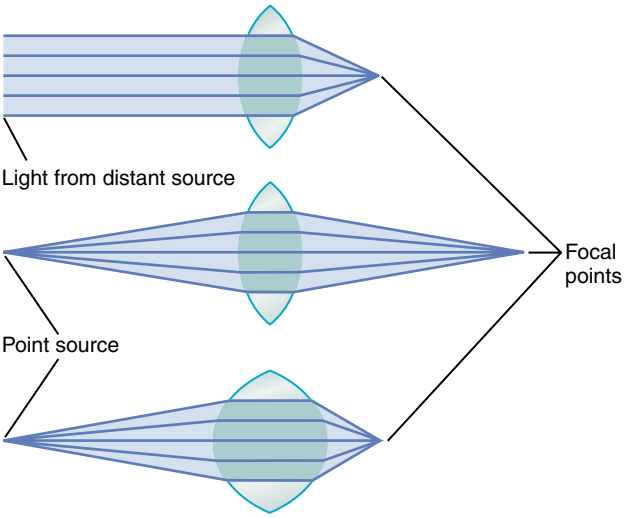


Figure 49-6 The two upper lenses of this figure have the same focal length, but the light rays entering the top lens are parallel, whereas those entering the middle lens are diverging; the effect of parallel versus diverging rays on the focal distance is shown. The bottom lens has far more refractive power than either of the other two lenses (i.e., has a much shorter focal length), demonstrating that the stronger the lens is, the nearer to the lens the point focus is.

in which f is the focal length of the lens for parallel rays, a is the distance of the point source of light from the lens, and b is the distance of focus on the other side of the lens.

Formation of an Image by a Convex Lens

Figure 49-7*A* shows a convex lens with two point sources of light to the left. Because light rays pass through the center of a convex lens without being refracted in either direction, the light rays from each point source of light are shown to come to a point focus on the opposite side of the lens *directly in line with the point source and the center of the lens.*

Any object in front of the lens is, in reality, a mosaic of point sources of light. Some of these points are very bright,

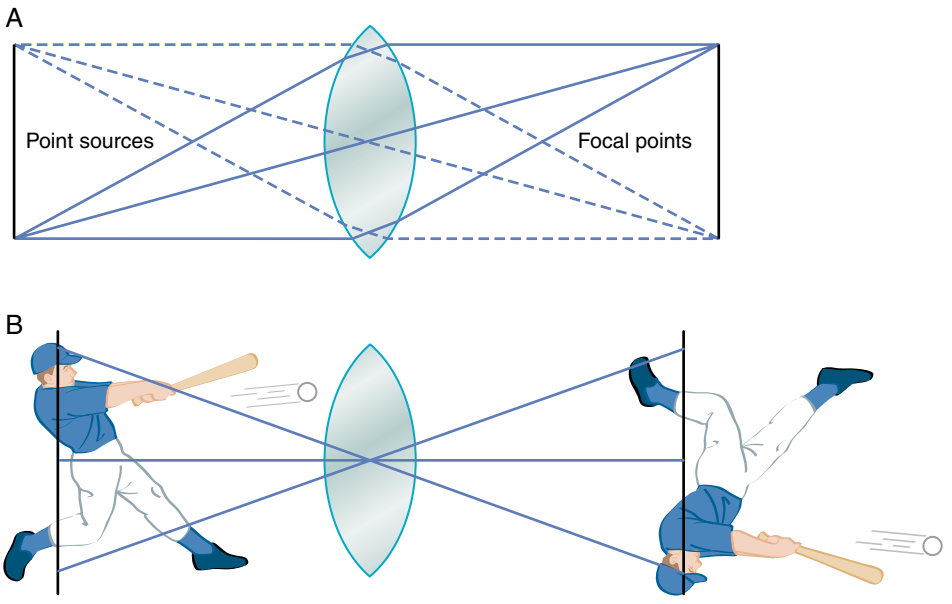


Figure 49-7 *A,* Two point sources of light focused at two separate points on opposite sides of the lens. *B,* Formation of an image by a convex spherical lens.

some are very weak, and they vary in color. Each point source of light on the object comes to a separate point focus on the opposite side of the lens in line with the lens center. If a white sheet of paper is placed at the focus distance from the lens, one can see an image of the object, as demonstrated in Figure 49-7*B*. However, this image is upside down with respect to the original object, and the two lateral sides of the image are reversed. This is the method by which the lens of a camera focuses images on film.

Measurement of the Refractive Power of a Lens—"Diopter"

The more a lens bends light rays, the greater is its "refractive power." This refractive power is measured in terms of *diopters*. The refractive power in diopters of a convex lens is equal to 1 meter divided by its focal length. Thus, a spherical lens that converges parallel light rays to a focal point 1 meter beyond the lens has a refractive power of +1 diopter, as shown in Figure 49-8. If the lens is capable of bending parallel light rays twice as much as a lens with a power of +1 diopter, it is said to have a strength of +2 diopters, and the light rays come to a focal point 0.5 meter beyond the lens. A lens capable of converging parallel light rays to a focal point only 10 centimeters (0.10 meter) beyond the lens has a refractive power of +10 diopters.

The refractive power of concave lenses cannot be stated in terms of the focal distance beyond the lens because the light rays diverge, rather than focus to a point. However, if a concave lens diverges light rays at the same rate that a 1-diopter convex lens converges them, the concave lens is said to have a dioptric strength of –1. Likewise, if the concave lens diverges light rays as much as a +10-diopter lens converges them, this lens is said to have a strength of –10 diopters.

Concave lenses "neutralize" the refractive power of convex lenses. Thus, placing a 1-diopter concave lens immediately in front of a 1-diopter convex lens results in a lens system with zero refractive power.

The strengths of cylindrical lenses are computed in the same manner as the strengths of spherical lenses, except that the *axis* of the cylindrical lens must be stated in addition to its strength. If a cylindrical lens focuses parallel light rays to a line focus 1 meter beyond the lens, it has a strength of +1 diopter. Conversely, if a cylindrical lens of a concave type *diverges* light rays as much as a +1-diopter cylindrical lens *converges* them, it has a strength of –1 diopter. If the focused line is horizontal, its axis is said to be 0 degrees. If it is vertical, its axis is 90 degrees.

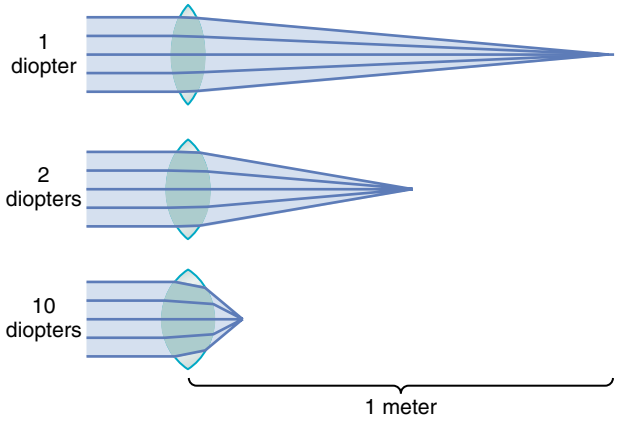


Figure 49-8 Effect of lens strength on the focal distance.

Optics of the Eye

The Eye as a Camera

The eye, shown in Figure 49-9, is optically equivalent to the usual photographic camera. It has a lens system, a variable aperture system (the pupil), and a retina that corresponds to the film. The lens system of the eye is composed of four refractive interfaces: (1) the interface between air and the anterior surface of the cornea, (2) the interface between the posterior surface of the cornea and the aqueous humor, (3) the interface between the aqueous humor and the anterior surface of the lens of the eye, and (4) the interface between the posterior surface of the lens and the vitreous humor. The internal index of air is 1; the cornea, 1.38; the aqueous humor, 1.33; the crystalline lens (on average), 1.40; and the vitreous humor, 1.34.

Consideration of All Refractive Surfaces of the Eye as a Single Lens—The "Reduced" Eye. If all the refractive surfaces of the eye are algebraically added together and then considered to be one single lens, the optics of the normal eye may be simplified and represented schematically as a "reduced eye." This is useful in simple calculations. In the reduced eye, a single refractive surface is considered to exist, with its central point 17 millimeters in front of the retina and a total refractive power of 59 diopters when the lens is accommodated for distant vision.

About two thirds of the 59 diopters of refractive power of the eye is provided by the anterior surface of the cornea (*not* by the eye lens). The principal reason for this is that the refractive index of the cornea is markedly different from that of air, whereas the refractive index of the eye lens is not greatly different from the indices of the aqueous humor and vitreous humor.

The total refractive power of the internal lens of the eye, as it normally lies in the eye surrounded by fluid on each side, is only 20 diopters, about one-third the total refractive power of the eye. But the importance of the internal lens is that, in response to nervous signals from the brain, *its curvature can be increased* markedly to provide "accommodation," which is discussed later in the chapter.

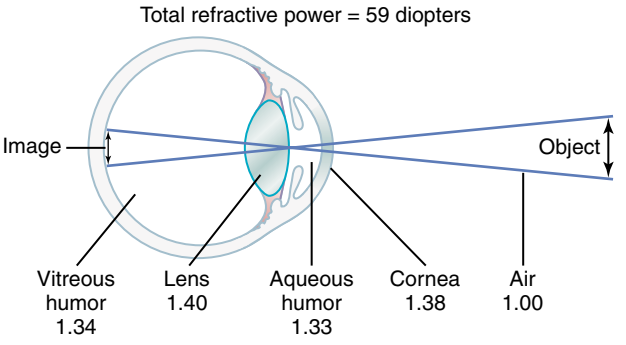


Figure 49-9 The eye as a camera. The numbers are the refractive indices.

Formation of an Image on the Retina. In the same manner that a glass lens can focus an image on a sheet of paper, the lens system of the eye can focus an image on the retina. The image is inverted and reversed with respect to the object. However, the mind perceives objects in the upright position despite the upside-down orientation on the retina because the brain is trained to consider an inverted image as normal.

Mechanism of "Accommodation"

In children, the refractive power of the lens of the eye can be increased voluntarily from 20 diopters to about 34 diopters; this in an "accommodation" of 14 diopters. To do this, the shape of the lens is changed from that of a moderately convex lens to that of a very convex lens. The mechanism is as follows.

In a young person, the lens is composed of a strong elastic capsule filled with viscous, proteinaceous, but transparent fluid. When the lens is in a relaxed state with no tension on its capsule, it assumes an almost spherical shape, owing mainly to the elastic retraction of the lens capsule. However, as shown in Figure 49-10, about 70 *suspensory ligaments* attach radially around the lens, pulling the lens edges toward the outer circle of the eyeball. These ligaments are constantly tensed by their attachments at the anterior border of the choroid and retina. The tension on the ligaments causes the lens to remain relatively flat under normal conditions of the eye.

However, also located at the lateral attachments of the lens ligaments to the eyeball is the *ciliary muscle,* which itself has two separate sets of smooth muscle fibers—*meridional fibers* and *circular fibers.* The meridional fibers extend from the peripheral ends of the suspensory ligaments to the corneoscleral junction. When these muscle fibers contract, the *peripheral insertions* of the lens ligaments are pulled medially toward the edges of the cornea, thereby releasing the ligaments' tension on the lens. The circular fibers are arranged circularly all the way around the ligament attachments so that when they contract, a sphincter-like action occurs, decreasing the diameter of the circle of ligament attachments; this also allows the ligaments to pull less on the lens capsule.

Thus, contraction of either set of smooth muscle fibers in the ciliary muscle relaxes the ligaments to the lens capsule, and the lens assumes a more spherical shape, like that of a balloon, because of the natural elasticity of the lens capsule.

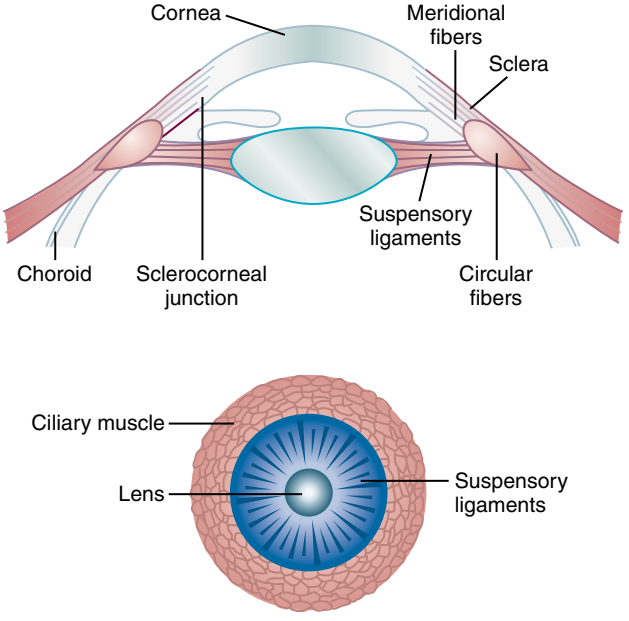


Figure 49-10 Mechanism of accommodation (focusing).

Accommodation Is Controlled by Parasympathetic Nerves. The ciliary muscle is controlled almost entirely by parasympathetic nerve signals transmitted to the eye through the third cranial nerve from the third nerve nucleus in the brain stem, as explained in Chapter 51. Stimulation of the parasympathetic nerves contracts both sets of ciliary muscle fibers, which relaxes the lens ligaments, thus allowing the lens to become thicker and increase its refractive power. With this increased refractive power, the eye focuses on objects nearer than when the eye has less refractive power. Consequently, as a distant object moves toward the eye, the number of parasympathetic impulses impinging on the ciliary muscle must be progressively increased for the eye to keep the object constantly in focus. (Sympathetic stimulation has an additional effect in relaxing the ciliary muscle, but this effect is so weak that it plays almost no role in the normal accommodation mechanism; the neurology of this is discussed in Chapter 51.)

Presbyopia—Loss of Accommodation by the Lens. As a person grows older, the lens grows larger and thicker and becomes far less elastic, partly because of progressive denaturation of the lens proteins. The ability of the lens to change shape decreases with age. The power of accommodation decreases from about 14 diopters in a child to less than 2 diopters by the time a person reaches 45 to 50 years; it then decreases to essentially 0 diopters at age 70 years. Thereafter, the lens remains almost totally nonaccommodating, a condition known as "presbyopia."

Once a person has reached the state of presbyopia, each eye remains focused permanently at an almost constant distance; this distance depends on the physical characteristics of each person's eyes. The eyes can no longer accommodate for both near and far vision. To see clearly both in the distance and nearby, an older person must wear bifocal glasses with the upper segment focused for far-seeing and the lower segment focused for near-seeing (e.g., for reading).

Pupillary Diameter

The major function of the iris is to increase the amount of light that enters the eye during darkness and to decrease the amount of light that enters the eye in daylight. The reflexes for controlling this mechanism are considered in the discussion of the neurology of the eye in Chapter 51.

The amount of light that enters the eye through the pupil is proportional to the *area* of the pupil or to the *square of the diameter* of the pupil. The pupil of the human eye can become as small as about 1.5 millimeters and as large as 8 millimeters in diameter. The quantity of light entering the eye can change about 30-fold as a result of changes in pupillary aperture.

"Depth of Focus" of the Lens System Increases with Decreasing Pupillary Diameter. Figure 49-11 shows two eyes that are exactly alike except for the diameters of the pupillary apertures. In the upper eye, the pupillary aperture is small, and in the lower eye, the aperture is large. In front of each of these two eyes are two small point sources of light; light from each passes through the pupillary aperture and focuses on the retina. Consequently, in both eyes, the retina sees two spots of light in perfect focus. It is evident from the diagrams, however, that if the retina is moved forward or backward to an out-of-focus position, the size of each spot will not change much in the upper eye, but in the lower eye the size of each spot will increase greatly, becoming a "blur circle." In other words, the upper lens system has far greater *depth of focus* than the bottom lens system. When a lens system has great depth of focus, the retina can be displaced considerably from the focal plane or the lens strength can change considerably from normal and the image will still remain nearly in sharp focus, whereas when a lens system has a "shallow" depth of focus, moving the retina only slightly away from the focal plane causes extreme blurring.

The greatest possible depth of focus occurs when the pupil is extremely small. The reason for this is that, with a very small aperture, almost all the rays pass through the center of the lens, and the central-most rays are always in focus, as explained earlier.

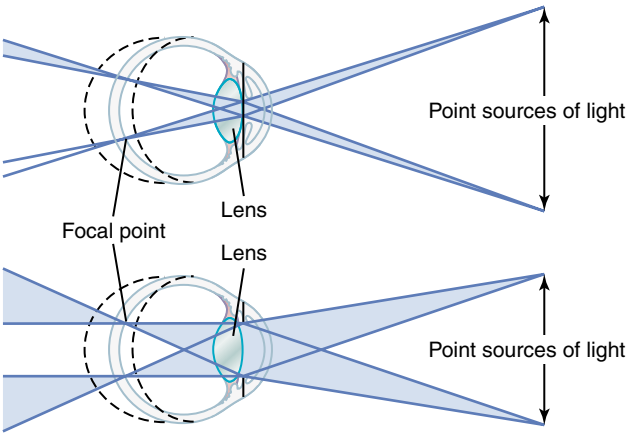


Figure 49-11 Effect of small (*top*) and large (*bottom*) pupillary apertures on "depth of focus."

Errors of Refraction

Emmetropia (Normal Vision). As shown in Figure 49-12, the eye is considered to be normal, or "emmetropic," if parallel light rays *from distant objects* are in sharp focus on the retina

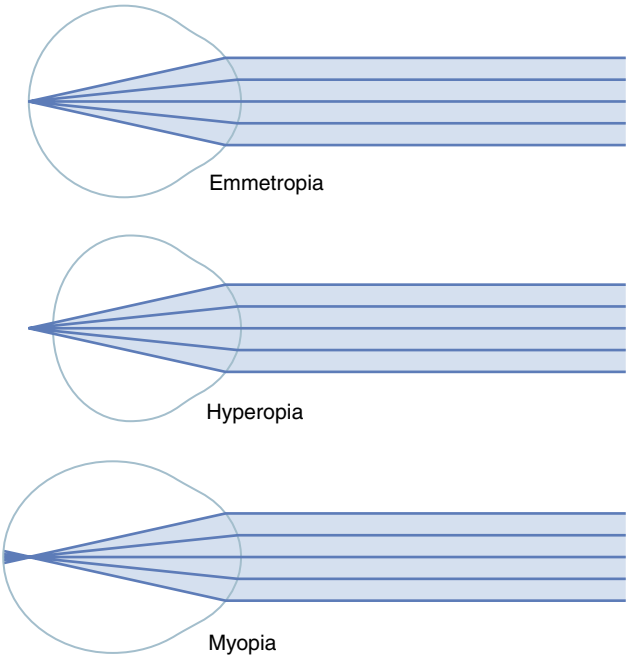


Figure 49-12 Parallel light rays focus on the retina in emmetropia, behind the retina in hyperopia, and in front of the retina in myopia.

when the ciliary muscle is completely relaxed. This means that the emmetropic eye can see all distant objects clearly with its ciliary muscle relaxed. However, to focus objects at close range, the eye must contract its ciliary muscle and thereby provide appropriate degrees of accommodation.

Hyperopia (Farsightedness). Hyperopia, which is also known as "farsightedness," is usually due to either an eyeball that is too short or, occasionally, a lens system that is too weak. In this condition, as seen in the middle panel of Figure 49-12, parallel light rays are not bent sufficiently by the relaxed lens system to come to focus by the time they reach the retina. To overcome this abnormality, the ciliary muscle must contract to increase the strength of the lens. By using the mechanism of accommodation, a farsighted person is capable of focusing distant objects on the retina. If the person has used only a small amount of strength in the ciliary muscle to accommodate for the distant objects, he or she still has much accommodative power left, and objects closer and closer to the eye can also be focused sharply until the ciliary muscle has contracted to its limit. In old age, when the lens becomes "presbyopic," a farsighted person is often unable to accommodate the lens sufficiently to focus even distant objects, much less near objects.

Myopia (Nearsightedness). In myopia, or "nearsightedness," when the ciliary muscle is completely relaxed, the light rays coming from distant objects are focused in front of the retina, as shown in the bottom panel of Figure 49-12. This is usually due to too long an eyeball, but it can result from too much refractive power in the lens system of the eye.

No mechanism exists by which the eye can decrease the strength of its lens to less than that which exists when the ciliary muscle is completely relaxed. A myopic person has no mechanism by which to focus distant objects sharply on the retina. However, as an object moves nearer to the person's eye, it finally gets close enough that its image can be focused. Then, when the object comes still closer to the eye,

the person can use the mechanism of accommodation to keep the image focused clearly. A myopic person has a definite limiting "far point" for clear vision.

Correction of Myopia and Hyperopia by Use of Lenses. It will be recalled that light rays passing through a concave lens diverge. If the refractive surfaces of the eye have too much refractive power, as in *myopia,* this excessive refractive power can be neutralized by placing in front of the eye a concave spherical lens, which will diverge rays. Such correction is demonstrated in the upper diagram of Figure 49-13.

Conversely, in a person who has *hyperopia*—that is, someone who has too weak a lens system—the abnormal vision can be corrected by adding refractive power using a convex lens in front of the eye. This correction is demonstrated in the lower diagram of Figure 49-13.

One usually determines the strength of the concave or convex lens needed for clear vision by "trial and error"—that is, by trying first a strong lens and then a stronger or weaker lens until the one that gives the best visual acuity is found.

Astigmatism. Astigmatism is a refractive error of the eye that causes the visual image in one plane to focus at a different distance from that of the plane at right angles. This most often results from too great a curvature of the cornea in one plane of the eye. An example of an astigmatic lens would be a lens surface like that of an egg lying sidewise to the incoming light. The degree of curvature in the plane through the long axis of the egg is not nearly as great as the degree of curvature in the plane through the short axis.

Because the curvature of the astigmatic lens along one plane is less than the curvature along the other plane, light rays striking the peripheral portions of the lens in one plane are not bent nearly as much as the rays striking the peripheral portions of the other plane. This is demonstrated in Figure 49-14, which shows rays of light originating from a point source and passing through an oblong, astigmatic lens. The light rays in the vertical plane, indicated by plane BD, are refracted greatly by the astigmatic lens because of the greater curvature in the vertical direction than in the horizontal direction. By contrast, the light rays in the horizontal plane, indicated by plane AC, are not bent nearly as much as the light rays in vertical plane BD. It is obvious that light rays passing through an astigmatic lens do not all come to a common focal point because the light rays passing through one plane focus far in front of those passing through the other plane.

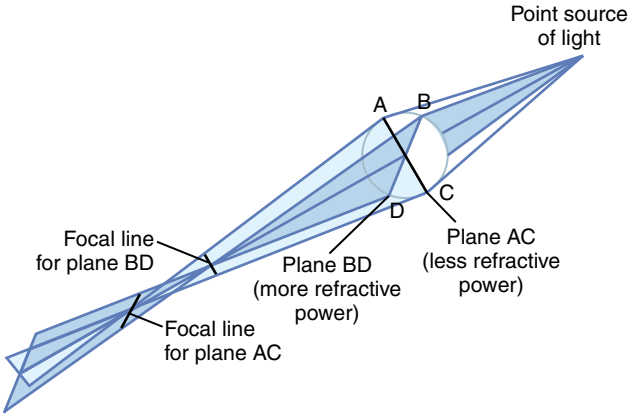


Figure 49-14 Astigmatism, demonstrating that light rays focus at one focal distance in one focal plane (*plane AC*) and at another focal distance in the plane at a right angle (*plane BD*).

The accommodative power of the eye can never compensate for astigmatism because, during accommodation, the curvature of the eye lens changes approximately equally in both planes; therefore, in astigmatism, each of the two planes requires a different degree of accommodation. Thus, without the aid of glasses, a person with astigmatism never sees in sharp focus.

Correction of Astigmatism with a Cylindrical Lens. One may consider an astigmatic eye as having a lens system made up of two cylindrical lenses of different strengths and placed at right angles to each other. To correct for astigmatism, the usual procedure is to find a spherical lens by trial and error that corrects the focus in one of the two planes of the astigmatic lens. Then an additional cylindrical lens is used to correct the remaining error in the remaining plane. To do this, both the *axis* and the *strength* of the required cylindrical lens must be determined.

Several methods exist for determining the axis of the abnormal cylindrical component of the lens system of an eye. One of these methods is based on the use of parallel black bars of the type shown in Figure 49-15. Some of these parallel bars are vertical, some horizontal, and some at various angles to the vertical and horizontal axes. After placing various spherical lenses in front of the astigmatic eye, a strength of lens that causes sharp focus of one set of parallel bars but does not correct the fuzziness of the set of bars at right angles to the sharp bars is usually found. It can be shown from the physical principles of optics discussed earlier in this chapter that the *axis* of the *out-of-focus* cylindrical component of the optical system is parallel to the bars that are fuzzy. Once this axis is found, the examiner tries progressively stronger and weaker positive or negative *cylindrical* lenses, the axes of which are placed in line with the out-of-focus bars, until the patient sees all the crossed bars with equal clarity. When this has been accomplished, the examiner directs the optician to grind a special lens combining both the spherical correction and the cylindrical correction at the appropriate axis.

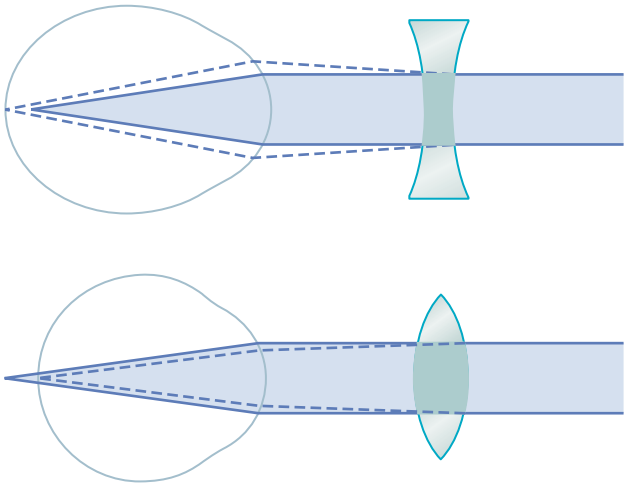


Figure 49-13 Correction of myopia with a concave lens, and correction of hyperopia with a convex lens.

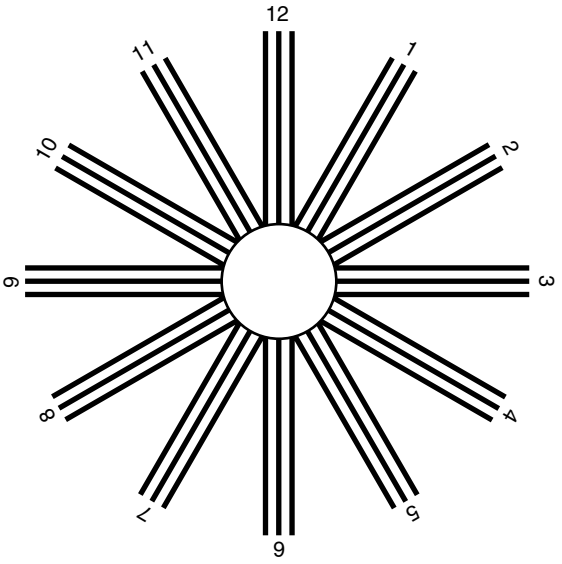


Figure 49-15 Chart composed of parallel black bars at different angular orientations for determining the axis of astigmatism.

Correction of Optical Abnormalities by Use of Contact Lenses

Glass or plastic contact lenses that fit snugly against the anterior surface of the cornea can be inserted. These lenses are held in place by a thin layer of tear fluid that fills the space between the contact lens and the anterior eye surface.

A special feature of the contact lens is that it nullifies almost entirely the refraction that normally occurs at the anterior surface of the cornea. The reason for this is that the tears between the contact lens and the cornea have a refractive index almost equal to that of the cornea, so the anterior surface of the cornea no longer plays a significant role in the eye's optical system. Instead, the outer surface of the contact lens plays the major role. Thus, the refraction of this surface of the contact lens substitutes for the cornea's usual refraction. This is especially important in people whose eye refractive errors are caused by an abnormally shaped cornea, such as those who have an odd-shaped, bulging cornea—a condition called *keratoconus.* Without the contact lens, the bulging cornea causes such severe abnormality of vision that almost no glasses can correct the vision satisfactorily; when a contact lens is used, however, the corneal refraction is neutralized and normal refraction by the outer surface of the contact lens is substituted.

The contact lens has several other advantages as well, including (1) the lens turns with the eye and gives a broader field of clear vision than glasses do, and (2) the contact lens has little effect on the size of the object the person sees through the lens, whereas lenses placed 1 centimeter or so in front of the eye do affect the size of the image, in addition to correcting the focus.

Cataracts—Opaque Areas in the Lens

"Cataracts" are an especially common eye abnormality that occurs mainly in older people. A cataract is a cloudy or opaque area or areas in the lens. In the early stage of cataract formation, the proteins in some of the lens fibers become denatured. Later, these same proteins coagulate to form opaque areas in place of the normal transparent protein fibers.

When a cataract has obscured light transmission so greatly that it seriously impairs vision, the condition can be corrected by surgical removal of the lens. When this is done, the eye loses a large portion of its refractive power, which must be replaced by a powerful convex lens in front of the eye; usually, however, an artificial plastic lens is implanted in the eye in place of the removed lens.

Visual Acuity

Theoretically, light from a distant point source, when focused on the retina, should be infinitely small. However, because the lens system of the eye is never perfect, such a retinal spot ordinarily has a total diameter of about 11 micrometers, even with maximal resolution of the normal eye optical system. The spot is brightest in its center and shades off gradually toward the edges, as shown by the two-point images in Figure 49-16.

The average diameter of the cones in the *fovea* of the retina—the central part of the retina, where vision is most highly developed—is about 1.5 micrometers, which is one-seventh the diameter of the spot of light. Nevertheless, because the spot of light has a bright center point and shaded edges, a person can normally distinguish two separate points if their centers lie up to 2 micrometers apart on the retina, which is slightly greater than the width of a foveal cone. This discrimination between points is also shown in Figure 49-16.

The normal visual acuity of the human eye for discriminating between point sources of light is about 25 seconds of arc. That is, when light rays from two separate points strike the eye with an angle of at least 25 seconds between them, they can usually be recognized as two points instead of one. This means that a person with normal visual acuity looking at two bright pinpoint spots of light 10 meters away can barely distinguish the spots as separate entities when they are 1.5 to 2 millimeters apart.

The fovea is less than 0.5 millimeter (<500 micrometers) in diameter, which means that maximum visual acuity occurs in less than 2 degrees of the visual field. Outside this foveal area, the visual acuity becomes progressively poorer, decreasing more than 10-fold as the periphery is approached. This is caused by the connection of more and more rods and cones to each optic nerve fiber in the

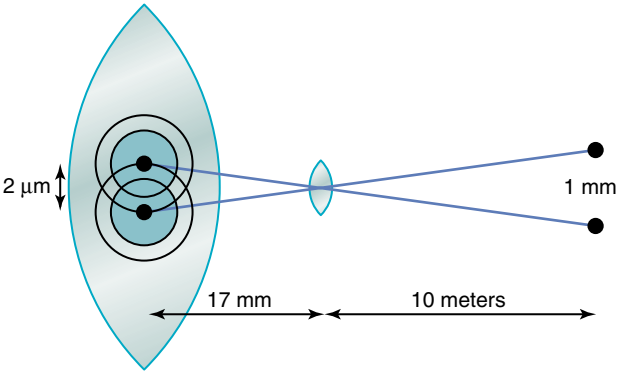


Figure 49-16 Maximum visual acuity for two-point sources of light.

nonfoveal, more peripheral parts of the retina, as discussed in Chapter 51.

Clinical Method for Stating Visual Acuity. The chart for testing eyes usually consists of letters of different sizes placed 20 feet away from the person being tested. If the person can see well the letters of a size that he or she should be able to see at 20 feet, the person is said to have 20/20 vision—that is, normal vision. If the person can see only letters that he or she should be able to see at 200 feet, the person is said to have 20/200 vision. In other words, the clinical method for expressing visual acuity is to use a mathematical fraction that expresses the ratio of two distances, which is also the ratio of one's visual acuity to that of a person with normal visual acuity.

Determination of Distance of an Object from the Eye—"Depth Perception"

A person normally perceives distance by three major means: (1) the sizes of the images of known objects on the retina, (2) the phenomenon of moving parallax, and (3) the phenomenon of stereopsis. This ability to determine distance is called *depth perception.*

Determination of Distance by Sizes of Retinal Images of Known Objects. If one knows that a person being viewed is 6 feet tall, one can determine how far away the person is simply by the size of the person's image on the retina. One does not consciously think about the size, but the brain has learned to calculate automatically from image sizes the distances of objects when the dimensions are known.

Determination of Distance by Moving Parallax. Another important means by which the eyes determine distance is that of moving parallax. If an individual looks off into the distance with the eyes completely still, he or she perceives no moving parallax, but when the person moves his or her head to one side or the other, the images of close-by objects move rapidly across the retinas, while the images of distant objects remain almost completely stationary. For instance, by moving the head 1 inch to the side when the object is only 1 inch in front of the eye, the image moves almost all the way across the retinas, whereas the image of an object 200 feet away from the eyes does not move perceptibly. Thus, by using this mechanism of moving parallax, one can tell the *relative distances* of different objects even though only one eye is used.

Determination of Distance by Stereopsis—Binocular Vision. Another method by which one perceives parallax is that of "binocular vision." Because one eye is a little more than 2 inches to one side of the other eye, the images on the two retinas are different from each other. For instance, an object 1 inch in front of the nose forms an image on the left side of the retina of the left eye but on the right side of the retina of the right eye, whereas a small object 20 feet in front of the nose has its image at closely corresponding points in the centers of the two retinas. This type of parallax is demonstrated in Figure 49-17, which shows the images of a red spot and a yellow square actually reversed on the two retinas because they are at different distances in front of the eyes. This gives a type of parallax that is present all the time when both eyes are being used. It is almost entirely this binocular parallax (or *stereopsis*) that gives a person with two eyes far greater ability to judge relative distances *when objects are nearby* than a person who has only one eye. However, stereopsis is virtually useless for depth perception at distances beyond 50 to 200 feet.

Ophthalmoscope

The ophthalmoscope is an instrument through which an observer can look into another person's eye and see the retina with clarity. Although the ophthalmoscope appears to be a relatively complicated instrument, its principles are simple. The basic components are shown in Figure 49-18 and can be explained as follows.

If a bright spot of light is on the retina of an *emmetropic eye,* light rays from this spot diverge toward the lens system of the eye. After passing through the lens system, they are parallel with one another because the retina is located one focal length distance behind the lens system. Then, when

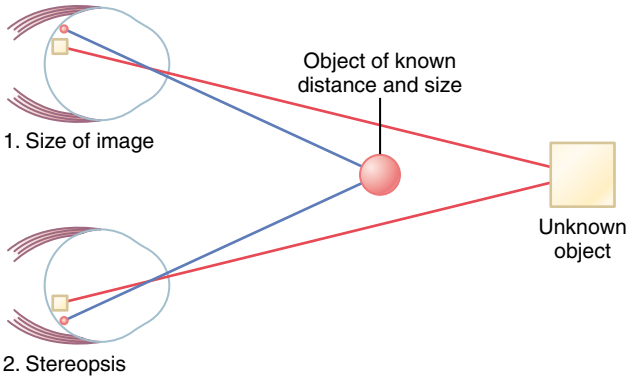


Figure 49-17 Perception of distance (*1*) by the size of the image on the retina and (*2*) as a result of stereopsis.

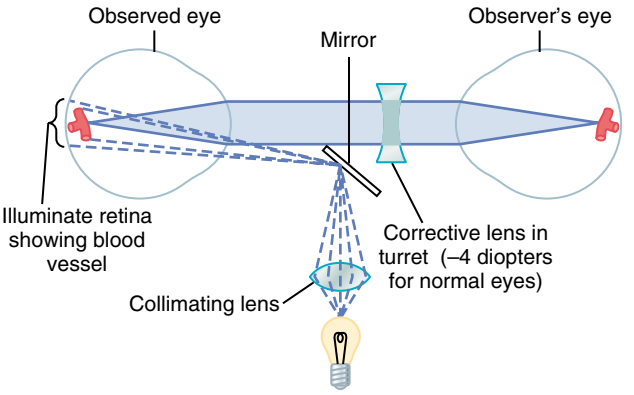


Figure 49-18 Optical system of the ophthalmoscope.

these parallel rays pass into an emmetropic eye of another person, they focus again to a point focus on the retina of the second person, because his or her retina is also one focal length distance behind the lens. Any spot of light on the retina of the observed eye projects to a focal spot on the retina of the observing eye. Thus, if the retina of one person is made to emit light, the image of his or her retina will be focused on the retina of the observer, provided the two eyes are emmetropic and are simply looking into each other.

To make an ophthalmoscope, one need only devise a means for illuminating the retina to be examined. Then, the reflected light from that retina can be seen by the observer simply by putting the two eyes close to each other. To illuminate the retina of the observed eye, an angulated mirror or a segment of a prism is placed in front of the observed eye in such a manner, as shown in Figure 49-18, that light from a bulb is reflected into the observed eye. Thus, the retina is illuminated through the pupil, and the observer sees into the subject's pupil by looking over the edge of the mirror or prism or *through* an appropriately designed prism.

It is clear that these principles apply only to people with completely emmetropic eyes. If the refractive power of either the observed eye or the observer's eye is abnormal, it is necessary to correct the refractive power for the observer to see a sharp image of the observed retina. The usual ophthalmoscope has a series of very small lenses mounted on a turret so that the turret can be rotated from one lens to another until the correction for abnormal refraction is made by selecting a lens of appropriate strength. In normal young adults, natural accommodative reflexes occur, causing an approximate +2-diopter increase in strength of the lens of each eye. To correct for this, it is necessary that the lens turret be rotated to approximately −4-diopter correction.

Fluid System of the Eye—Intraocular Fluid

The eye is filled with *intraocular fluid,* which maintains sufficient pressure in the eyeball to keep it distended. Figure 49-19 demonstrates that this fluid can be divided into two portions—*aqueous humor,* which lies in front of the lens, and *vitreous humor,* which is between the posterior surface of the lens and the retina. The aqueous humor is a freely flowing fluid, whereas the vitreous humor, sometimes called the *vitreous body,* is a gelatinous mass held together by a fine fibrillar network composed primarily of greatly elongated proteoglycan molecules. Both water and dissolved substances can *diffuse* slowly in the vitreous humor, but there is little *flow* of fluid.

Aqueous humor is continually being formed and reabsorbed. The balance between formation and reabsorption of aqueous humor regulates the total volume and pressure of the intraocular fluid.

Formation of Aqueous Humor by the Ciliary Body

Aqueous humor is formed in the eye *at an average rate of 2 to 3 microliters each minute.* Essentially all of it is secreted by the *ciliary processes,* which are linear folds projecting from the *ciliary body* into the space behind the iris where the lens ligaments and ciliary muscle attach to the eyeball. A cross section of these ciliary processes is shown in Figure 49-20, and their relation to the fluid chambers of the eye can be seen in Figure 49-19. Because of their folded architecture, the total surface area of the ciliary processes is about 6 square centimeters in each eye—a large area, considering the small size of the ciliary body. The surfaces of these processes are covered by highly secretory epithelial cells, and immediately beneath them is a highly vascular area.

Aqueous humor is formed almost entirely as an active secretion by the epithelium of the ciliary processes. Secretion begins with active transport of sodium ions into the spaces between the epithelial cells. The sodium ions pull chloride and bicarbonate ions along with them to maintain electrical neutrality. Then all these ions together cause osmosis of water from the blood capillaries lying below into the same epithelial intercellular spaces, and the resulting solution washes from the spaces of the ciliary processes into the anterior chamber of the eye. In addition, several nutrients are transported across the epithelium by active transport or facilitated diffusion; they include amino acids, ascorbic acid, and glucose.

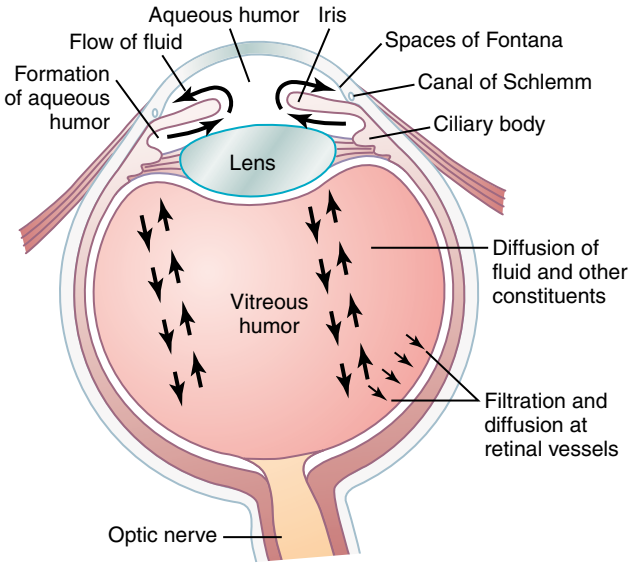


Figure 49-19 Formation and flow of fluid in the eye.

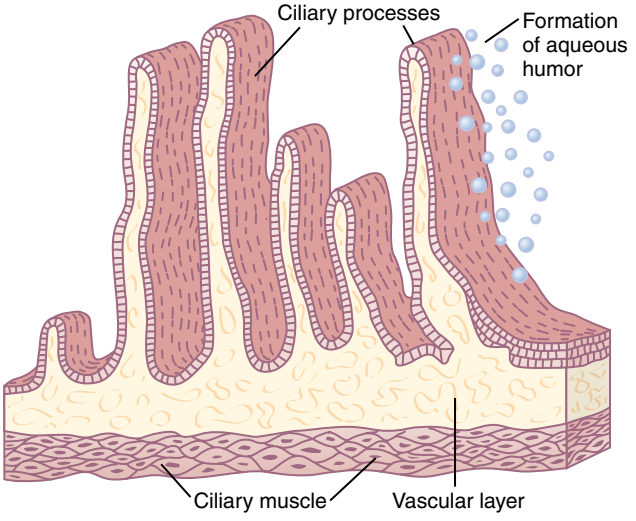


Figure 49-20 Anatomy of the ciliary processes. Aqueous humor is formed on surfaces.

Outflow of Aqueous Humor from the Eye

After aqueous humor is formed by the ciliary processes, it first flows, as shown in Figure 49-19, *through the pupil into the anterior chamber of the eye.* From here, the fluid flows *anterior to the lens and* into the *angle between the cornea and the iris,* then through a meshwork of *trabeculae,* finally entering the *canal of Schlemm,* which empties into extraocular veins. Figure 49-21 demonstrates the anatomical structures at this iridocorneal angle, showing that the spaces between the trabeculae extend all the way from the anterior chamber to the canal of Schlemm. The canal of Schlemm is a thin-walled vein that extends circumferentially all the way around the eye. Its endothelial membrane is so porous that even large protein molecules, as well as small particulate matter up to the size of red blood cells, can pass from the anterior chamber into the canal of Schlemm. Even though the canal of Schlemm is actually a venous blood vessel, so much aqueous humor normally flows into it that it is filled only with aqueous humor rather than with blood. The small veins that lead from the canal of Schlemm to the larger veins of the eye usually contain only aqueous humor, and they are called *aqueous veins.*

Intraocular Pressure

The average normal intraocular pressure is about 15 mm Hg, with a range from 12 to 20 mm Hg.

Tonometry. Because it is impractical to pass a needle into a patient's eye to measure intraocular pressure, this pressure is measured clinically by using a "tonometer," the principle of which is shown in Figure 49-22. The cornea of the eye is anesthetized with a local anesthetic, and the footplate of the tonometer is placed on the cornea. A small force is then applied to a central plunger, causing the part of the cornea beneath the plunger to be displaced inward. The amount of displacement is recorded on the scale of the tonometer, and this is calibrated in terms of intraocular pressure.

Regulation of Intraocular Pressure. Intraocular pressure remains constant in the normal eye, usually within ±2 mm Hg of its normal level, which averages about 15 mm Hg. The level of this pressure is determined mainly by the resistance to outflow of aqueous humor from the anterior chamber into the canal of Schlemm. This outflow resistance results from the meshwork of trabeculae through which the fluid must percolate on its way from the lateral angles of the anterior chamber to the wall of the canal of Schlemm. These trabeculae have minute openings of only 2 to 3 micrometers. The rate of fluid flow into the canal increases markedly as the pressure rises. At about 15 mm Hg in the normal eye, the amount of fluid leaving the eye by way of the canal of Schlemm usually averages 2.5 µl/min and equals the inflow of fluid from the ciliary body. The pressure normally remains at about this level of 15 mm Hg.

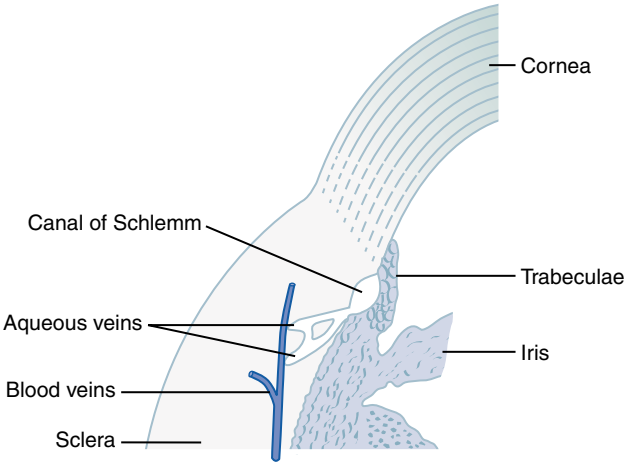


Figure 49-21 Anatomy of the iridocorneal angle, showing the system for outflow of aqueous humor from the eyeball into the conjunctival veins.

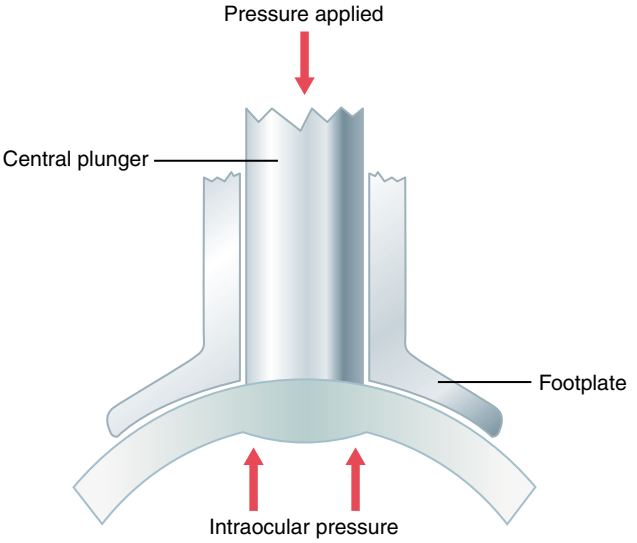


Figure 49-22 Principles of the tonometer.

Mechanism for Cleansing the Trabecular Spaces and Intraocular Fluid. When large amounts of debris are present in the aqueous humor, as occurs after hemorrhage into the eye or during intraocular infection, the debris is likely to accumulate in the trabecular spaces leading from the anterior chamber to the canal of Schlemm; this debris can prevent adequate reabsorption of fluid from the anterior chamber, sometimes causing "glaucoma," as explained subsequently. However, on the surfaces of the trabecular plates are large numbers of phagocytic cells. Immediately outside the canal of Schlemm is a layer of interstitial gel that contains large numbers of reticuloendothelial cells that have an extremely high capacity for engulfing debris and digesting it into small molecular substances that can then be absorbed. Thus, this phagocytic system keeps the trabecular spaces cleaned. The surface of the iris and other surfaces of the eye behind the iris are covered with an epithelium that is capable of phagocytizing proteins and small particles from the aqueous humor, thereby helping to maintain a clear fluid.

"Glaucoma"—a Principal Cause of Blindness. Glaucoma is one of the most common causes of blindness. It is a disease of the eye in which the intraocular pressure becomes pathologically high, sometimes rising acutely to 60 to 70 mm Hg. Pressures above 25 to 30 mm Hg can cause loss of vision when maintained for long periods. Extremely high pressures can cause blindness within days or even hours. As the pressure rises, the axons of the optic nerve are compressed where they leave the eyeball at the optic disc. This compression is believed to block axonal flow of cytoplasm from the retinal neuronal cell bodies into the optic nerve fibers leading to the brain. The result is lack of appropriate nutrition of the fibers,

which eventually causes death of the involved fibers. It is possible that compression of the retinal artery, which enters the eyeball at the optic disc, also adds to the neuronal damage by reducing nutrition to the retina.

In most cases of glaucoma, the abnormally high pressure results from increased resistance to fluid outflow through the trabecular spaces into the canal of Schlemm at the iridocorneal junction. For instance, in acute eye inflammation, white blood cells and tissue debris can block these trabecular spaces and cause an acute increase in intraocular pressure. In chronic conditions, especially in older individuals, fibrous occlusion of the trabecular spaces appears to be the likely culprit.

Glaucoma can sometimes be treated by placing drops in the eye that contain a drug that diffuses into the eyeball and reduces the secretion or increases the absorption of aqueous humor. When drug therapy fails, operative techniques to open the spaces of the trabeculae or to make channels to allow fluid to flow directly from the fluid space of the eyeball into the subconjunctival space outside the eyeball can often effectively reduce the pressure.

Bibliography

Buisseret P: Influence of extraocular muscle proprioception on vision, *Physiol Rev* 75:323, 1995.

Buznego C, Trattler WB: Presbyopia-correcting intraocular lenses, *Curr Opin Ophthalmol* 20:13, 2009.

Candia OA, Alvarez LJ: Fluid transport phenomena in ocular epithelia, *Prog Retin Eye Res* 27:197, 2008.

Congdon NG, Friedman DS, Lietman T: Important causes of visual impairment in the world today, *JAMA* 290:2057, 2003.

Doane JF: Accommodating intraocular lenses, *Curr Opin Ophthalmol* 15:16, 2004.

Khaw PT, Shah P, Elkington AR: Glaucoma-1: diagnosis, *BMJ* 328:97, 2004.

Krag S, Andreassen TT: Mechanical properties of the human lens capsule, *Prog Retin Eye Res* 22:749, 2003.

Kwon YH, Fingert JH, Kuehn MH, et al: Primary open-angle glaucoma, *N Engl J Med* 360:1113, 2009.

Mathias RT, Rae JL, Baldo GJ: Physiological properties of the normal lens, *Physiol Rev* 77:21, 1997.

Sakimoto T, Rosenblatt MI, Azar DT: Laser eye surgery for refractive errors, *Lancet* 367:1432, 2006.

Schaeffel F, Simon P, Feldkaemper M, et al: Molecular biology of myopia, *Clin Exp Optom* 86:295, 2003.

Schwartz K, Budenz D: Current management of glaucoma, *Curr Opin Ophthalmol* 15:119, 2004.

Smith G: The optical properties of the crystalline lens and their significance, *Clin Exp Optom* 86:3, 2003.

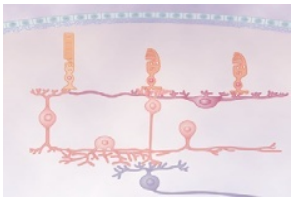
Tan JC, Peters DM, Kaufman PL: Recent developments in understanding the pathophysiology of elevated intraocular pressure, *Curr Opin Ophthalmol* 17:168, 2006.

Weber AJ, Harman CD, Viswanathan S: Effects of optic nerve injury, glaucoma, and neuroprotection on the survival, structure, and function of ganglion cells in the mammalian retina, *J Physiol* 586:4393, 2008.

Weinreb RN, Khaw PT: Primary open-angle glaucoma, *Lancet* 363:1711, 2004.

CHAPTER 50

The Eye: II. Receptor and Neural Function of the Retina



The retina is the light-sensitive portion of the eye that contains (1) the *cones,* which are responsible for color vision, and (2) the *rods,* which can detect dim light and are mainly responsible for black and white vision and vision in the dark. When either rods or cones are excited, signals are transmitted first through successive layers of neurons in the retina and, finally, into optic nerve fibers and the cerebral cortex. The purpose of this chapter is to explain the mechanisms by which the rods and cones detect light and color and convert the visual image into optic nerve signals.

Anatomy and Function of the Structural Elements of the Retina

Layers of the Retina. Figure 50-1 shows the functional components of the retina, which are arranged in layers from the outside to the inside as follows: (1) pigmented layer, (2) layer of rods and cones projecting to the pigment, (3) outer nuclear layer containing the cell bodies of the rods and cones, (4) outer plexiform layer, (5) inner nuclear layer, (6) inner plexiform layer, (7) ganglionic layer, (8) layer of optic nerve fibers, and (9) inner limiting membrane.

After light passes through the lens system of the eye and then through the vitreous humor, it *enters the retina from the inside of the eye* (see Figure 50-1); that is, it passes first through the ganglion cells and then through the plexiform and nuclear layers before it finally reaches the layer of rods and cones located all the way on the outer edge of the retina. This distance is a thickness of several hundred micrometers; visual acuity is decreased by this passage through such nonhomogeneous tissue. However, in the *central foveal region of the retina,* as discussed subsequently, the inside layers are pulled aside to decrease this loss of acuity.

Foveal Region of the Retina and Its Importance in Acute Vision. The *fovea* is a minute area in the center of the retina, shown in Figure 50-2, occupying a total area a little more than 1 square millimeter; it is especially capable of acute and detailed vision. The *central fovea,* only 0.3 millimeter in diameter, is composed almost entirely of cones; these cones have a special structure that aids their detection of detail in the visual image. That is, the foveal cones have especially long and slender bodies, in contradistinction to the much fatter cones located more peripherally in the retina. Also, in the foveal region, the blood vessels, ganglion cells, inner nuclear layer of cells, and plexiform layers are all displaced to one side rather than resting directly on top of the cones. This allows light to pass unimpeded to the cones.

Rods and Cones. Figure 50-3 is a diagrammatic representation of the essential components of a photoreceptor (either a rod or a cone). As shown in Figure 50-4, the outer segment of the cone is conical in shape. In general, the rods are narrower and longer than the cones, but this is not always the case. In the peripheral portions of the retina, the rods are 2 to 5 micrometers in diameter, whereas the cones are 5 to 8 micrometers in diameter; in the central part of the retina, in the fovea, there are rods, and the cones are slender and have a diameter of only 1.5 micrometers.

The major functional segments of either a rod or cone are shown in Figure 50-3: (1) the *outer segment,* (2) the *inner segment,* (3) the *nucleus,* and (4) the *synaptic body*. The light-sensitive photochemical is found in the outer segment. In the case of the rods, this is *rhodopsin;* in the cones, it is one of three "color" photochemicals, usually called simply *color pigments,* that function almost exactly the same as rhodopsin except for differences in spectral sensitivity.

Note in the *outer segments* of the rods and cones in Figures 50-3 and 50-4 the large numbers of *discs*. Each disc is actually an infolded shelf of cell membrane. There are as many as 1000 discs in each rod or cone.

Both rhodopsin and the color pigments are conjugated proteins. They are incorporated into the membranes of the discs in the form of transmembrane proteins. The concentrations of these photosensitive pigments in the discs are so great that the pigments themselves constitute about 40 percent of the entire mass of the outer segment.

The *inner segment* of the rod or cone contains the usual cytoplasm with cytoplasmic organelles. Especially important are the mitochondria, which, as explained later, play the important role of providing energy for function of the photoreceptors.

The *synaptic body* is the portion of the rod or cone that connects with subsequent neuronal cells, the *horizontal* and *bipolar cells,* which represent the next stages in the vision chain.

Pigment Layer of the Retina. The black pigment *melanin* in the pigment layer prevents light reflection throughout the globe of the eyeball; this is extremely important for clear

Figure 50-1 Layers of retina.

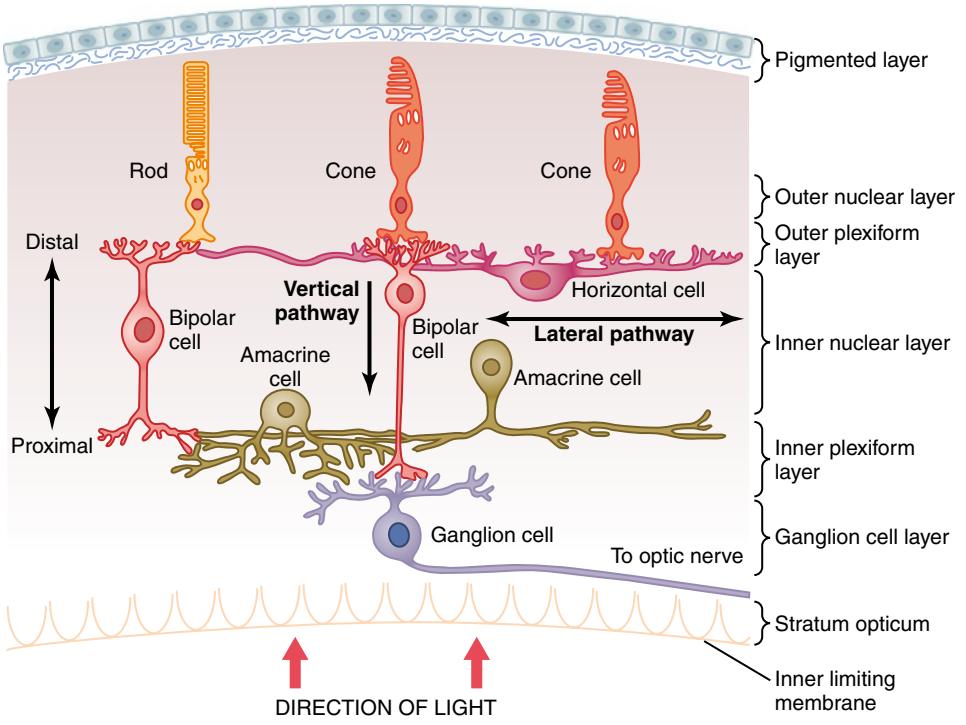
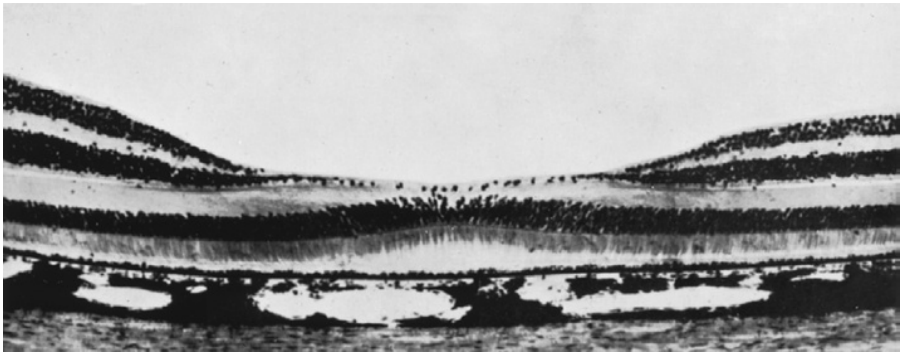


Figure 50-2 Photomicrograph of the macula and of the fovea in its center. Note that the inner layers of the retina are pulled to the side to decrease interference with light transmission. (From Fawcett DW: Bloom and Fawcett: A Textbook of Histology, 11th ed. Philadelphia: WB Saunders, 1986; courtesy H. Mizoguchi.)



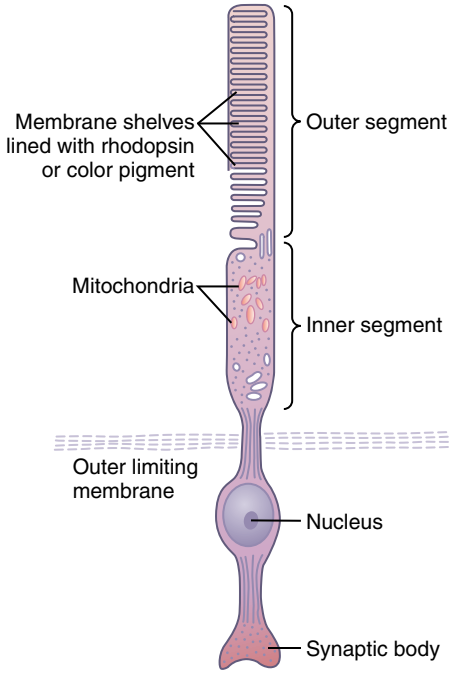


Figure 50-3 Schematic drawing of the functional parts of the rods and cones.

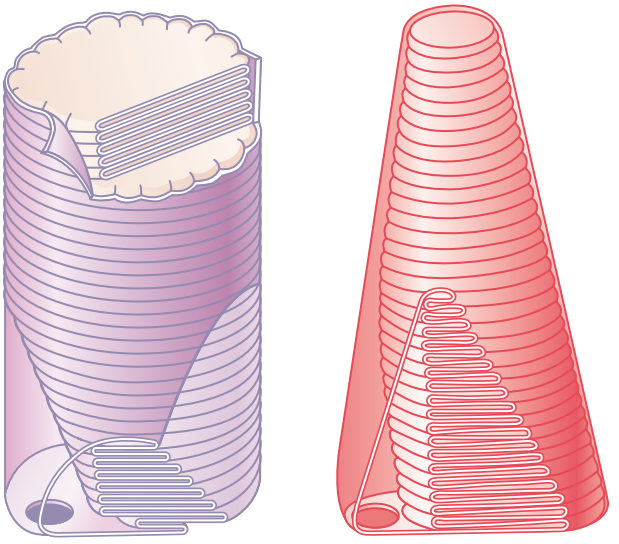


Figure 50-4 Membranous structures of the outer segments of a rod (*left*) and a cone (*right*). (Courtesy Dr. Richard Young.)

vision. This pigment performs the same function in the eye as the black coloring inside the bellows of a camera. Without it, light rays would be reflected in all directions within the eyeball and would cause diffuse lighting of the retina rather than the normal contrast between dark and light spots required for formation of precise images.

The importance of melanin in the pigment layer is well illustrated by its absence in *albinos,* people who are hereditarily lacking in melanin pigment in all parts of their bodies. When an albino enters a bright room, light that impinges on the retina is reflected in all directions inside the eyeball by the unpigmented surfaces of the retina and by the underlying sclera, so a single discrete spot of light that would normally excite only a few rods or cones is reflected everywhere and excites many receptors. Therefore, the visual acuity of albinos, even with the best optical correction, is seldom better than 20/100 to 20/200 rather than the normal 20/20 values.

The pigment layer also stores large quantities of *vitamin A.* This vitamin A is exchanged back and forth through the cell membranes of the outer segments of the rods and cones, which themselves are embedded in the pigment. We show later that vitamin A is an important precursor of the photosensitive chemicals of the rods and cones.

Blood Supply of the Retina—The Central Retinal Artery and the Choroid. The nutrient blood supply for the internal layers of the retina is derived from the central retinal artery, which enters the eyeball through the center of the optic nerve and then divides *to supply the entire inside retinal surface.* Thus, the inner layers of the retina have their own blood supply independent of the other structures of the eye.

However, the outermost layer of the retina is adherent to the *choroid,* which is also a highly vascular tissue lying between the retina and the sclera. The outer layers of the retina, especially the outer segments of the rods and cones, depend mainly on diffusion from the choroid blood vessels for their nutrition, especially for their oxygen.

Retinal Detachment. The neural retina occasionally *detaches from the pigment epithelium.* In some instances, the cause of such detachment is injury to the eyeball that allows fluid or blood to collect between the neural retina and the pigment epithelium. Detachment is occasionally caused by contracture of fine collagenous fibrils in the vitreous humor, which pull areas of the retina toward the interior of the globe.

Partly because of diffusion across the detachment gap and partly because of the independent blood supply to the neural retina through the retinal artery, the detached retina can resist degeneration for days and can become functional again if it is surgically replaced in its normal relation with the pigment epithelium. If it is not replaced soon, however, the retina will be destroyed and will be unable to function even after surgical repair.

Photochemistry of Vision

Both rods and cones contain chemicals that decompose on exposure to light and, in the process, excite the nerve fibers leading from the eye. The light-sensitive chemical in the *rods* is called *rhodopsin;* the light-sensitive chemicals in the *cones,* called *cone pigments* or *color pigments,* have compositions only slightly different from that of rhodopsin.

In this section, we discuss principally the photochemistry of rhodopsin, but the same principles can be applied to the cone pigments.

Rhodopsin-Retinal Visual Cycle, and Excitation of the Rods

Rhodopsin and Its Decomposition by Light Energy. The outer segment of the rod that projects into the pigment layer of the retina has a concentration of about 40 percent of the light-sensitive pigment called *rhodopsin,* or *visual purple.* This substance is a combination of the protein *scotopsin* and the carotenoid pigment *retinal* (also called "retinene"). Furthermore, the retinal is a particular type called 11-*cis* retinal. This *cis* form of retinal is important because only this form can bind with scotopsin to synthesize rhodopsin.

When light energy is absorbed by rhodopsin, the rhodopsin begins to decompose within a very small fraction of a second, as shown at the top of Figure 50-5. The cause of this is photoactivation of electrons in the retinal portion of the rhodopsin, which leads to instantaneous change of the *cis* form of retinal into an all-*trans* form that still has the same chemical structure as the *cis* form but has a different physical structure—a straight molecule rather than an angulated molecule. Because the three-dimensional orientation of the reactive sites of the all-*trans* retinal no longer fits with the orientation of the reactive sites on the protein *scotopsin,* the all-*trans* retinal begins to pull away from the scotopsin. The immediate product is *bathorhodopsin,* which is a partially split combination of the all-*trans* retinal and scotopsin. Bathorhodopsin is extremely unstable and decays in nanoseconds to *lumirhodopsin.* This then decays in microseconds to *metarhodopsin I,* then in about a millisecond to *metarhodopsin II,* and finally, much more slowly

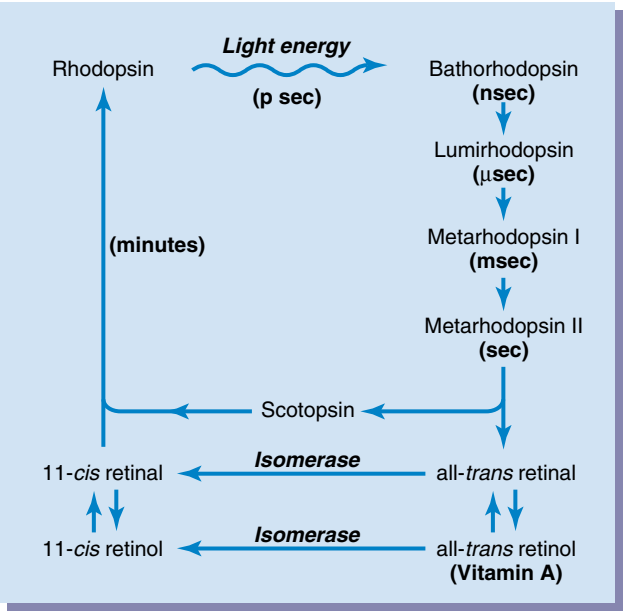


Figure 50-5 Rhodopsin-retinal visual cycle in the rod, showing decomposition of rhodopsin during exposure to light and subsequent slow re-formation of rhodopsin by the chemical processes.

(in seconds), into the completely split products *scotopsin* and all-*trans* retinal.

It is the metarhodopsin II, also called *activated rhodopsin,* that excites electrical changes in the rods, and the rods then transmit the visual image into the central nervous system in the form of optic nerve action potential, as we discuss later.

Re-formation of Rhodopsin. The first stage in re-formation of rhodopsin, as shown in Figure 50-5, is to reconvert the all-*trans* retinal into 11-*cis* retinal. This process requires metabolic energy and is catalyzed by the enzyme *retinal isomerase.* Once the 11-*cis* retinal is formed, it automatically recombines with the scotopsin to re-form rhodopsin, which then remains stable until its decomposition is again triggered by absorption of light energy.

Role of Vitamin A for Formation of Rhodopsin. Note in Figure 50-5 that there is a second chemical route by which all-*trans* retinal can be converted into 11-*cis* retinal. This is by conversion of the all-*trans* retinal first into all-*trans* retinol, which is one form of vitamin A. Then the all-*trans* retinol is converted into 11-*cis* retinol under the influence of the enzyme isomerase. Finally, the 11-*cis* retinol is converted into 11-*cis* retinal, which combines with scotopsin to form new rhodopsin.

Vitamin A is present both in the cytoplasm of the rods and in the pigment layer of the retina. Therefore, vitamin A is normally always available to form new retinal when needed. Conversely, when there is excess retinal in the retina, it is converted back into vitamin A, thus reducing the amount of light-sensitive pigment in the retina. We shall see later that this interconversion between retinal and vitamin A is especially important in long-term adaptation of the retina to different light intensities.

Night Blindness. Night blindness occurs in any person with severe vitamin A deficiency. The reason for this is that without vitamin A, the amounts of retinal and rhodopsin that can be formed are severely depressed. This condition is called *night blindness* because the amount of light available at night is too little to permit adequate vision in vitamin A–deficient persons.

For night blindness to occur, a person usually must remain on a vitamin A–deficient diet for months because large quantities of vitamin A are normally stored in the liver and can be made available to the eyes. Once night blindness develops, it can sometimes be reversed in less than 1 hour by intravenous injection of vitamin A.

Excitation of the Rod When Rhodopsin Is Activated by Light

The Rod Receptor Potential Is Hyperpolarizing, Not Depolarizing. When the rod is exposed to light, the resulting receptor potential is different from the receptor potentials in almost all other sensory receptors. That is, excitation of the rod causes *increased negativity* of the intrarod membrane potential, which is a state of *hyperpolarization,* meaning that there is more negativity than normal *inside* the rod membrane. This is exactly opposite to the decreased negativity (the process of "depolarization") that occurs in almost all other sensory receptors.

How does activation of rhodopsin cause hyperpolarization? The answer is that *when rhodopsin decomposes, it decreases the rod membrane conductance for sodium ions in the outer segment of the rod.* This causes hyperpolarization of the entire rod membrane in the following way.

Figure 50-6 shows movement of sodium and potassium ions in a complete electrical circuit through the inner and outer segments of the rod. The inner segment continually pumps sodium from inside the rod to the outside and potassium ions are pumped to the inside of the

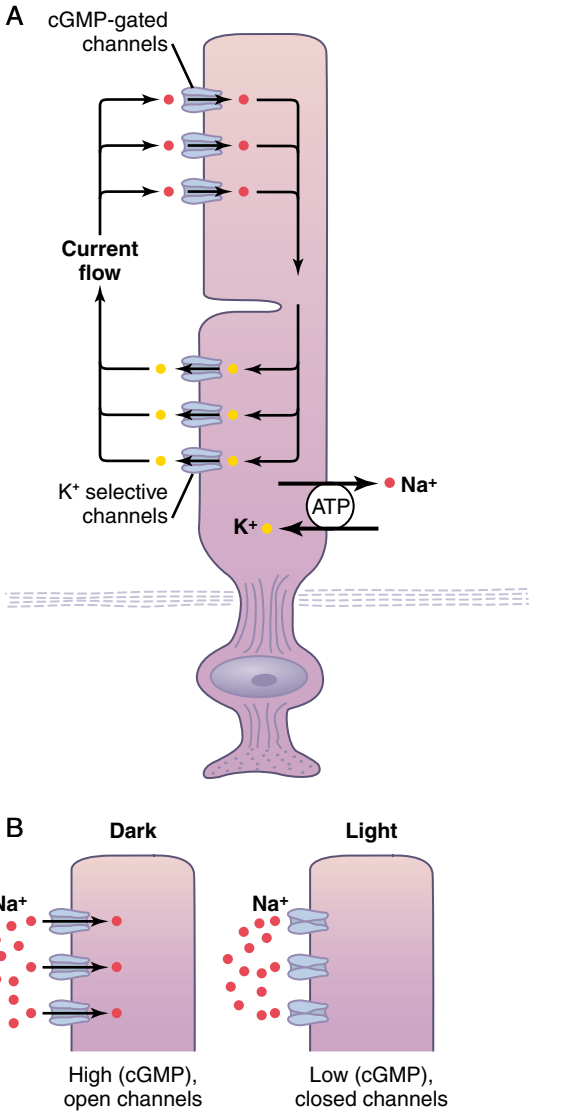


Figure 50-6 Sodium flows into a photoreceptor (e.g., rod) through cGMP-gated channels. Potassium flows out of the cell through nongated potassium channels. A sodium-potassium pump maintains steady levels of sodium and potassium inside the cell. In the dark, cGMP levels are high and the sodium channels are open. In the light, cGMP levels are reduced and the sodium channels close, causing the cell to hyperpolarize.

cell. Potassium ions leak out of the cell through nongated potassium channels that are confined to the inner segment of the rod. As in other cells, this sodium-potassium pump creates a negative potential on the inside of the entire cell. However, the outer segment of the rod, where the photoreceptor discs are located, is entirely different; here, the rod membrane, in the *dark* state, is leaky to sodium ions that flow through cGMP-gated channels. In the dark state, cGMP levels are high, permitting positively charged sodium ions to continually leak back to the inside of the rod and thereby neutralize much of the negativity on the inside of the entire cell. Thus, *under normal dark conditions, when the rod is not excited, there is reduced electronegativity* inside the membrane of the rod, measuring about −40 millivolts rather than the usual −70 to −80 millivolts found in most sensory receptors.

Then, when the rhodopsin in the outer segment of the rod is exposed to light, it is activated and begins to decompose, the cGMP gated sodium channels are closed, and the outer segment membrane conductance of sodium to the interior of the rod is reduced by a three-step process (Figure 50-7): (1) Light is absorbed by the rhodopsin, causing photoactivation of the electrons in the retinal portion, as previously described; (2) the activated rhodopsin stimulates a G-protein called *transducin,* which then activates cGMP phosphodiesterase; this enzyme catalyzes the breakdown of cGMP to 5′-cGMP; and (3) the reduction in cGMP closes the cGMP-gated sodium channels and reduces the inward sodium current. Sodium ions continue to be pumped outward through the membrane of the inner segment. Thus, more sodium ions now leave the rod than leak back in. Because they are positive ions, their loss from inside the rod creates increased negativity

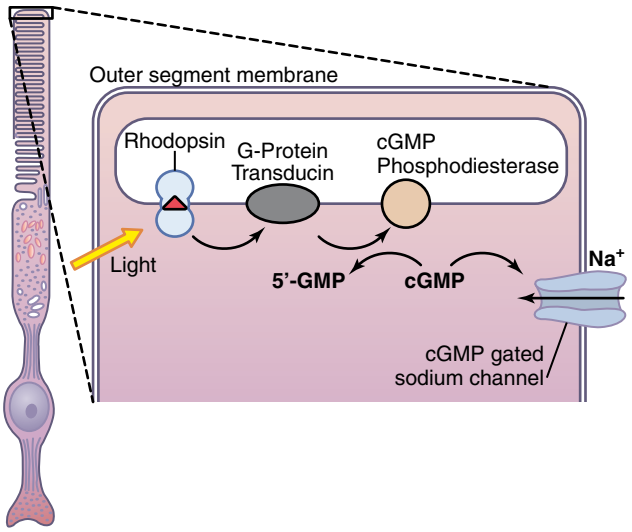


Figure 50-7 Phototransduction in the outer segment of the photoreceptor (rod or cone) membrane. When light hits the photoreceptor (e.g., rod cell), the light-absorbing retinal portion of rhodopsin is activated. This stimulates transducin, a G-protein, which then activates cGMP phosphodiesterase. This enzyme catalyzes the degradation of cGMP into 5′-GMP. The reduction in cGMP then causes closure of the sodium channels, which, in turn, causes hyperpolarization of the photoreceptor.

inside the membrane, and the greater the amount of light energy striking the rod, the greater the electronegativity becomes—that is, the greater is the degree of *hyperpolarization.* At maximum light intensity, the membrane potential approaches −70 to −80 millivolts, which is near the equilibrium potential for potassium ions across the membrane.

Duration of the Receptor Potential, and Logarithmic Relation of the Receptor Potential to Light Intensity. When a sudden pulse of light strikes the retina, the transient hyperpolarization that occurs in the rods—that is, the *receptor potential* that occurs—reaches a peak in about 0.3 second and lasts for more than a second. In cones, the change occurs four times as fast as in the rods. A visual image impinged on the rods of the retina for only one millionth of a second can sometimes cause the sensation of seeing the image for longer than a second.

Another characteristic of the receptor potential is that it is approximately proportional to the logarithm of the light intensity. This is exceedingly important because it allows the eye to discriminate light intensities through a range many thousand times as great as would be possible otherwise.

Mechanism by Which Rhodopsin Decomposition *Decreases* Membrane Sodium Conductance—The Excitation "Cascade." Under optimal conditions, a single photon of light, the smallest possible quantal unit of light energy, can cause a measurable receptor potential in a rod of about 1 millivolt. Only 30 photons of light will cause half saturation of the rod. How can such a small amount of light cause such great excitation? The answer is that the photoreceptors have an extremely sensitive chemical cascade that amplifies the stimulatory effects about a millionfold, as follows:

1. The *photon activates an electron* in the 11-*cis* retinal portion of the rhodopsin; this leads to the formation of *metarhodopsin II,* which is the active form of rhodopsin, as already discussed and shown in Figure 50-5.
2. The *activated rhodopsin* functions as an enzyme to activate many molecules of *transducin,* a protein present in an inactive form in the membranes of the discs and cell membrane of the rod.
3. The *activated transducin* activates many more molecules of *phosphodiesterase.*
4. *Activated phosphodiesterase* is another enzyme; it immediately hydrolyzes many molecules of *cyclic guanosine monophosphate* (cGMP), thus destroying it. Before being destroyed, the cGMP had been bound with the sodium channel protein of the rod's outer membrane in a way that "splints" it in the open state. But in light, when phosphodiesterase hydrolyzes the cGMP, this removes the splinting and allows the sodium channels to close. Several hundred channels close for each originally activated molecule of rhodopsin. Because the sodium flux through each of these channels has been extremely rapid, flow of more than a million sodium

ions is blocked by the channel closure before the channel opens again. This diminution of sodium ion flow is what excites the rod, as already discussed.

5. Within about a second, another enzyme, *rhodopsin kinase,* which is always present in the rod, inactivates the activated rhodopsin (the metarhodopsin II), and the entire cascade reverses back to the normal state with open sodium channels.

Thus, the rods have developed an important chemical cascade that amplifies the effect of a single photon of light to cause movement of millions of sodium ions. This explains the extreme sensitivity of the rods under dark conditions.

The cones are about 30 to 300 times less sensitive than the rods, but even this allows color vision at any intensity of light greater than extremely dim twilight.

Photochemistry of Color Vision by the Cones

It was pointed out at the outset of this discussion that the photochemicals in the cones have almost exactly the same chemical composition as that of rhodopsin in the rods. The only difference is that the protein portions, or the opsins—called *photopsins* in the cones—are slightly different from the scotopsin of the rods. The *retinal* portion of all the visual pigments is exactly the same in the cones as in the rods. The color-sensitive pigments of the cones, therefore, are combinations of retinal and photopsins.

In the discussion of color vision later in the chapter, it will become evident that only one of three types of color pigments is present in each of the different cones, thus making the cones selectively sensitive to different colors: blue, green, or red. These color pigments are called, respectively, *blue-sensitive pigment, green-sensitive pigment,* and *red-sensitive pigment.* The absorption characteristics of the pigments in the three types of cones show peak absorbencies at light wavelengths of 445, 535, and 570 nanometers, respectively. These are also the wavelengths for peak light sensitivity for each type of cone, which begins to explain how the retina differentiates the colors. The approximate absorption curves for these three pigments are shown in Figure 50-8. Also shown is the absorption curve for the rhodopsin of the rods, with a peak at 505 nanometers.

Automatic Regulation of Retinal Sensitivity—Light and Dark Adaptation

Light and Dark Adaptation. If a person has been in bright light for hours, large portions of the photochemicals in both the rods and the cones will have been reduced to retinal and opsins. Furthermore, much of the retinal of both the rods and the cones will have been converted into vitamin A. Because of these two effects, the concentrations of the photosensitive chemicals remaining in the rods and cones are considerably reduced, and the sensitivity of the eye to light is correspondingly reduced. This is called *light adaptation.*

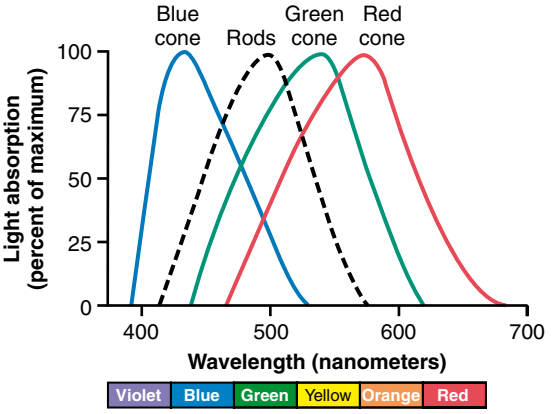


Figure 50-8 Light absorption by the pigment of the rods and by the pigments of the three color-receptive cones of the human retina. (Drawn from curves recorded by Marks WB, Dobelle WH, MacNichol EF Jr: Visual pigments of single primate cones. Science 143:1181, 1964, and by Brown PK, Wald G: Visual pigments in single rods and cones of the human retina: direct measurements reveal mechanisms of human night and color vision. Science 144:45, 1964. ©1964 by the American Association for the Advancement of Science.)

Conversely, if a person remains in darkness for a long time, the retinal and opsins in the rods and cones are converted back into the light-sensitive pigments. Furthermore, vitamin A is converted back into retinal to increase light-sensitive pigments, the final limit being determined by the amount of opsins in the rods and cones to combine with the retinal. This is called *dark adaptation.*

Figure 50-9 shows the course of dark adaptation when a person is exposed to total darkness after having been exposed to bright light for several hours. Note that the sensitivity of the retina is very low on first entering the darkness, but within 1 minute, the sensitivity has already increased 10-fold—that is, the retina can respond to light of one tenth the previously required intensity. At the end of 20 minutes, the sensitivity has increased about 6000-fold, and at the end of 40 minutes, about 25,000-fold.

The resulting curve of Figure 50-9 is called the *dark adaptation curve.* Note, however, the inflection in the

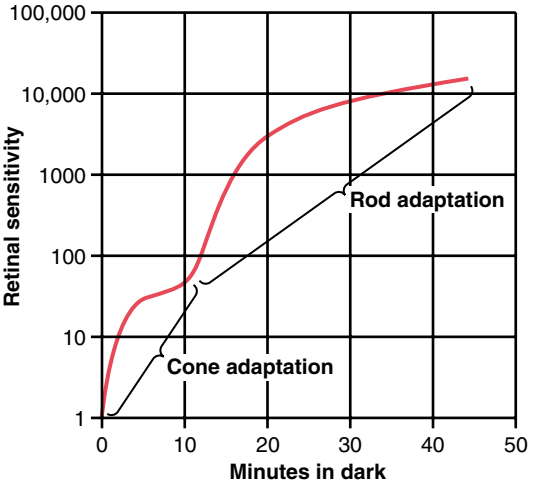


Figure 50-9 Dark adaptation, demonstrating the relation of cone adaptation to rod adaptation.

curve. The early portion of the curve is caused by adaptation of the cones because all the chemical events of vision, including adaptation, occur about four times as rapidly in cones as in rods. However, the cones do not achieve anywhere near the same degree of sensitivity change in darkness as the rods do. Therefore, despite rapid adaptation, the cones cease adapting after only a few minutes, while the slowly adapting rods continue to adapt for many minutes and even hours, their sensitivity increasing tremendously. In addition, still more sensitivity of the rods is caused by neuronal signal convergence of 100 or more rods onto a single ganglion cell in the retina; these rods summate to increase their sensitivity, as discussed later in the chapter.

Other Mechanisms of Light and Dark Adaptation. In addition to adaptation caused by changes in concentrations of rhodopsin or color photochemicals, the eye has two other mechanisms for light and dark adaptation. The first of these is a *change in pupillary size,* as discussed in Chapter 49. This can cause adaptation of approximately 30-fold within a fraction of a second because of changes in the amount of light allowed through the pupillary opening.

The other mechanism is *neural adaptation,* involving the neurons in the successive stages of the visual chain in the retina itself and in the brain. That is, when light intensity first increases, the signals transmitted by the bipolar cells, horizontal cells, amacrine cells, and ganglion cells are all intense. However, most of these signals decrease rapidly at different stages of transmission in the neural circuit. Although the degree of adaptation is only a fewfold rather than the many thousandfold that occurs during adaptation of the photochemical system, neural adaptation occurs in a fraction of a second, in contrast to the many minutes to hours required for full adaptation by the photochemicals.

Value of Light and Dark Adaptation in Vision. Between the limits of maximal dark adaptation and maximal light adaptation, the eye can change its sensitivity to light as much as 500,000 to 1 million times, the sensitivity automatically adjusting to changes in illumination.

Because registration of images by the retina requires detection of both dark and light spots in the image, it is essential that the sensitivity of the retina always be adjusted so that the receptors respond to the lighter areas but not to the darker areas. An example of maladjustment of retinal adaptation occurs when a person leaves a movie theater and enters the bright sunlight. Then, even the dark spots in the images seem exceedingly bright, and as a consequence, the entire visual image is bleached, having little contrast among its different parts. This is poor vision, and it remains poor until the retina has adapted sufficiently so that the darker areas of the image no longer stimulate the receptors excessively.

Conversely, when a person first enters darkness, the sensitivity of the retina is usually so slight that even the light spots in the image cannot excite the retina. After dark adaptation, the light spots begin to register. As an example of the extremes of light and dark adaptation, the intensity of sunlight is about 10 billion times that of starlight, yet the eye can function both in bright sunlight after light adaptation and in starlight after dark adaptation.

Color Vision

From the preceding sections, we have learned that different cones are sensitive to different colors of light. This section is a discussion of the mechanisms by which the retina detects the different gradations of color in the visual spectrum.

Tricolor Mechanism of Color Detection

All theories of color vision are based on the well-known observation that the human eye can detect almost all gradations of colors when only red, green, and blue monochromatic lights are appropriately mixed in different combinations.

Spectral Sensitivities of the Three Types of Cones. On the basis of color vision tests, the spectral sensitivities of the three types of cones in humans have proved to be essentially the same as the light absorption curves for the three types of pigment found in the cones. These curves are shown in Figure 50-8 and slightly differently in Figure 50-10. They can explain most of the phenomena of color vision.

Interpretation of Color in the Nervous System. Referring to Figure 50-10, one can see that an orange monochromatic light with a wavelength of 580 nanometers stimulates the red cones to a value of about 99 (99 percent of the peak stimulation at optimum wavelength); it stimulates the green cones to a value of about 42, but the blue cones not at all. Thus, the ratios of stimulation of the three types of cones in this instance are 99:42:0. The nervous system interprets this set of ratios as the sensation of orange. Conversely, a monochromatic blue light with a wavelength of 450 nanometers stimulates the red cones to a stimulus value of 0, the green cones to

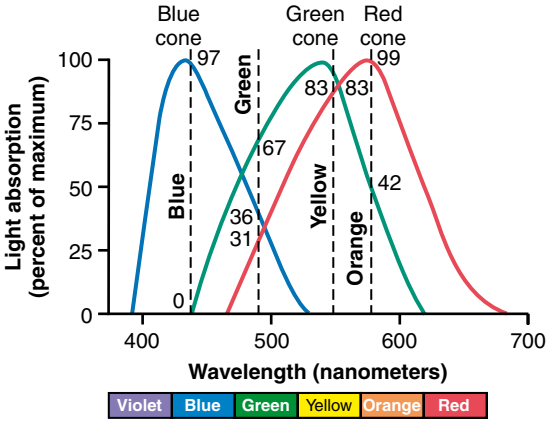


Figure 50-10 Demonstration of the degree of stimulation of the different color-sensitive cones by monochromatic lights of four colors: blue, green, yellow, and orange.

a value of 0, and the blue cones to a value of 97. This set of ratios—0:0:97—is interpreted by the nervous system as blue. Likewise, ratios of 83:83:0 are interpreted as yellow, and 31:67:36 as green.

Perception of White Light. About equal stimulation of all the red, green, and blue cones gives one the sensation of seeing white. Yet there is no single wavelength of light corresponding to white; instead, white is a combination of all the wavelengths of the spectrum. Furthermore, the perception of white can be achieved by stimulating the retina with a proper combination of only three chosen colors that stimulate the respective types of cones about equally.

Color Blindness

Red-Green Color Blindness. When a single group of color-receptive cones is missing from the eye, the person is unable to distinguish some colors from others. For instance, one can see in Figure 50-10 that green, yellow, orange, and red colors, which are the colors between the wavelengths of 525 and 675 nanometers, are normally distinguished from one another by the red and green cones. If either of these two cones is missing, the person cannot use this mechanism for distinguishing these four colors; the person is especially unable to distinguish red from green and is therefore said to have *red-green color blindness.*

A person with loss of red cones is called a *protanope;* the overall visual spectrum is noticeably shortened at the long wavelength end because of a lack of the red cones. A color-blind person who lacks green cones is called a *deuteranope;* this person has a perfectly normal visual spectral width because red cones are available to detect the long wavelength red color.

Red-green color blindness is a genetic disorder that occurs almost exclusively in males. That is, genes in the female X chromosome code for the respective cones. Yet color blindness almost never occurs in females because at least one of the two X chromosomes almost always has a normal gene for each type of cone. Because the male has only one X chromosome, a missing gene can lead to color blindness.

Because the X chromosome in the male is always inherited from the mother, never from the father, color blindness is passed from mother to son, and the mother is said to be a *color blindness carrier;* this is true in about 8 percent of all women.

Blue Weakness. Only rarely are blue cones missing, although sometimes they are underrepresented, which is a genetically inherited state giving rise to the phenomenon called *blue weakness.*

Color Test Charts. A rapid method for determining color blindness is based on the use of spot charts such as those shown in Figure 50-11. These charts are arranged with a confusion of spots of several different colors. In the top chart, the person with normal color vision reads "74," whereas the red-green color-blind person reads "21." In the bottom chart, the person with normal color vision reads "42," whereas the red-blind person reads "2," and the green-blind person reads "4."

If one studies these charts while at the same time observing the spectral sensitivity curves of the different cones depicted in Figure 50-10, it can be readily understood how excessive emphasis can be placed on spots of certain colors by color-blind people.

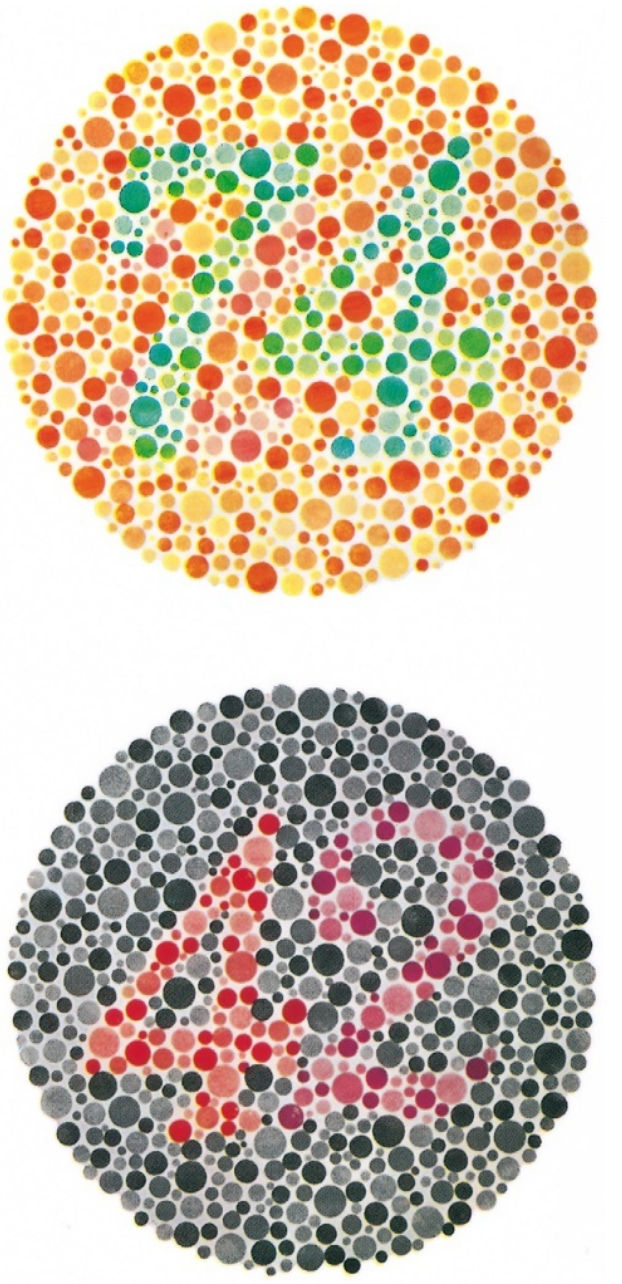


Figure 50-11 Two Ishihara charts. *Upper:* In this chart, the normal person reads "74," but the red-green color-blind person reads "21." *Lower:* In this chart, the red-blind person (protanope) reads "2," but the green-blind person (deuteranope) reads "4." The normal person reads "42." (Reproduced from Ishihara's Tests for Colour Blindness. Tokyo: Kanehara & Co., but tests for color blindness cannot be conducted with this material. For accurate testing, the original plates should be used.)

Neural Function of the Retina

Neural Circuitry of the Retina

Figure 50-12 presents the essentials of the retina's neural connections, showing at the left the circuit in the peripheral retina and at the right the circuit in the foveal retina. The different neuronal cell types are as follows:

1. The photoreceptors themselves—the *rods* and *cones*—which transmit signals to the outer plexiform

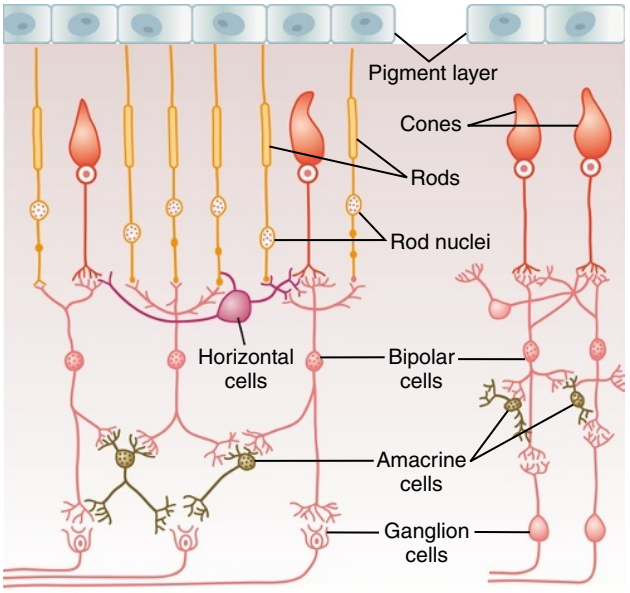


Figure 50-12 Neural organization of the retina: peripheral area to the left, foveal area to the right.

layer, where they synapse with bipolar cells and horizontal cells

2. The *horizontal cells,* which transmit signals horizontally in the outer plexiform layer from the rods and cones to bipolar cells
3. The *bipolar cells,* which transmit signals vertically from the rods, cones, and horizontal cells to the inner plexiform layer, where they synapse with ganglion cells and amacrine cells
4. The *amacrine cells,* which transmit signals in two directions, either directly from bipolar cells to ganglion cells or horizontally within the inner plexiform layer from axons of the bipolar cells to dendrites of the ganglion cells or to other amacrine cells
5. The *ganglion cells,* which transmit output signals from the retina through the optic nerve into the brain

A sixth type of neuronal cell in the retina, not very prominent and not shown in the figure, is the *interplexiform* cell. This cell transmits signals in the retrograde direction from the inner plexiform layer to the outer plexiform layer. These signals are inhibitory and are believed to control lateral spread of visual signals by the horizontal cells in the outer plexiform layer. Their role may be to help control the degree of contrast in the visual image.

The Visual Pathway from the Cones to the Ganglion Cells Functions Differently from the Rod Pathway. As is true for many of our other sensory systems, the retina has both an old type of vision based on rod vision and a new type of vision based on cone vision. The neurons and nerve fibers that conduct the visual signals for cone vision are considerably larger than those that conduct the visual signals for rod vision, and the signals are conducted to the brain two to five times as rapidly. Also, the circuitry for the two systems is slightly different, as follows.

To the right in Figure 50-12 is the visual pathway from the *foveal portion of the retina,* representing the new, fast cone system. This shows three neurons in the direct pathway: (1) cones, (2) bipolar cells, and (3) ganglion cells. In addition, horizontal cells transmit inhibitory signals laterally in the outer plexiform layer, and amacrine cells transmit signals laterally in the inner plexiform layer.

To the left in Figure 50-12 are the neural connections for the peripheral retina, where both rods and cones are present. Three bipolar cells are shown; the middle of these connects only to rods, representing the type of visual system present in many lower animals. The output from the bipolar cell passes only to amacrine cells, which relay the signals to the ganglion cells. Thus, for pure rod vision, there are four neurons in the direct visual pathway: (1) rods, (2) bipolar cells, (3) amacrine cells, and (4) ganglion cells. Also, horizontal and amacrine cells provide lateral connectivity.

The other two bipolar cells shown in the peripheral retinal circuitry of Figure 50-12 connect with both rods and cones; the outputs of these bipolar cells pass both directly to ganglion cells and by way of amacrine cells.

Neurotransmitters Released by Retinal Neurons. Not all the neurotransmitter chemical substances used for synaptic transmission in the retina have been entirely delineated. However, both the rods and the cones release *glutamate* at their synapses with the bipolar cells.

Histological and pharmacological studies have proven there are many types of amacrine cells secreting at least eight types of transmitter substances, including *gamma-aminobutyric acid, glycine, dopamine, acetylcholine,* and *indolamine,* all of which normally function as inhibitory transmitters. The transmitters of the bipolar, horizontal, and interplexiform cells are unclear, but at least some of the horizontal cells release inhibitory transmitters.

Transmission of Most Signals Occurs in the Retinal Neurons by Electrotonic Conduction, Not by Action Potentials. The only retinal neurons that always transmit visual signals by means of action potentials are the ganglion cells, and they send their signals all the way to the brain through the optic nerve. Occasionally, action potentials have also been recorded in amacrine cells, although the importance of these action potentials is questionable. Otherwise, all the retinal neurons conduct their visual signals by *electrotonic conduction,* which can be explained as follows.

Electrotonic conduction means direct flow of electric current, not action potentials, in the neuronal cytoplasm and nerve axons from the point of excitation all the way to the output synapses. Even in the rods and cones, conduction from their outer segments, where the visual signals are generated, to the synaptic bodies is by electrotonic conduction. That is, when hyperpolarization occurs in response to light in the outer segment of a rod or a cone,

almost the same degree of hyperpolarization is conducted by direct electric current flow in the cytoplasm all the way to the synaptic body, and no action potential is required. Then, when the transmitter from a rod or cone stimulates a bipolar cell or horizontal cell, once again the signal is transmitted from the input to the output by direct electric current flow, not by action potentials.

The importance of electrotonic conduction is that it allows *graded conduction* of signal strength. Thus, for the rods and cones, the strength of the hyperpolarizing output signal is directly related to the intensity of illumination; the signal is not all or none, as would be the case for each action potential.

Lateral Inhibition to Enhance Visual Contrast—Function of the Horizontal Cells

The horizontal cells, shown in Figure 50-12, connect laterally between the synaptic bodies of the rods and cones, as well as connecting with the dendrites of the bipolar cells. The outputs of the horizontal cells *are always inhibitory.* Therefore, this lateral connection provides the same phenomenon of lateral inhibition that is important in all other sensory systems—that is, helping to ensure transmission of visual patterns with proper visual contrast. This phenomenon is demonstrated in Figure 50-13, which shows a minute spot of light focused on the retina. The visual pathway from the central most area where the light strikes is excited, whereas an area to the side is inhibited. In other words, instead of the excitatory signal spreading widely in the retina because of spreading dendritic and axonal trees in the plexiform layers, transmission through the horizontal cells puts a stop to this by providing lateral inhibition in the surrounding areas. This is essential to allow high visual accuracy in transmitting contrast borders in the visual image.

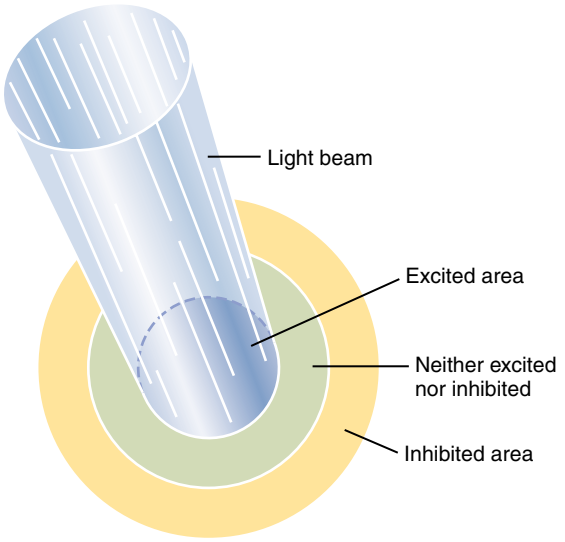


Figure 50-13 Excitation and inhibition of a retinal area caused by a small beam of light, demonstrating the principle of lateral inhibition.

Some of the amacrine cells probably provide additional lateral inhibition and further enhancement of visual contrast in the inner plexiform layer of the retina as well.

Excitation of Some Bipolar Cells and Inhibition of Others—The Depolarizing and Hyperpolarizing Bipolar Cells

Two types of bipolar cells provide opposing excitatory and inhibitory signals in the visual pathway: (1) the *depolarizing bipolar cell* and (2) the *hyperpolarizing bipolar cell.* That is, some bipolar cells depolarize when the rods and cones are excited, and others hyperpolarize.

There are two possible explanations for this difference. One explanation is that the two bipolar cells are of entirely different types—one responding by depolarizing in response to the glutamate neurotransmitter released by the rods and cones, and the other responding by hyperpolarizing. The other possibility is that one of the bipolar cells receives direct excitation from the rods and cones, whereas the other receives its signal indirectly through a horizontal cell. Because the horizontal cell is an inhibitory cell, this would reverse the polarity of the electrical response.

Regardless of the mechanism for the two types of bipolar responses, the importance of this phenomenon is that it allows half the bipolar cells to transmit positive signals and the other half to transmit negative signals. We shall see later that both positive and negative signals are used in transmitting visual information to the brain.

Another important aspect of this reciprocal relation between depolarizing and hyperpolarizing bipolar cells is that it provides a second mechanism for lateral inhibition, in addition to the horizontal cell mechanism. Because depolarizing and hyperpolarizing bipolar cells lie immediately against each other, this provides a mechanism for separating contrast borders in the visual image, even when the border lies exactly between two adjacent photoreceptors. In contrast, the horizontal cell mechanism for lateral inhibition operates over a much greater distance.

Amacrine Cells and Their Functions

About 30 types of amacrine cells have been identified by morphological or histochemical means. The functions of about half a dozen types of amacrine cells have been characterized, and all of them are different. One type of amacrine cell is part of the direct pathway for rod vision—that is, from rod to bipolar cells to amacrine cells to ganglion cells.

Another type of amacrine cell responds strongly at the onset of a continuing visual signal, but the response dies rapidly.

Other amacrine cells respond strongly at the offset of visual signals, but again, the response fades quickly.

Still other amacrine cells respond when a light is turned either on or off, signaling simply a change in illumination, irrespective of direction.

Still another type of amacrine cell responds to movement of a spot across the retina in a specific direction;

therefore, these amacrine cells are said to be *directional sensitive.*

In a sense, then, many or most amacrine cells are interneurons that help analyze visual signals before they ever leave the retina.

Ganglion Cells and Optic Nerve Fibers

Each retina contains about 100 million rods and 3 million cones; yet the number of ganglion cells is only about 1.6 million. Thus, an average of 60 rods and 2 cones converge on each ganglion cell and the optic nerve fiber leading from the ganglion cell to the brain.

However, major differences exist between the peripheral retina and the central retina. As one approaches the fovea, fewer rods and cones converge on each optic fiber, and the rods and cones also become more slender. These effects progressively increase the acuity of vision in the central retina. In the center, in the *central fovea,* there are only slender cones—about 35,000 of them—and no rods. Also, the number of optic nerve fibers leading from this part of the retina is almost exactly equal to the number of cones, as shown to the right in Figure 50-12. This explains the high degree of visual acuity in the central retina in comparison with the much poorer acuity peripherally.

Another difference between the peripheral and central portions of the retina is the much greater sensitivity of the peripheral retina to weak light. This results partly from the fact that rods are 30 to 300 times more sensitive to light than cones are, but it is further magnified by the fact that as many as 200 rods converge on a single optic nerve fiber in the more peripheral portions of the retina, so signals from the rods summate to give even more intense stimulation of the peripheral ganglion cells and their optic nerve fibers.

Three Types of Retinal Ganglion Cells and Their Respective Fields

There are three distinct types of ganglion cells, designated W, X, and Y cells. Each of these serves a different function.

Transmission of Rod Vision by the W Cells. The W cells, constituting about 40 percent of all the ganglion cells, are small, having a diameter less than 10 micrometers, and they transmit signals in their optic nerve fibers at the slow velocity of only 8 m/sec. These ganglion cells receive most of their excitation from rods, transmitted by way of small bipolar cells and amacrine cells. They have broad fields in the peripheral retina because the dendrites of the ganglion cells spread widely in the inner plexiform layer, receiving signals from broad areas.

On the basis of histology, as well as physiological experiments, the W cells seem to be especially sensitive for detecting directional movement in the field of vision, and they are probably important for much of our crude rod vision under dark conditions.

Transmission of the Visual Image and Color by the X Cells. The most numerous of the ganglion cells are the X cells, representing 55 percent of the total. They are of medium diameter, between 10 and 15 micrometers, and transmit signals in their optic nerve fibers at about 14 m/sec.

The X cells have small fields because their dendrites do not spread widely in the retina. Because of this, their signals represent discrete retinal locations. Therefore, it is mainly through the X cells that the fine details of the visual image are transmitted. Also, because every X cell receives input from at least one cone, X cell transmission is probably responsible for all color vision.

Function of the Y Cells to Transmit Instantaneous Changes in the Visual Image. The Y cells are the largest of all, up to 35 micrometers in diameter, and they transmit their signals to the brain at 50 m/sec or faster. They are the least numerous of all the ganglion cells, representing only 5 percent of the total. Also, they have broad dendritic fields, so signals are picked up by these cells from widespread retinal areas.

The Y ganglion cells respond, like many of the amacrine cells, to rapid changes in the visual image—either rapid movement or rapid change in light intensity—sending bursts of signals for only small fractions of a second. These ganglion cells presumably apprise the central nervous system almost instantaneously when a new visual event occurs anywhere in the visual field, but without specifying with great accuracy the location of the event, other than to give appropriate clues that make the eyes move toward the exciting vision.

Excitation of the Ganglion Cells

Spontaneous, Continuous Action Potentials in the Ganglion Cells. It is from the ganglion cells that the long fibers of the optic nerve lead into the brain. Because of the distance involved, the electrotonic method of conduction employed in the rods, cones, and bipolar cells within the retina is no longer appropriate; therefore, ganglion cells transmit their signals by means of repetitive action potentials instead. Furthermore, even when unstimulated, they still transmit continuous impulses at rates varying between 5 and 40 per second. The visual signals, in turn, are superimposed onto this background ganglion cell firing.

Transmission of Changes in Light Intensity—The On-Off Response. As noted previously, many ganglion cells are specifically excited by *changes* in light intensity. This is demonstrated by the records of nerve impulses in Figure 50-14. The upper panel shows rapid impulses for a fraction of a second when a light is first turned on, but decreasing rapidly in the next fraction of a second. The lower tracing is from a ganglion cell located lateral to the spot of light; this cell is markedly inhibited when the light is turned on because of lateral inhibition. Then, when the light is turned off, opposite effects occur. Thus, these records are called "on-off" and "off-on" responses. The opposite directions of these responses to light are caused, respectively, by the depolarizing and hyperpolarizing bipolar cells, and

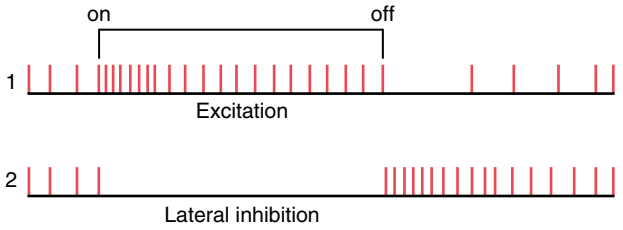


Figure 50-14 Responses of a ganglion cell to light in (*1*) an area excited by a spot of light and (*2*) an area adjacent to the excited spot; the ganglion cell in this area is inhibited by the mechanism of *lateral inhibition*. (Modified from Granit R: Receptors and Sensory Perception: A Discussion of Aims, Means, and Results of Electrophysiological Research into the Process of Reception. New Haven, Conn: Yale University Press, 1955.)

the transient nature of the responses is probably at least partly generated by the amacrine cells, many of which have similar transient responses themselves.

This capability of the eyes to detect *change* in light intensity is strongly developed in both the peripheral retina and the central retina. For instance, a minute gnat flying across the field of vision is instantaneously detected. Conversely, the same gnat sitting quietly remains below the threshold of visual detection.

Transmission of Signals Depicting Contrasts in the Visual Scene—The Role of Lateral Inhibition

Many ganglion cells respond mainly to contrast borders in the scene. Because this seems to be the major means by which the pattern of a scene is transmitted to the brain, let us explain how this process occurs.

When flat light is applied to the entire retina—that is, when all the photoreceptors are stimulated equally by the incident light—the contrast type of ganglion cell is neither stimulated nor inhibited. The reason for this is that signals transmitted *directly* from the photoreceptors through depolarizing bipolar cells are excitatory, whereas the signals transmitted *laterally* through hyperpolarizing bipolar cells, as well as through horizontal cells, are mainly inhibitory. Thus, the direct excitatory signal through one pathway is likely to be neutralized by inhibitory signals through lateral pathways. One circuit for this is demonstrated in Figure 50-15, which shows at the top three photoreceptors. The central receptor excites a depolarizing bipolar cell. The two receptors on each side are connected to the same bipolar cell through inhibitory horizontal cells that neutralize the direct excitatory signal if all three receptors are stimulated simultaneously by light.

Now, let us examine what happens when a contrast border occurs in the visual scene. Referring again to Figure 50-15, assume that the central photoreceptor is stimulated by a bright spot of light while one of the two lateral receptors is in the dark. The bright spot of light excites the direct pathway through the bipolar cell. The fact that one of the lateral photoreceptors is in the dark causes one of the horizontal cells to remain unstimulated. Therefore, this cell does not inhibit the bipolar cell, and this allows extra excitation of the bipolar cell. Thus, where

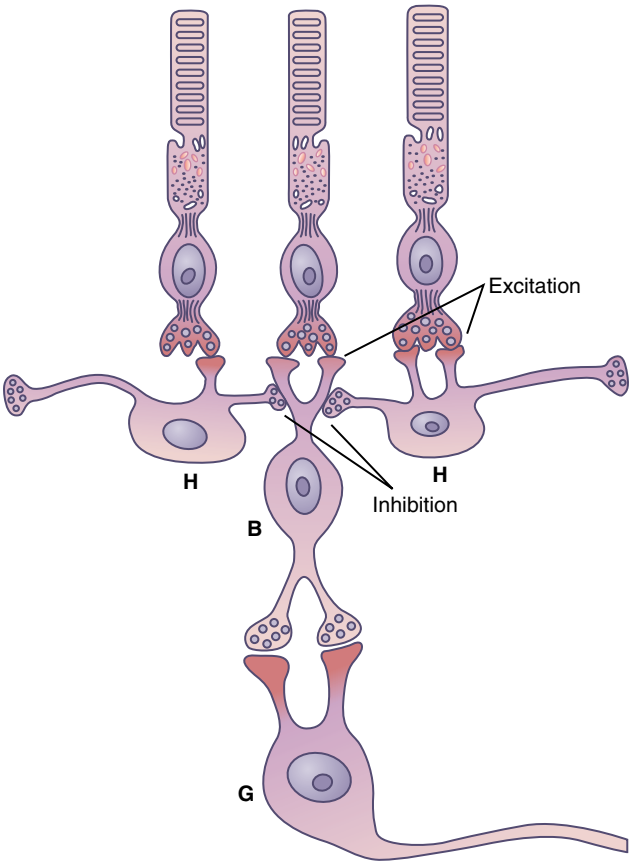


Figure 50-15 Typical arrangement of rods, horizontal cells (H), a bipolar cell (B), and a ganglion cell (G) in the retina, showing excitation at the synapses between the rods and the bipolar cell and horizontal cells, but inhibition from the horizontal cells to the bipolar cell.

visual contrasts occur, the signals through the direct and lateral pathways accentuate one another.

In summary, the mechanism of lateral inhibition functions in the eye in the same way that it functions in most other sensory systems—to provide contrast detection and enhancement.

Transmission of Color Signals by the Ganglion Cells

A single ganglion cell may be stimulated by several cones or by only a few. When all three types of cones—the red, blue, and green types—stimulate the same ganglion cell, the signal transmitted through the ganglion cell is the same for any color of the spectrum. Therefore, the signal from the ganglion cell plays no role in the detection of different colors. Instead, it is a "white" signal.

Conversely, some of the ganglion cells are excited by only one color type of cone but inhibited by a second type. For instance, this frequently occurs for the red and green cones, with red causing excitation and green causing inhibition, or vice versa.

The same type of reciprocal effect occurs between blue cones on the one hand and a combination of red and green cones (both of which are excited by yellow) on the other hand, giving a reciprocal excitation-inhibition relation between the blue and yellow colors.

The mechanism of this opposing effect of colors is the following: One color type of cone excites the ganglion cell by the direct excitatory route through a depolarizing bipolar cell, whereas the other color type inhibits the ganglion cell by the indirect inhibitory route through a hyperpolarizing bipolar cell.

The importance of these color-contrast mechanisms is that they represent a means by which the retina itself begins to differentiate colors. Thus, each color-contrast type of ganglion cell is excited by one color but inhibited by the "opponent" color. Therefore, color analysis begins in the retina and is not entirely a function of the brain.

Bibliography

Artemyev NO: Light-dependent compartmentalization of transducin in rod photoreceptors, *Mol Neurobiol* 37:44, 2008.

Bloomfield SA, Völgyi B: The diverse functional roles and regulation of neuronal gap junctions in the retina, *Nat Rev Neurosci* 10:495, 2009.

Bowmaker JK: Evolution of vertebrate visual pigments, *Vision Res* 48:2022, 2008.

Carroll J: Focus on molecules: the cone opsins, *Exp Eye Res* 86:865, 2008.

D'Amico DJ: Clinical practice. Primary retinal detachment, *N Engl J Med* 359:2346, 2008.

Fain GL, Matthews HR, Cornwall MC, Koutalos Y: Adaptation in vertebrate photoreceptors, *Physiol Rev* 81:117, 2001.

Garriga P, Manyosa J: The eye photoreceptor protein rhodopsin: structural implications for retinal disease, *FEBS Lett* 528:17, 2002.

Gegenfurtner KR: Cortical mechanisms of colour vision, *Nat Rev Neurosci* 4:563, 2003.

Gegenfurtner KR, Kiper DC: Color vision, *Annu Rev Neurosci* 26:181, 2003.

Hankins MW, Peirson SN, Foster RG: Melanopsin: an exciting photopigment, *Trends Neurosci* 31:27, 2008.

Hardie RC: Phototransduction: shedding light on translocation, *Curr Biol* 13:R775, 2003.

Hartzell HC, Qu Z, Yu K, et al: Molecular physiology of bestrophins: multifunctional membrane proteins linked to Best disease and other retinopathies, *Physiol Rev* 88:639, 2008.

Kandel ER, Schwartz JH, Jessell TM: *Principles of Neural Science*, 4th ed, New York, 2000, McGraw-Hill.

Kolb H, Nelson R, Ahnelt P, Cuenca N: Cellular organization of the vertebrate retina, *Prog Brain Res* 131:3, 2001.

Luo DG, Xue T, Yau KW: How vision begins: an odyssey, *Proc Natl Acad Sci U S A* 105:9855, 2008.

Masland RH: The fundamental plan of the retina, *Nat Neurosci* 4:877, 2001.

Okawa H, Sampath AP: Optimization of single-photon response transmission at the rod-to-rod bipolar synapse, *Physiology (Bethesda)* 22:279, 2007.

Schwartz EA: Transport-mediated synapses in the retina, *Physiol Rev* 82:875, 2002.

Solomon SG, Lennie P: The machinery of colour vision, *Nat Rev Neurosci* 8:276, 2007.

Taylor WR, Vaney DI: New directions in retinal research, *Trends Neurosci* 26:379, 2003.

Wensel TG: Signal transducing membrane complexes of photoreceptor outer segments, *Vision Res* 48:2052, 2008.

Westheimer G: The ON-OFF dichotomy in visual processing: from receptors to perception, *Prog Retin Eye Res* 26:636, 2007.