Crystal begins her run along the beach just as the sun is rising over the ocean. She loves this time of day, because it is cool and the mist rising from the sand creates a mystical effect. She looks down the beach and notices something about 100 yards away that wasn't there yesterday. "What an interesting piece of driftwood," she thinks, although it is difficult to see because of the mist and dim lighting (**Figure 3.1a**). As she approaches the object, she begins to doubt her initial perception, and just as she is wondering whether it might not be driftwood, she realizes that it is, in fact, the old beach umbrella that was lying under the lifeguard stand yesterday (**Figure 3.1b**). "Driftwood transformed into an umbrella, right before my eyes," she thinks.
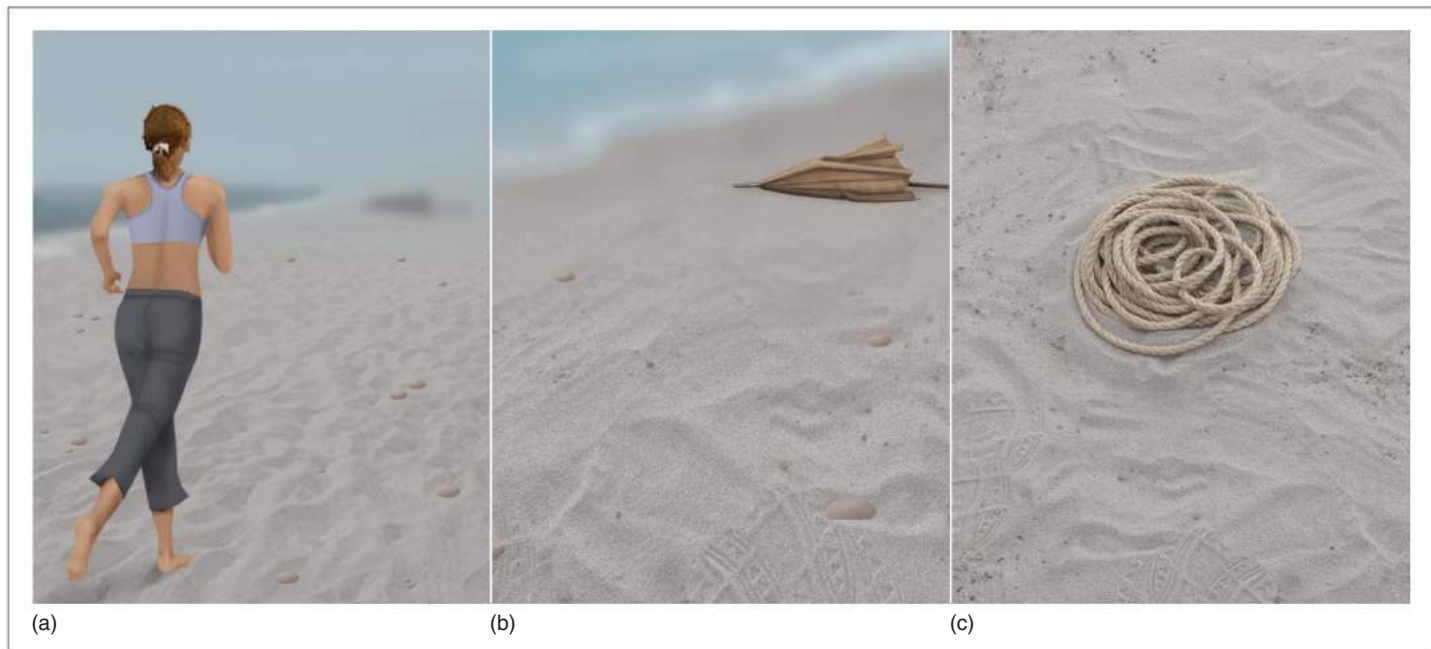
Continuing down the beach, she passes some coiled rope that appears to be abandoned (**Figure 3.1c**). She stops to check it out. Grabbing one end, she flips the rope and sees that, as she suspected, it is one continuous strand. But she needs to keep running, because she is supposed to meet a friend at Beach Java, a coffee shop far down the beach. Later, sitting in the coffeehouse, she tells her friend about the piece of magic driftwood that was transformed into an umbrella.

▶ The Nature of Perception

We define **perception** as experiences resulting from stimulation of the senses. To appreciate how these experiences are created, let's return to Crystal on the beach.

Some Basic Characteristics of Perception

Crystal's experiences illustrate a number of things about perception. Her experience of seeing what she thought was driftwood turn into an umbrella illustrates how perceptions

▶ **Figure 3.1** (a) Initially Crystal thinks she sees a large piece of driftwood far down the beach. (b) Eventually she realizes she is looking at an umbrella. (c) On her way down the beach, she passes some coiled rope.

Bruce Goldstein

can change based on added information (Crystal's view became better as she got closer to the umbrella) and how perception can involve a process similar to reasoning or problem solving (Crystal figured out what the object was based partially on remembering having seen the umbrella the day before). (Another example of an initially erroneous perception followed by a correction is the famous pop culture line, "It's a bird. It's a plane. It's Superman!") Crystal's guess that the coiled rope was continuous illustrates how perception can be based on a perceptual rule (*when objects overlap, the one underneath usually continues behind the one on top*), which may be based on the person's past experiences.

Crystal's experience also demonstrates how arriving at a perception can involve a *process*. It took some time for Crystal to realize that what she thought was driftwood was actually an umbrella, so it is possible to describe her perception as involving a "reasoning" process. In most cases, perception occurs so rapidly and effortlessly that it appears to be automatic. But, as we will see in this chapter, perception is far from automatic. It involves complex, and usually invisible, processes that resemble reasoning, although they occur much more rapidly than Crystal's realization that the driftwood was actually an umbrella.

Finally, Crystal's experience also illustrates how perception occurs in conjunction with action. Crystal is running and perceiving at the same time; later, at the coffee shop, she easily reaches for her cup of coffee, a process that involves coordination of seeing the coffee cup, determining its location, physically reaching for it, and grasping its handle. This aspect of Crystal's experiences is what happens in everyday perception. We are usually moving, and even when we are just sitting in one place watching TV, a movie, or a sporting event, our eyes are constantly in motion as we shift our attention from one thing to another to perceive what is happening. We also grasp and pick up things many times a day, whether it is a cup of coffee, a phone, or this book. Perception, therefore, is more than just "seeing" or "hearing." It is central to our ability to organize the actions that occur as we interact with the environment.

It is important to recognize that while perception creates a picture of our environment and helps us take action within it, it also plays a central role in cognition in general. When we consider that perception is essential for creating memories, acquiring knowledge, solving problems, communicating with other people, recognizing someone you met last week, and answering questions on a cognitive psychology exam, it becomes clear that perception is the gateway to all the other cognitions we will be describing in this book.

The goal of this chapter is to explain the mechanisms responsible for perception. To begin, we move from Crystal's experience on the beach and in the coffee shop to what happens when perceiving a city scene: Pittsburgh as seen from the upper deck of PNC Park, home of the Pittsburgh Pirates.

A Human Perceives Objects and a Scene

Sitting in the upper deck of PNC Park, Roger looks out over the city (**Figure 3.2**). He sees a group of about 10 buildings on the left and can easily tell one building from another. Looking straight ahead, he sees a small building in front of a larger one, and has no trouble telling that they are two separate buildings. Looking down toward the river, he notices a horizontal yellow band above the right field bleachers. It is obvious to him that this is not part of the ballpark but is located across the river.

All of Roger's perceptions come naturally to him and require little effort. But when we look closely at the scene, it becomes apparent that the scene poses many "puzzles." The following demonstration points out a few of them.

➤ **Figure 3.2** It is easy to tell that there are a number of different buildings on the left and that straight ahead there is a low rectangular building in front of a taller building. It is also possible to tell that the horizontal yellow band above the bleachers is across the river. These perceptions are easy for humans but would be quite difficult for a computer vision system. The letters on the left indicate areas referred to in the Demonstration.

DEMONSTRATION   Perceptual Puzzles in a Scene

The following questions refer to the areas labeled in Figure 3.2. Your task is to answer each question and indicate the reasoning behind each answer:

What is the dark area at A?

Are the surfaces at B and C facing in the same or different directions?

Are areas B and C on the same building or on different buildings?

Does the building at D extend behind the one at A?

Although it may have been easy to answer the questions, it was probably somewhat more challenging to indicate what your "reasoning" was. For example, how did you know the dark area at A is a shadow? It could be a dark-colored building that is in front of a light-colored building. On what basis might you have decided that building D extends behind building A? It could, after all, simply end right where A begins. We could ask similar questions about everything in this scene because, as we will see, a particular pattern of shapes can be created by a wide variety of objects.

One of the messages of this demonstration is that to determine what is "out there," it is necessary to go beyond the pattern of light and dark that a scene creates on the retina—the

structure that lines the back of the eye and contains the receptors for seeing. One way to appreciate the importance of this "going beyond" process is to consider how difficult it has been to program even the most powerful computers to accomplish perceptual tasks that humans achieve with ease.

A Computer-Vision System Perceives Objects and a Scene

A computer that can perceive has been a dream that dates back to early science fiction and movies. Because movies can make up things, it was easy to show the droids R2-D2 and C3PO having a conversation on the desert planet Tatooine in the original *Star Wars* (1977). Although C3PO did most of the talking (R2D2 mainly beeped), both could apparently navigate through their environment with ease, and recognize objects along the way.

But designing a computer vision system that can actually perceive the environment and recognize objects and scenes is more complicated than making a *Star Wars* movie. In the 1950s, when digital computers became available to researchers, it was thought that it would take perhaps a decade to design a machine-vision system that would rival human vision. But the early systems were primitive and took minutes of calculations to identify simple isolated objects that a young child could name in seconds. Perceiving objects and scenes was, the researchers realized, still the stuff of science fiction.

It wasn't until 1987 that the *International Journal of Computer Vision*, the first journal devoted solely to computer vision, was founded. Papers from the first issues considered topics such as how to interpret line drawings of curved objects (Malik, 1987) and how to determine the three-dimensional layout of a scene based on a film of movement through the scene (Bolles et al., 1987). These papers and others in the journal had to resort to complex mathematical formulas to solve perceptual problems that are easy for humans.

Flash-forward to March 13, 2004. Thirteen robotic vehicles are lined up in the Mojave Desert in California for the Defense Advanced Projects Agency's (DARPA) Grand Challenge. The task was to drive 150 miles from the starting point to Las Vegas, using only GPS coordinates to define the course and computer vision to avoid obstacles. The best performance was achieved by a vehicle entered by Carnegie-Mellon University, which traversed only 7.3 miles before getting stuck.

Progress continued through the next decade, however, with thousands of researchers and multi-million-dollar investments, until now, when driverless cars are no longer a novelty. As I write this, a fleet of driverless Uber vehicles are finding their way around the winding streets of Pittsburgh, San Francisco, and other cities (**Figure 3.3**).

One message of the preceding story is that although present accomplishments of computer-vision systems are impressive, it turned out to be extremely difficult to create the systems that made driverless cars possible. But as impressive as driverless cars are, computer-vision systems still make mistakes in naming objects. For example, **Figure 3.4** shows three objects that a computer identified as a tennis ball.


iStock.com/JasonDoiy

➤ **Figure 3.3** Driverless car on the streets of San Francisco.
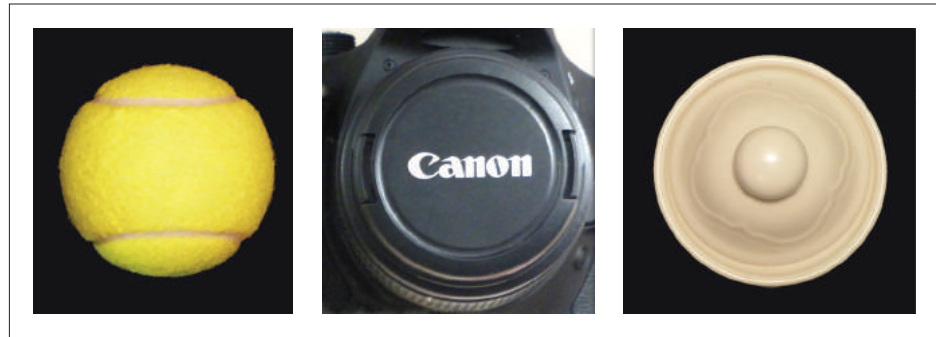
➤ **Figure 3.4** Even computer-vision programs that are able to recognize objects fairly accurately make mistakes, such as confusing objects that share features. In this example, the lens cover and the top of the teapot are erroneously classified as a "tennis ball." (Source: Based on K. Simonyan et al., 2012)

In another area of computer-vision research, programs have been created that can describe pictures of real scenes. For example, a computer accurately identified a scene similar to the one in **Figure 3.5** as "a large plane sitting on a runway." But mistakes still occur, as when a picture similar to the one in **Figure 3.6** was identified as "a young boy holding a baseball bat" (Fei-Fei, 2015). The computer's problem is that it doesn't have the huge storehouse of information about the world that humans begin accumulating as soon as they are born. If a computer has never seen a toothbrush, it identifies it as something with a similar shape. And, although the computer's response to the airplane picture is accurate, it is beyond the computer's capabilities to recognize that this is a picture of airplanes on display, perhaps at an air show, and that the people are not passengers but are visiting the air show. So on one hand, we have come a very long way from the first attempts in the 1950s to design computer-vision systems, but to date, humans still out-perceive computers. In the next section, we consider some of the reasons perception is so difficult for computers to master.
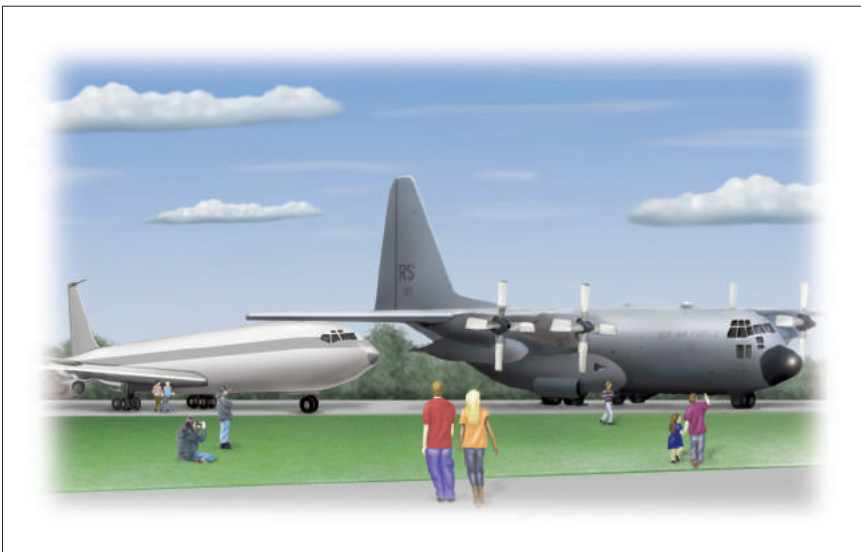
➤ **Figure 3.5** Picture similar to one that a computer vision program identified as "a large plane sitting on a runway."

➤ **Figure 3.6** Picture similar to one that a computer vision program identified as "a young boy holding a baseball bat."

Why Is It So Difficult to Design a Perceiving Machine?

We will now describe a few of the difficulties involved in designing a "perceiving machine." Remember that although the problems we describe pose difficulties for computers, humans solve them easily.

The Stimulus on the Receptors Is Ambiguous

When you look at the page of this book, the image cast by the borders of the page on your retina is ambiguous. It may seem strange to say that, because (1) the rectangular shape of the page is obvious, and (2) once we know the page's shape and its distance from the eye, determining its image on the retina is a simple geometry problem, which, as shown in **Figure 3.7**, can be solved by extending "rays" from the corners of the page (in red) into the eye.

But the perceptual system is not concerned with determining an object's image on the retina. It *starts* with the image on the retina, and its job is to determine what object "out there" created the image. The task of determining the object responsible for a particular image on the retina is called the **inverse projection problem**, because it involves starting with the retinal image and extending rays out from the eye. When we do this, as shown by extending the lines in Figure 3.7 out from the eye, we see that the retinal image created by the rectangular page could have also been created by a number of other objects, including a tilted trapezoid, a much larger rectangle, and an infinite number of other objects, located at different distances. When we consider that a particular image on the retina can be created by many different objects in the environment, it is easy to see why we say that the image on the retina is ambiguous. Nonetheless, humans typically solve the inverse projection problem easily, even though it still poses serious challenges to computer-vision systems.

Objects Can Be Hidden or Blurred

Sometimes objects are hidden or blurred. Look for the pencil and eyeglasses in **Figure 3.8** before reading further. Although it might take a little searching, people can find the pencil in the foreground and the glasses frame sticking out from behind the computer next to the picture, even though only a small portion of these objects is visible. People also easily perceive the book, scissors, and paper as whole objects, even though they are partially hidden by other objects.

Image on retina

Objects that create the same image on the retina
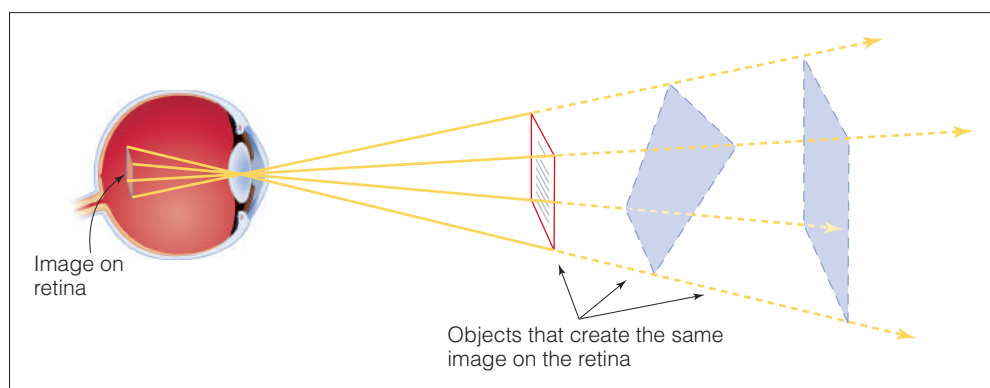
➤ **Figure 3.7** The projection of the book (red object) onto the retina can be determined by extending rays (solid lines) from the book into the eye. The principle behind the inverse projection problem is illustrated by extending rays out from the eye past the book (dashed lines). When we do this, we can see that the image created by the book can be created by an infinite number of objects, among them the tilted trapezoid and large rectangle shown here. This is why we say that the image on the retina is ambiguous.

➤ **Figure 3.8** A portion of the mess on the author's desk. Can you locate the hidden pencil (easy) and the author's glasses (hard)?

This problem of hidden objects occurs any time one object obscures part of another object. ==This occurs frequently in the environment, but people easily understand that the part of an object that is covered continues to exist, and they are able to use their knowledge of the environment to determine what is likely to be present.==

People are also able to recognize objects that are not in sharp focus, such as the faces in **Figure 3.9**. See how many of these people you can identify, and then consult the answers on page 91. Despite the degraded nature of these images, people can often identify most of them, whereas computers perform poorly on this task (Sinha, 2002).

## Objects Look Different from Different Viewpoints

Another problem facing any perceiving machine is that objects are often viewed from different angles, so their images are continually changing, as in **Figure 3.10**. ==People's ability to recognize an object even when it is seen from different viewpoints is called **viewpoint invariance**==. Computer-vision systems can achieve viewpoint invariance only by a laborious process that involves complex calculations designed to determine which points on an object match in different views (Vedaldi, Ling, & Soatto, 2010).
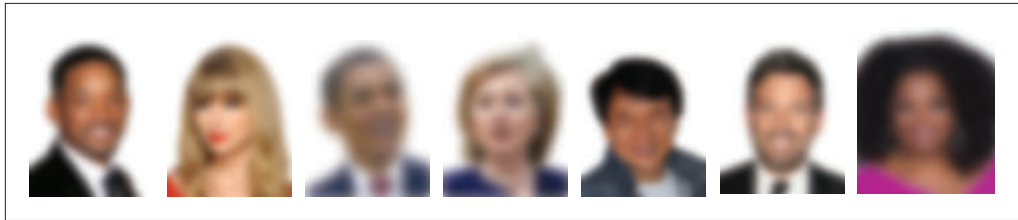
➤ **Figure 3.9** Who are these people? See page 91 for the answers.
(Source: Based on Sinha, 2002)

(a)   (b)   (c)
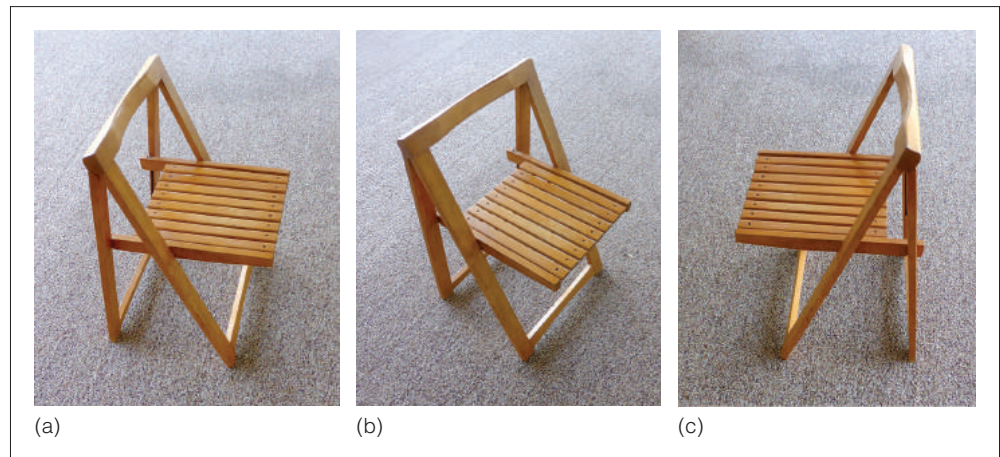
➤ **Figure 3.10** Your ability to recognize each of these views as being of the same chair is an example of viewpoint invariance.

Scenes Contain High-Level Information

Moving from objects to scenes adds another level of complexity. Not only are there often many objects in a scene, but they may be providing information about the scene that requires some reasoning to figure out. Consider, for example, the airplane picture in Figure 3.5. What is the basis for deciding the planes are probably on display at an air show? One answer is knowing that the plane on the right is an older-looking military plane that is most likely no longer in service. We also know that the people aren't passengers waiting to board, because they are walking on the grass and aren't carrying any luggage. Cues like this, although obvious to a person, would need to be programmed into a computer.

The difficulties facing any perceiving machine illustrate that the process of perception is more complex than it seems. Our task, therefore, in describing perception is to explain this process, focusing on how our human perceiving machine operates. We begin by considering two types of information used by the human perceptual system: (1) environmental energy stimulating the receptors and (2) knowledge and expectations that the observer brings to the situation.

▶ Information for Human Perception

Perception is built on a foundation of information from the environment. Looking at something creates an image on the retina. This image generates electrical signals that are transmitted through the retina, and then to the visual receiving area of the brain. This sequence of events from eye to brain is called **bottom-up processing**, because it starts at the "bottom" or beginning of the system, when environmental energy stimulates the receptors.

But perception involves information in addition to the foundation provided by activation of the receptors and bottom-up processing. Perception also involves factors such as a person's knowledge of the environment, and the expectations people bring to the perceptual situation. For example, remember the experiment described in Chapter 1, which showed that people identify a rapidly flashed object in a kitchen scene more accurately when that object fits the scene (Figure 1.13)? This knowledge we have of the environment is the basis of **top-down processing**—processing that originates in the brain, at the "top" of the perceptual system. It is this knowledge that enables people to rapidly identify objects and scenes, and also to go beyond mere identification of objects to determining the story behind a scene. We will now consider two additional examples of top-down processing: perceiving objects and hearing words in a sentence.

Perceiving Objects

An example of top-down processing, illustrated in **Figure 3.11**, is called "the multiple personalities of a blob," because even though all of the blobs are identical, they are perceived as different objects depending on their orientation and the context within which they are seen (Oliva & Torralba, 2007). The blob appears to be an object on a table in (b), a shoe on a person bending down in (c), and a car and a person crossing the street in (d). We perceive the blob as different objects because of our knowledge of the kinds of objects that are likely to be found in different types of scenes. The human advantage over computers is therefore due, in part, to the additional top-down knowledge available to humans.

blob

(a)   (b)

(c)   (d)
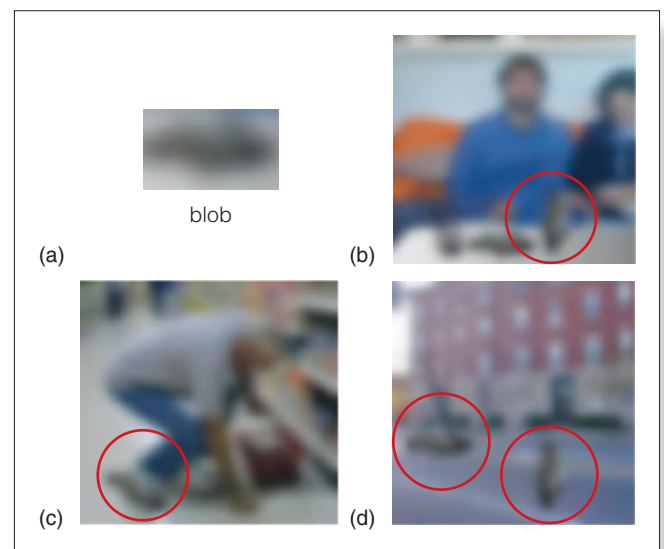
➤ **Figure 3.11** "Multiple personalities of a blob." What we expect to see in different contexts influences our interpretation of the identity of the "blob" inside the circles. (Source: Adapted from A. Oliva & A. Torralba, 2007)
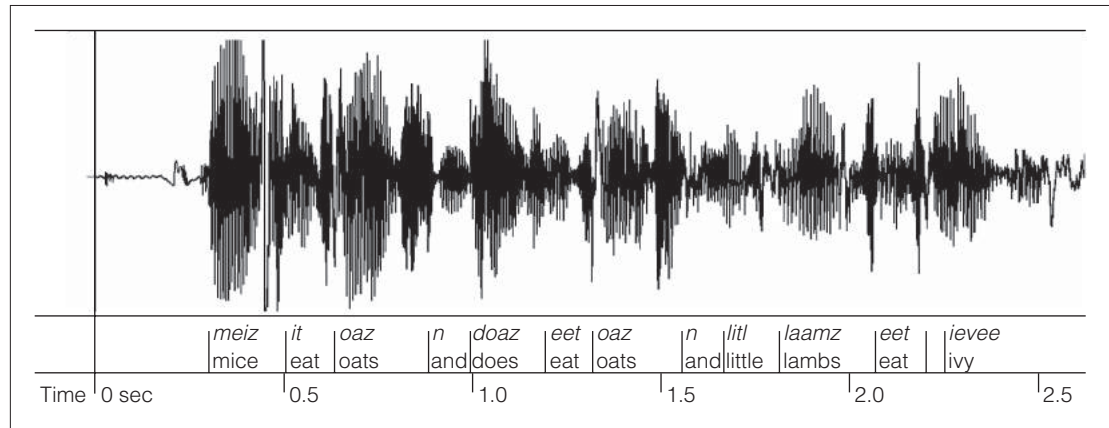
➤ **Figure 3.12** Sound energy for the sentence "Mice eat oats and does eat oats and little lambs eat ivy." The italicized words just below the sound record indicate how this sentence was pronounced by the speaker. The vertical lines next to the words indicate where each word begins. Note that it is difficult or impossible to tell from the sound record where one word ends and the next one begins.

(Source: Speech signal courtesy of Peter Howell)

Hearing Words in a Sentence

An example of how top-down processing influences speech perception occurs for me as I sit in a restaurant listening to people speaking Spanish at the next table. Unfortunately, I don't understand what they are saying because I don't understand Spanish. To me, the dialogue sounds like an unbroken string of sound, except for occasional pauses and when a familiar word like *gracias* pops out. My perception reflects the fact that the physical sound signal for speech is generally continuous, and when there are breaks in the sound, they do not necessarily occur between words. You can see this in **Figure 3.12** by comparing the place where each word in the sentence begins with the pattern of the sound signal.

The ability to tell when one word in a conversation ends and the next one begins is a phenomenon called **speech segmentation**. The fact that a listener familiar only with English and another listener familiar with Spanish can receive identical sound stimuli but experience different perceptions means that each listener's experience with language (or lack of it!) is influencing his or her perception. The continuous sound signal enters the ears and triggers signals that are sent toward the speech areas of the brain (bottom-up processing); if a listener understands the language, their knowledge of the language creates the perception of individual words (top-down processing).

While segmentation is aided by knowing the meanings of words, listeners also use other information to achieve segmentation. As we learn a language, we are learning more than the meaning of the words. Without even realizing it we are learning **transitional probabilities**—the likelihood that one sound will follow another within a word. For example, consider the words *pretty baby*. In English it is likely that *pre* and *ty* will be in the same word (***pre-tty***) but less likely that *ty* and *ba* will be in the same word (pre***tty ba***by).

Every language has transitional probabilities for different sounds, and the process of learning about transitional probabilities and about other characteristics of language is called **statistical learning**. Research has shown that infants as young as 8 months of age are capable of statistical learning.

Jennifer Saffran and coworkers (1996) carried out an early experiment that demonstrated statistical learning in young infants. **Figure 3.13a** shows the design of
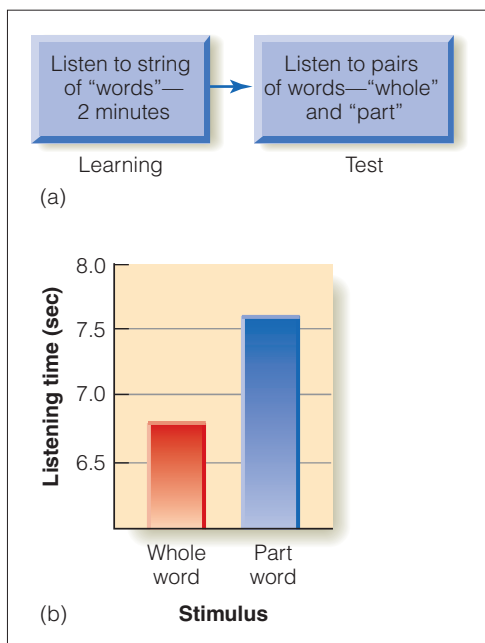
➤ **Figure 3.13** (a) Design of the experiment by Saffran and coworkers (1996), in which infants listened to a continuous string of nonsense syllables and were then tested to see which sounds they perceived as belonging together. (b) The results, indicating that infants listened longer to the "part-word" stimuli.

this experiment. During the learning phase of the experiment, the infants heard four non-sense "words" such as *bidaku, padoti, golabu*, and *tupiro*, which were combined in random order to create 2 minutes of continuous sound. An example of part of a string created by combining these words is *bidaku**padoti**golabu**tupiro**padoti**bidaku.…* In this string, every other word is printed in boldface in order to help you pick out the words. However, when the infants heard these strings, all the words were pronounced with the same into-nation, and there were no breaks between the words to indicate where one word ended and the next one began.

The transitional probabilities between two syllables that appeared *within* a word were always 1.0. For example, for the word *bidaku*, when /*bi*/ was presented, /*da*/ always fol-lowed it. Similarly, when /*da*/ was presented, /*ku*/ always followed it. In other words, these three sounds always occurred together and in the same order, to form the word *bidaku*.

The transitional probabilities between the *end* of one word and the *beginning* of an-other were only 0.33. For example, there was a 33 percent chance that the last sound, /*ku*/ from *bidaku*, would be followed by the first sound, /*pa*/, from *padoti*, a 33 percent chance that it would be followed by /tu/ from *tupiro*, and a 33 percent chance it would be followed by /*go*/ from *golabu*.

If Saffran's infants were sensitive to transitional probabilities, they would perceive stim-uli like *bidaku* or *padoti* as words, because the three syllables in these words are linked by transitional probabilities of 1.0. In contrast, stimuli like ***tibida*** (the end of *padoti* plus the beginning of ***bidaku***) would not be perceived as words, because the transitional probabili-ties were much smaller.

To determine whether the infants did, in fact, perceive stimuli like *bidaku* and *padoti* as words, the infants were tested by being presented with pairs of three-syllable stimuli. Some of the stimuli were "words" that had been presented before, such as *padoti*. These were the "whole-word" stimuli. The other stimuli were created from the end of one word and the beginning of another, such as *tibida*. These were the "part-word" stimuli.

The prediction was that the infants would choose to listen to the part-word stimuli longer than to the whole-word stimuli. This prediction was based on previous research that showed that infants tend to lose interest in stimuli that are repeated, and so become familiar, but pay more attention to novel stimuli that they haven't experienced before. Thus, if the infants perceived the whole-word stimuli as words that had been repeated over and over during the 2-minute learning session, they would pay less attention to these familiar stimuli than to the more novel part-word stimuli that they did not perceive as being words.

Saffran measured how long the infants listened to each sound by presenting a blinking light near the speaker where the sound was coming from. When the light attracted the in-fant's attention, the sound began, and it continued until the infant looked away. Thus, the infants controlled how long they heard each sound by how long they looked at the light. **Figure 3.13b** shows that the infants did, as predicted, listen longer to the part-word stimuli. From results such as these, we can conclude that the ability to use transitional probabilities to segment sounds into words begins at an early age.

The examples of how context affects our perception of the blob and how knowledge of the statistics of speech affects our ability to create words from a continuous speech stream illustrate that top-down processing based on knowledge we bring to a situation plays an important role in perception.

We have seen that perception depends on two types of information: bottom-up (in-formation stimulating the receptors) and top-down (information based on knowledge). Exactly how the perceptual system uses this information has been conceived of in different ways by different people. We will now describe four prominent approaches to perceiving objects, which will take us on a journey that begins in the 1800s and ends with modern conceptions of object perception.
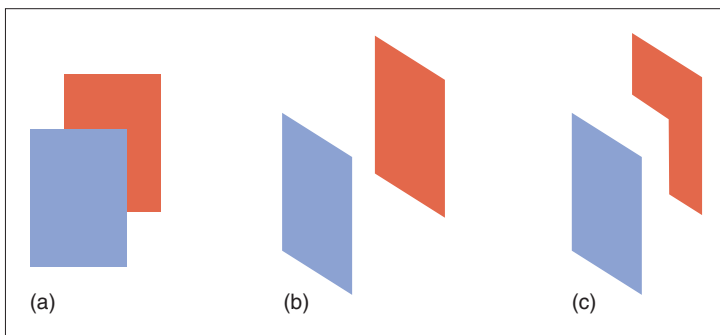
▶ Conceptions of Object Perception

An early idea about how people use information was proposed by 19th-century physicist and physiologist Hermann von Helmholtz (1866/1911).

Helmholtz's Theory of Unconscious Inference

Hermann von Helmholtz (1821–1894) was a physicist who made important contributions to fields as diverse as thermodynamics, nerve physiology, visual perception, and aesthetics. He also invented the ophthalmoscope, versions of which are still used today to enable physicians to examine the blood vessels inside the eye.

One of Helmholtz's contributions to perception was based on his realization that the image on the retina is ambiguous. We have seen that ambiguity means that a particular pattern of stimulation on the retina can be caused by a large number of objects in the environment (see Figure 3.7). For example, what does the pattern of stimulation in **Figure 3.14a** represent? For most people, this pattern on the retina results in the perception of a blue rectangle in front of a red rectangle, as shown in **Figure 3.14b**. But as **Figure 3.14c** indicates, this display could also have been caused by a six-sided red shape positioned behind or right next to the blue rectangle.

Helmholtz's question was, How does the perceptual system "decide" that this pattern on the retina was created by overlapping rectangles? His answer was the **likelihood principle**, which states that we perceive the object that is *most likely* to have caused the pattern of stimuli we have received. This judgment of what is most likely occurs, according to Helmholtz, by a process called **unconscious inference**, in which our perceptions are the result of unconscious assumptions, or inferences, that we make about the environment. Thus, we *infer* that it is likely that Figure 3.14a is a rectangle covering another rectangle because of experiences we have had with similar situations in the past.

Helmholtz's description of the process of perception resembles the process involved in solving a problem. For perception, the problem is to determine which object has caused a

▶ **Figure 3.14** The display in (a) is usually interpreted as being (b) a blue rectangle in front of a red rectangle. It could, however, be (c) a blue rectangle and an appropriately positioned six-sided red figure.

particular pattern of stimulation, and this problem is solved by a process in which the perceptual system applies the observer's knowledge of the environment in order to infer what the object might be.

An important feature of Helmholtz's proposal is that this process of perceiving what is most likely to have caused the pattern on the retina happens rapidly and unconsciously. These unconscious assumptions, which are based on the likelihood principle, ==result in perceptions that seem "instantaneous," even though they are the outcome of a rapid process.== Thus, although you might have been able to solve the perceptual puzzles in the scene in Figure 3.2 without much effort, this ability, according to Helmholtz, is the outcome of processes of which we are unaware. (See Rock, 1983, for a more recent version of this idea.)

The Gestalt Principles of Organization

We will now consider an approach to perception proposed by a group called the **Gestalt psychologists** about 30 years after Helmholtz proposed his theory of unconscious inference. The goal of the Gestalt approach was the same as Helmholtz's—to explain how we perceive objects—but they approached the problem in a different way.

The Gestalt approach to perception originated, in part, as a reaction to Wilhelm Wundt's structuralism (see page 7). Remember from Chapter 1 that ==Wundt proposed that our overall experience could be understood by combining basic elements of experience called *sensations*. According to this idea, our perception of the face in **Figure 3.15** is created by adding up many sensations, represented as dots in this figure.==

The Gestalt psychologists rejected the idea that perceptions were formed by "adding up" sensations. One of the origins of the Gestalt idea that perceptions could not be explained by adding up small sensations has been attributed to the experience of psychologist Max Wertheimer, who while on vacation in 1911 took a train ride through Germany (Boring, 1942). When he got off the train to stretch his legs at Frankfurt, he bought a stroboscope from a toy vendor on the train platform. The stroboscope, a mechanical device that created an illusion of movement by rapidly alternating two slightly different pictures, caused Wertheimer to wonder how the structuralist idea that experience is created from sensations could explain the illusion of movement he observed.

Figure 3.16 diagrams the principle behind the illusion of movement created by the stroboscope, which is called ==**apparent movement**== because, although movement is perceived, nothing is actually moving. There are three components to stimuli that create apparent movement: (1) One light flashes on and off (**Figure 3.16a**); (2) there is a period of darkness, lasting a fraction of a second (**Figure 3.16b**); and (3) the second light flashes on and off (**Figure 3.16c**). Physically, therefore, there are two lights flashing on and off separated by a period of darkness. But we don't see the darkness because our perceptual system adds something during the period of darkness—the perception of a light moving through the space between the flashing lights (**Figure 3.16d**). Modern examples of apparent movement are electronic signs that display moving advertisements or news headlines, and movies. The perception of movement in these displays is so compelling that it is difficult to imagine that they are made up of stationary lights flashing on and off (for the news headlines) or still images flashed one after the other (for the movies).

Wertheimer drew two conclusions from the phenomenon of apparent movement. ==His first conclusion was that apparent movement cannot be explained by sensations, because there is nothing in the dark space between the flashing light==s. His second conclusion became one of the basic principles of Gestalt psychology: ==*The whole is different than the sum of its parts*.== This conclusion follows from the fact that the perceptual system creates the perception of movement from stationary images. This idea that the whole is different than the sum of its parts led the Gestalt psychologists to propose a number of **principles of perceptual organization** to explain the way elements are grouped together to create larger

➤ **Figure 3.15** According to structuralism, a number of sensations (represented by the dots) add up to create our perception of the face.

(a) One light flashes

(b) Darkness

(c) The second light flashes

(d) Flash—dark—flash
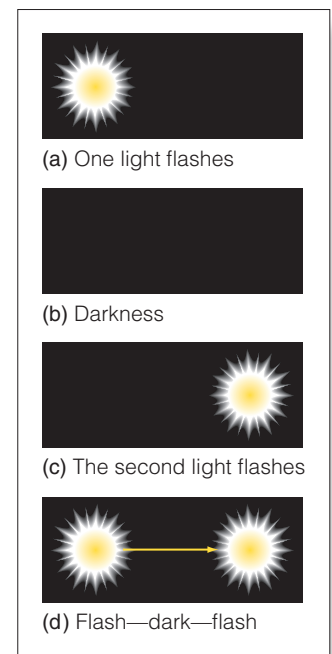
➤ **Figure 3.16** The conditions for creating apparent movement. (a) One light flashes, followed by (b) a short period of darkness, followed by (c) another light flashing at a different position. The resulting perception, symbolized in (d), is a light moving from left to right. Movement is seen between the two lights even though there is only darkness in the space between them.
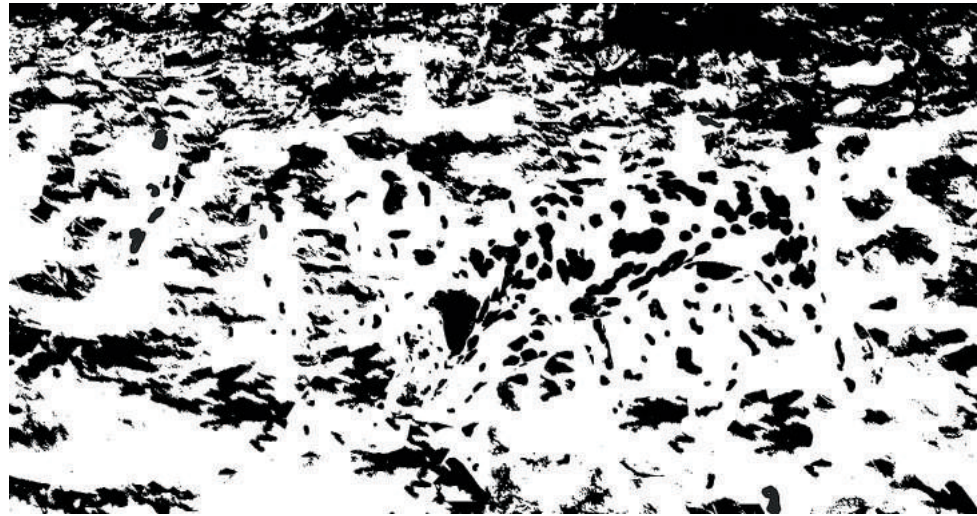
➤ **Figure 3.17** Some black and white shapes that become perceptually organized into a Dalmatian. (See page 91 for an outline of the Dalmatian.)

objects. For example, in **Figure 3.17**, some of the black areas become grouped to form a Dalmatian and others are seen as shadows in the background. We will describe a few of the Gestalt principles, beginning with one that brings us back to Crystal's run along the beach.

Good Continuation   The **principle of good continuation** states the following: *Points that, when connected, result in straight or smoothly curving lines are seen as belonging together, and the lines tend to be seen in such a way as to follow the smoothest path. Also, objects that are overlapped by other objects are perceived as continuing behind the overlapping object*. Thus, when Crystal saw the coiled rope in Figure 3.1c, she wasn't surprised that when she grabbed one end of the rope and flipped it, it turned out to be one continuous strand (**Figure 3.18**). The reason this didn't surprise her is that even though there were many places where one part of the rope overlapped another part, she didn't perceive the rope as consisting of a number of separate pieces; rather, she perceived the rope as continuous. (Also consider your shoelaces!)

Pragnanz   Pragnanz, roughly translated from the German, means "good figure." The **law of pragnanz**, also called the **principle of good figure** or the **principle of simplicity**, states: *Every stimulus pattern is seen in such a way that the resulting structure is as simple as possible.*


(a)   (b)
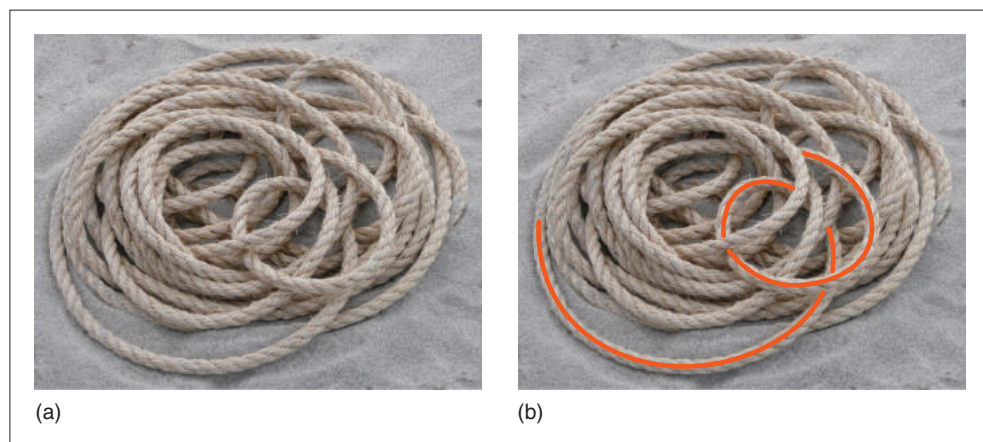
➤ **Figure 3.18** (a) Rope on the beach. (b) Good continuation helps us perceive the rope as a single strand.

The familiar Olympic symbol in **Figure 3.19a** is an example of the law of simplicity at work. We see this display as five circles and not as a larger number of more complicated shapes such as the ones shown in the "exploded" view of the Olympic symbol in **Figure 3.19b**. (The law of good continuation also contributes to perceiving the five circles. Can you see why this is so?)

Similarity   Most people perceive **Figure 3.20a** as either horizontal rows of circles, vertical columns of circles, or both. But when we change the color of some of the columns, as in **Figure 3.20b**, most people perceive vertical columns of circles. This perception illustrates the **principle of similarity**: *Similar things appear to be grouped together.* A striking example of grouping by similarity of color is shown in **Figure 3.21**. Grouping can also occur because of similarity of size, shape, or orientation.

There are many other principles of organization, proposed by the original Gestalt psychologists (Helson, 1933) as well as by modern psychologists (Palmer, 1992; Palmer & Rock, 1994), but the main message, for our discussion, is that the Gestalt psychologists realized that perception is based on more than just the pattern of light and dark on the retina. In their conception, perception is determined by specific organizing principles.

But where do these organizing principles come from? Max Wertheimer (1912) describes these principles as "intrinsic laws," which implies that they are built into the system. This idea that the principles are "built in" is consistent with the Gestalt psychologists' idea that although a person's experience can *influence* perception, the role of experience is minor compared to the perceptual principles (also see Koffka, 1935). This idea that experience plays only a minor role in perception differs from Helmholtz's likelihood principle, which proposes that our knowledge of the environment enables us to determine what is most likely to have created the pattern on the retina and also differs from modern approaches to object perception, which propose that our experience with the environment is a central component of the process of perception.
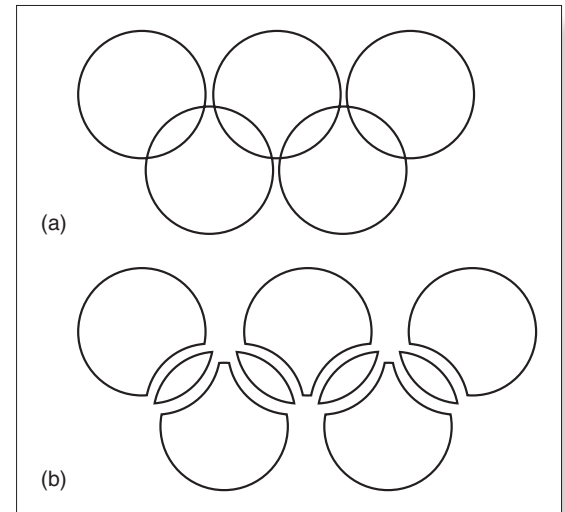
➤ **Figure 3.19** The Olympic symbol is perceived as five circles (a), not as the nine shapes in (b).
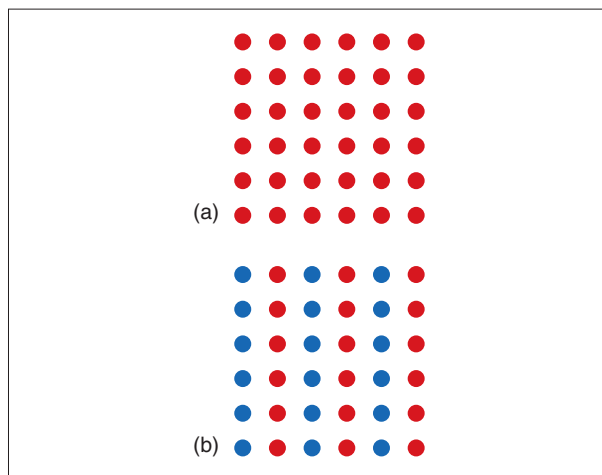
➤ **Figure 3.20** (a) This pattern of dots is perceived as horizontal rows, vertical columns, or both. (b) This pattern of dots is perceived as vertical columns.
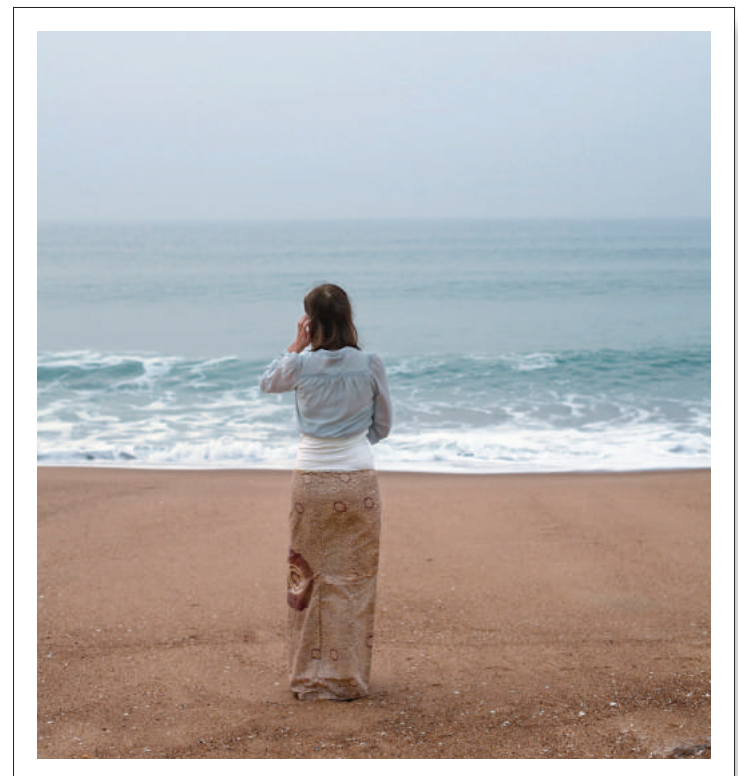
➤ **Figure 3.21** This photograph, Waves, by Wilma Hurskainen, was taken at the exact moment that the front of the white water aligned with the white area on the woman's clothing. Similarity of color causes grouping; differently colored areas of the dress are perceptually grouped with the same colors in the scene. Also notice how the front edge of the water creates grouping by good continuation across the woman's dress.

(Source: Courtesy of Wilma Hurskainen)

Taking Regularities of the Environment into Account

Modern perceptual psychologists take experience into account by noting that certain characteristics of the environment occur frequently. For example, blue is associated with open sky, landscapes are often green and smooth, and verticals and horizontals are often associated with buildings. These frequently occurring characteristics are called **regularities in the environment**. There are two types of regularities: *physical regularities* and *semantic regularities*.

Physical Regularities **Physical regularities** are regularly occurring physical properties of the environment. For example, there are more vertical and horizontal orientations in the environment than oblique (angled) orientations. This occurs in human-made environments (for example, buildings contain lots of horizontals and verticals) and also in natural environments (trees and plants are more likely to be vertical or horizontal than slanted) (Coppola et al., 1998) (**Figure 3.22**). It is therefore no coincidence that people can perceive horizontals and verticals more easily than other orientations, an effect called the **oblique effect** (Appelle, 1972; Campbell et al., 1966; Orban et al., 1984). Another example of a physical regularity is that when one object partially covers another one, the contour of the partially covered object "comes out the other side," as occurs for the rope in Figure 3.18.

Another physical regularity is illustrated by **Figure 3.23a**, which shows indentations created by people walking in the sand. But turning this picture upside down, as in **Figure 3.23b**, transforms the indentations into rounded mounds. Our perception in these two situations has been explained by the **light-from-above assumption**: We usually assume that light is coming from above, because light in our environment, including the sun and most artificial light, usually comes from above (Kleffner & Ramachandran, 1992). **Figure 3.23c** shows how light coming from above and from the left illuminates an indentation, leaving a shadow on the left. **Figure 3.23d** shows how the same light illuminates a bump, leaving a shadow on the right. Our perception of illuminated shapes is influenced by how they are shaded, combined with the brain's assumption that light is coming from above.

One of the reasons humans are able to perceive and recognize objects and scenes so much better than computer-guided robots is that our system is adapted to respond to the physical characteristics of our environment, such as the orientations of objects and the direction of light. But this adaptation goes beyond physical characteristics. It also occurs because, as we saw when we considered the multiple personalities of a blob (page 67), we have learned about what types of objects typically occur in specific types of scenes.

Semantic Regularities In language, *semantics* refers to the meanings of words or sentences. Applied to perceiving scenes, *semantics* refers to the meaning of a scene. This meaning is often related to what happens within a scene. For example, food preparation, cooking, and perhaps eating occur in a kitchen; waiting around, buying tickets, checking luggage, and going through security checkpoints happen in airports. **Semantic regularities** are the characteristics associated with the functions carried out in different types of scenes.

Bruce Goldstein

➤ **Figure 3.22** In these two scenes from nature, horizontal and vertical orientations are more common than oblique orientations. These scenes are special examples, picked because of the large proportion of verticals. However, randomly selected photos of natural scenes also contain more horizontal and vertical orientations than oblique orientations. This also occurs for human-made buildings and objects.
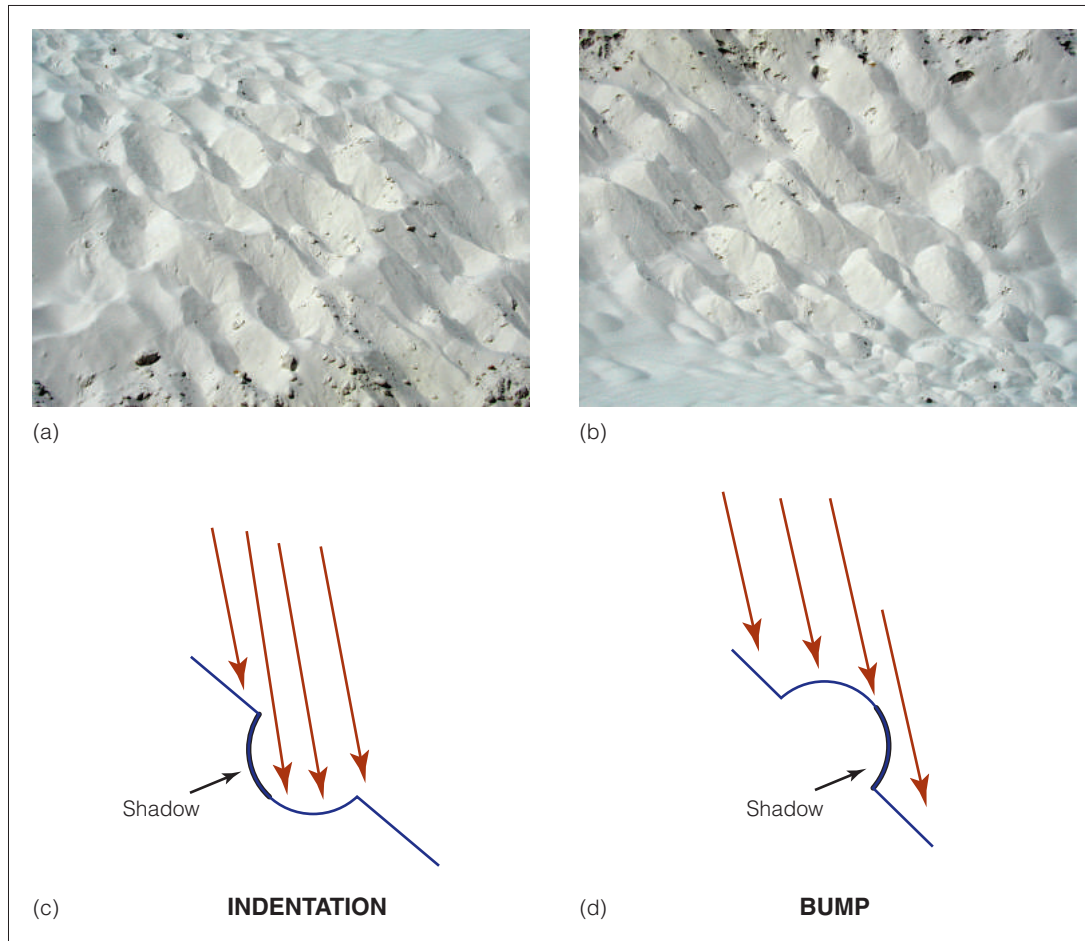
➤ **Figure 3.23** (a) Indentations made by people walking in the sand. (b) ==Turning the picture upside down turns indentations into rounded mounds.== (c) How light from above and to the left illuminates an indentation, causing a shadow on the left. (d) The same light illuminating a bump causes a shadow on the right.

One way to demonstrate that people are aware of semantic regularities is simply to ask them to imagine a particular type of scene or object, as in the following demonstration.

DEMONSTRATION   Visualizing Scenes and Objects

Your task in this demonstration is simple. Close your eyes and then visualize or simply think about the following scenes and objects:

1. An office
2. The clothing section of a department store
3. A microscope
4. A lion

Most people who have grown up in modern society have little trouble visualizing an office or the clothing section of a department store. What is important about this ability, for our purposes, is that part of this visualization involves details within these scenes. Most people see an office as having a desk with a computer on it, bookshelves, and a chair. The department store scene contains racks of clothes, a changing room, and perhaps a cash register. What did you see when you visualized the microscope or the lion? Many people report seeing not just a

single object, but an object within a setting. Perhaps you perceived the microscope sitting on a lab bench or in a laboratory and the lion in a forest, on a savannah, or in a zoo. The point of this demonstration is that our visualizations contain information based on our knowledge of different kinds of scenes. This knowledge of what a given scene typically contains is called a **scene schema**, and the expectations created by scene schemas contribute to our ability to perceive objects and scenes. For example, Palmer's (1975) experiment (Figure 1.13), in which people identified the bread, which fit the kitchen scene, faster than the mailbox, which didn't fit the scene, is an example of the operation of people's scene schemas for "kitchen." In connection with this, how do you think your scene schemas for "airport" might contribute to your interpretation of what is happening in the scene in Figure 3.5?

Although people make use of regularities in the environment to help them perceive, they are often unaware of the specific information they are using. This aspect of perception is similar to what occurs when we use language. Even though we aren't aware of transitional probabilities in language, we use them to help perceive words in a sentence. Even though we may not think about regularities in visual scenes, we use them to help perceive scenes and the objects within scenes.
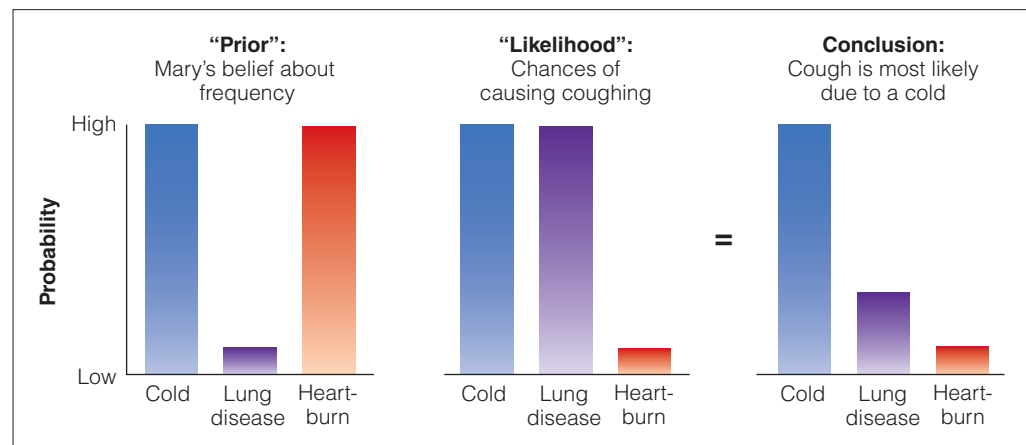
Bayesian Inference

Two of the ideas we have described—(1) Helmholtz's idea that we resolve the ambiguity of the retinal image by inferring what is most likely, given the situation, and (2) the idea that regularities in the environment provide information we can use to resolve ambiguities—are the starting point for our last approach to object perception: *Bayesian inference* (Geisler, 2008, 2011; Kersten et al., 2004; Yuille & Kersten, 2006).

Bayesian inference was named after Thomas Bayes (1701–1761), who proposed that our estimate of the probability of an outcome is determined by two factors: (1) the **prior probability**, or simply the **prior**, which is our initial belief about the probability of an outcome, and (2) the extent to which the available evidence is consistent with the outcome. This second factor is called the **likelihood** of the outcome.

To illustrate Bayesian inference, let's first consider **Figure 3.24a**, which shows Mary's *priors* for three types of health problems. Mary believes that having a cold or heartburn is likely to occur, but having lung disease is unlikely. With these priors in her head (along with lots of other beliefs about health-related matters), Mary notices that her friend Charles has a bad cough. She guesses that three possible causes could be a cold, heartburn, or lung disease. Looking further into possible causes, she does some research and finds that coughing is often associated with having either a cold or lung disease, but isn't associated with heartburn (**Figure 3.24b**). This additional information, which is the *likelihood*, is combined with Mary's *prior* to produce the conclusion that Charles probably has a cold (**Figure 3.24c**)

➤ **Figure 3.24** These graphs present hypothetical probabilities to illustrate the principle behind Bayesian inference. (a) Mary's beliefs about the relative frequency of having a cold, lung disease, and heartburn. These beliefs are her *priors*. (b) Further data indicate that colds and lung disease are associated with coughing, but heartburn is not. These data contribute to the *likelihood*. (c) Taking the priors and likelihood together results in the conclusion that Charles's cough is probably due to a cold.

(Tenenbaum et al., 2011). In practice, Bayesian inference involves a mathematical procedure in which the prior is multiplied by the likelihood to determine the probability of the outcome. Thus, people start with a prior and then use additional evidence to update the prior and reach a conclusion (Körding & Wolpert, 2006).

Applying this idea to object perception, let's return to the inverse projection problem from Figure 3.7. Remember that the inverse projection problem occurs because a huge number of possible objects could be associated with a particular image on the retina. So, the problem is how to determine what is "out there" that is causing a particular retinal image. Luckily, we don't have to rely only on the retinal image, because we come to most perceptual situations with prior probabilities based on our past experiences.

One of the *priors* you have in your head is that books are rectangular. Thus, when you look at a book on your desk, your initial belief is that it is likely that the book is rectangular. The *likelihood* that the book is rectangular is provided by additional evidence such as the book's retinal image, combined with your perception of the book's distance and the angle at which you are viewing the book. If this additional evidence is consistent with your prior that the book is rectangular, the likelihood is high and the perception "rectangular" is strengthened. Additional testing by changing your viewing angle and distance can further strengthen the conclusion that the shape is a rectangle. Note that you aren't necessarily conscious of this testing process—it occurs automatically and rapidly. The important point about this process is that while the retinal image is still the starting point for perceiving the shape of the book, adding the person's prior beliefs reduces the possible shapes that could be causing that image.

What Bayesian inference does is to restate Helmholtz's idea—that we perceive what is most likely to have created the stimulation we have received—in terms of probabilities. It isn't always easy to specify these probabilities, particularly when considering complex perceptions. However, because Bayesian inference provides a specific procedure for determining what might be out there, researchers have used it to develop computer-vision systems that can apply knowledge about the environment to more accurately translate the pattern of stimulation on their sensors into conclusions about the environment. (Also see Goldreich & Tong, 2013, for an example of how Bayesian inference has been applied to tactile perception.)

Comparing the Four Approaches

Now that we have described four conceptions of object perception (Helmholtz's unconscious inference, the Gestalt laws of organization, regularities in the environment, and Bayesian inference), here's a question: Which one is different from the other three? After you've figured out your answer, look at the bottom of the page.

The approaches of Helmholtz, regularities, and Bayes all have in common the idea that we use data about the environment, gathered through our past experiences in perceiving, to determine what is out there. Top-down processing is therefore an important part of these approaches.

The Gestalt psychologists, in contrast, emphasized the idea that the principles of organization are built in. They acknowledged that perception is affected by experience but argued that built-in principles can override experience, thereby assigning bottom-up processing a central role in perception. The Gestalt psychologist Max Wertheimer (1912) provided the following example to illustrate how built-in principles could override experience: Most people recognize **Figure 3.25a** as *W* and *M* based on their past experience with these letters. However, when the letters are arranged as in **Figure 3.25b**, most people see two uprights plus a

(a)   (b)
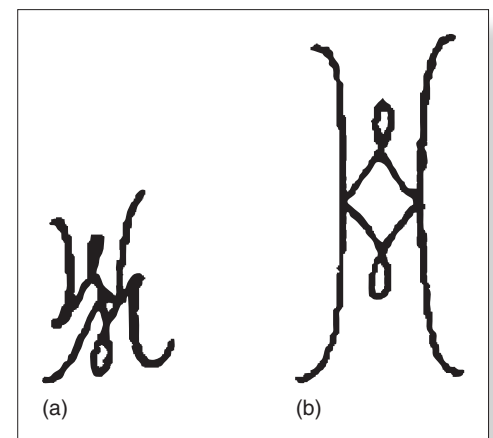
➤ **Figure 3.25** (a) W on top of M. (b) When combined, a new pattern emerges, overriding the meaningful letters.

(Source: From M. Wertheimer, 1912)

Answer: The Gestalt approach.

Bruce Goldstein

➤ **Figure 3.26** A usual occurrence in the environment: Objects (the men's legs) are partially hidden by another object (the grey boards). In this example, the men's legs continue in a straight line and are the same color above and below the boards, so it is highly likely that they continue behind the boards.

pattern between them. The uprights, which are created by the principle of good continuation, are the dominant perception and override the effects of past experience we have had with *W*s and *M*s.

Although the Gestalt psychologists deemphasized experience, using arguments like the preceding one, modern psychologists have pointed out that the laws of organization could, in fact, have been created by experience. For example, it is possible that the principle of good continuation has been determined by experience with the environment. Consider the scene in **Figure 3.26**. From years of experience in seeing objects that are partially covered by other objects, we know that when two visible parts of an object (like the men's legs) have the same color (principle of similarity) and are "lined up" (principle of good continuation), they belong to the same object and extend behind whatever is blocking it. Thus, one way to look at the Gestalt principles is that they describe the operating characteristics of the human perceptual system, *which happen to be determined at least partially by experience*. In the next section, we will consider physiological evidence that experiencing certain stimuli over and over can actually shape the way neurons respond.

TEST YOURSELF 3.2

1. Describe Helmholtz's theory of unconscious inference. What is the likelihood principle?

2. Describe the Gestalt approach to perception, focusing on the principles of organization. How do these principles originate, according to the Gestalt psychologists?

3. What are regularities of the environment, and how do they influence perception? Distinguish between physical regularities and semantic regularities. What is a scene schema?

4. Describe Bayesian inference in terms of how it would explain the "coughing" example and the inverse projection problem.

5. How does the Gestalt approach differ from the other three? How do modern psychologists explain the relation between experience and the principles of organization?

▶ Neurons and Knowledge About the Environment

We will now follow up on the idea that experience can shape the way neurons respond. Our starting point is the finding that there are more neurons in the animal and human visual cortex that respond to horizontal and vertical orientations than to oblique (slanted) orientations.

Neurons That Respond to Horizontals and Verticals

When we described physical regularities in the environment, we mentioned that horizontals and verticals are common features of the environment (Figure 3.22), and behavioral experiments have shown that people are more sensitive to these orientations than to other

orientations that are not as common (the *oblique effect*; see page 74). It is not a coincidence, therefore, that when researchers have recorded the activity of single neurons in the visual cortex of monkeys and ferrets, they have found more neurons that respond best to horizontals and verticals than neurons that respond best to oblique orientations (Coppola et al., 1998; DeValois et al., 1982). Evidence from brain-scanning experiments suggests that this occurs in humans as well (Furmanski & Engel, 2000).

Why are there more neurons that respond to horizontals and verticals? One possible answer is based on the **theory of natural selection**, which states that characteristics that enhance an animal's ability to survive, and therefore reproduce, will be passed on to future generations. Through the process of evolution, organisms whose visual systems contained neurons that fired to important things in the environment (such as verticals and horizontals, which occur frequently in the forest, for example) would be more likely to survive and pass on an enhanced ability to sense verticals and horizontals than would an organism with a visual system that did not contain these specialized neurons. Through this evolutionary process, the visual system may have been shaped to contain neurons that respond to things that are found frequently in the environment.

Although there is no question that perceptual functioning has been shaped by evolution, there is also a great deal of evidence that *learning* can shape the response properties of neurons through the process of *experience-dependent plasticity* that we introduced in Chapter 2 (page 34).

Experience-Dependent Plasticity

In Chapter 2, we described Blakemore and Cooper's (1970) experiment in which they showed that rearing cats in horizontal or vertical environments can cause neurons in the cat's cortex to fire preferentially to horizontal or vertical stimuli. This shaping of neural responding by experience, which is called *experience-dependent plasticity*, provides evidence that experience can shape the nervous system.

Experience-dependent plasticity has also been demonstrated in humans using the brain imaging technique of fMRI (see Method: Brain Imaging, page 41). The starting point for this research is the finding that there is an area in the temporal lobe called the fusiform face area (FFA) that contains many neurons that respond best to faces (see Chapter 2, page 42). Isabel Gauthier and coworkers (1999) showed that experience-dependent plasticity may play a role in determining these neurons' response to faces by measuring the level of activity in the FFA in response to faces and also to objects called Greebles (**Figure 3.27a**). Greebles are families of computer-generated "beings" that all have the same basic configuration but differ in the shapes of their parts (just like faces). The left pair of bars in **Figure 3.27b** show that for "Greeble novices" (people who have had little experience in perceiving Greebles), the faces cause more FFA activity than the Greebles.

Gauthier then gave her subjects extensive training over a 4-day period in "Greeble recognition." These training sessions, which required that each Greeble be labeled with a specific name, turned the participants into "Greeble experts." The right bars in Figure 3.27b show that after the training, the FFA responded almost as well to Greebles as to faces. Apparently, the FFA contains neurons that respond not just to faces but to other complex objects as well. The particular objects to which the neurons respond best are established by experience with the objects. In fact, Gauthier has also shown that neurons in the FFA
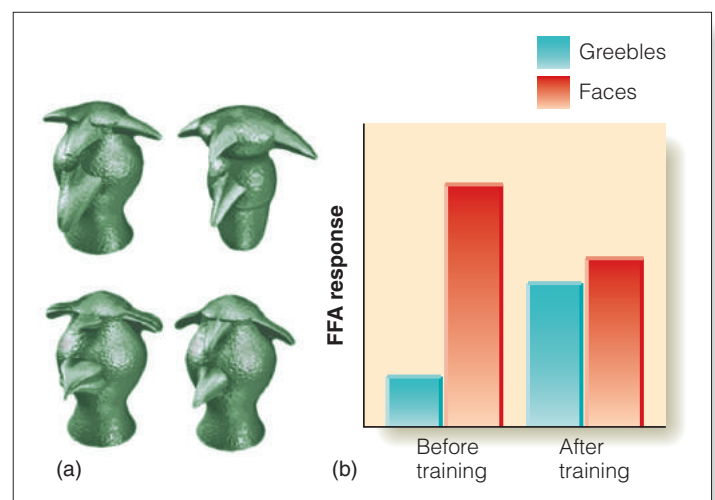
➤ **Figure 3.27** (a) Greeble stimuli used by Gauthier. Participants were trained to name each different Greeble. (b) Magnitude of FFA responses to faces and Greebles before and after Greeble training.

(Source: Based on I. Gauthier et al., 1999)

of people who are experts in recognizing cars and birds respond well not only to human faces but to cars (for the car experts) and to birds (for the bird experts) (Gauthier et al., 2000). Just as rearing kittens in a vertical environment increased the number of neurons that responded to verticals, training humans to recognize Greebles, cars, or birds causes the FFA to respond more strongly to these objects. These results support the idea that neurons in the FFA respond strongly to faces because we have a lifetime of experience perceiving faces.

These demonstrations of experience-dependent plasticity in kittens and humans show that the brain's functioning can be "tuned" to operate best within a specific environment. Thus, continued exposure to things that occur regularly in the environment can cause neurons to become adapted to respond best to these regularities. Looked at in this way, it is not unreasonable to say that neurons can reflect knowledge about properties of the environment.

We have come a long way from thinking about perception as something that happens automatically in response to activation of sensory receptors. We've seen that perception is the outcome of an interaction between bottom-up information, which flows from receptors to brain, and top-down information, which usually involves knowledge about the environment or expectations related to the situation.

At this point in our description of perception, how would you answer the question: "What is the purpose of perception?" One possible answer is that the purpose of perception is to create our awareness of what is happening in the environment, as when we see objects in scenes or we perceive words in a conversation. But it becomes obvious that this answer doesn't go far enough, when we ask, *why* it is important that we are able to experience objects in scenes and words in conversations?

The answer to that question is that an important purpose of perception is to enable us to interact with the environment. The key word here is *interact*, because interaction implies taking action. We are taking action when we pick something up, when we walk across campus, when we have an interaction with someone we are talking with. Interactions such as these are essential for accomplishing what we want to accomplish, and often are essential for our very survival. We end this chapter by considering the connection between perception and action, first by considering behavior and then physiology.

▶ Perception and Action: Behavior

The approach to perception we have described so far could be called the "sitting in a chair" approach to studying perception, because most of the situations we have described could occur as a person sits in a chair viewing various stimuli. In fact, that is probably what you are doing as you read this book—reading words, looking at pictures, doing "demonstrations," all while sitting still. We will now consider how movement helps us perceive, and how action and perception interact.

Movement Facilitates Perception

Although movement adds a complexity to perception that isn't there when we are sitting in one place, movement also helps us perceive objects in the environment more accurately. One reason this occurs is that moving reveals aspects of objects that are not apparent from a single viewpoint. For example, consider the "horse" in **Figure 3.28**. From one viewpoint, this object looks like a metal sculpture of a fairly normal horse (**Figure 3.28a**). However, walking around the horse reveals that it isn't as normal as it first appeared (**Figures 3.28b** and **3.28c**). Thus, seeing an object from different viewpoints provides added information that results in more accurate perception, especially for objects that are out of the ordinary, such as the distorted horse.

➤ **Figure 3.28** Three views of a "horse." Moving around an object can reveal its true shape.

The Interaction of Perception and Action

Our concern with movement extends beyond noting that it helps us perceive objects by revealing additional information about them. Movement is also important because of the coordination that is continually occurring between perceiving stimuli and taking action toward these stimuli. Consider, for example, what happens when Crystal, resting in the coffee shop after her run, reaches out to pick up her cup of coffee (**Figure 3.29**). She first identifies the coffee cup among the flowers and other objects on the table (**Figure 3.29a**). Once the coffee cup is perceived, she reaches for it, taking into account its location on the table (**Figure 3.29b**). As she reaches, avoiding the flowers, she positions her fingers to grasp the cup, taking into account her perception of the cup's handle (**Figure 3.29c**); then she lifts the cup with just the right amount of force, taking into account her estimate of how heavy it is based on her perception of its fullness. This simple action requires continually perceiving the position of the cup, and of her hand and fingers relative to the cup, while calibrating her actions in order to accurately grasp the cup and then pick it up without spilling any coffee (Goodale, 2010). All this just to pick up a cup of coffee! What's amazing about this sequence is that it happens almost automatically, without much effort at all. But as with everything else about perception, this ease and apparent simplicity are achieved with the aid of complex underlying mechanisms. We will now describe the physiology behind these mechanisms.

(a) Perceive cup   (b) Reach for cup   (c) Grasp cup

➤ **Figure 3.29** Picking up a cup of coffee: (a) perceiving and recognizing the cup; (b) reaching for it; (c) grasping and picking it up. This action involves coordination between perceiving and action that is carried out by two separate streams in the brain, as described in the text.

▶ Perception and Action: Physiology

Psychologists have long recognized the close connection between perceiving objects and interacting with them, but the details of this link between perception and action have become clearer as a result of physiological research that began in the 1980s. This research has shown that there are two processing streams in the brain—one involved with perceiving objects, and the other involved with locating and taking action toward these objects. This physiological research involves two methods: *brain ablation*—the study of the effect of removing parts of the brain in animals, and *neuropsychology*—the study of the behavior of people with brain damage, which we described in Chapter 2 (see page 38). Both of these methods demonstrate how studying the functioning of animals and humans with brain damage can reveal important principles about the functioning of the normal (intact) brain.

What and *Where* Streams

In a classic experiment, Leslie Ungerleider and Mortimer Mishkin (1982) studied how removing part of a monkey's brain affected its ability to identify an object and to determine the object's location. This experiment used a technique called **brain ablation**—removing part of the brain.

METHOD   Brain Ablation

The goal of a brain ablation experiment is to determine the function of a particular area of the brain. This is accomplished by first determining an animal's capacity by testing it behaviorally. Most ablation experiments studying perception have used monkeys because of the similarity of the monkey's visual system to that of humans and because monkeys can be trained to demonstrate perceptual capacities such as acuity, color vision, depth perception, and object perception.

   Once the animal's perception has been measured, a particular area of the brain is ablated (removed or destroyed), either by surgery or by injecting a chemical in the area to be removed. Ideally, one particular area is removed and the rest of the brain remains intact. After ablation, the monkey is tested to determine which perceptual capacities remain and which have been affected by the ablation. Ablation is also called *lesioning*.

Ungerleider and Mishkin presented monkeys with two tasks: (1) an object discrimination problem and (2) a landmark discrimination problem. In the **object discrimination problem**, a monkey was shown one object, such as a rectangular solid, and was then presented with a two-choice task like the one shown in **Figure 3.30a**, which included the "target" object (the rectangular solid) and another stimulus, such as the triangular solid. If the monkey pushed aside the target object, it received the food reward that was hidden in a well under the object. The **landmark discrimination problem** is shown in **Figure 3.30b**. Here, the tall cylinder is the landmark, which indicates the food well that contains food. The monkey received food if it removed the food well cover closer to the tall cylinder.

   In the ablation part of the experiment, part of the temporal lobe was removed in some monkeys. Behavioral testing showed that the object discrimination problem became very difficult for the monkeys when their temporal lobes were removed. This result indicates that the neural pathway that reaches the temporal lobes is responsible for determining an object's identity. Ungerleider and Mishkin therefore called the pathway leading from the striate cortex to the temporal lobe the **what pathway** (**Figure 3.31**).
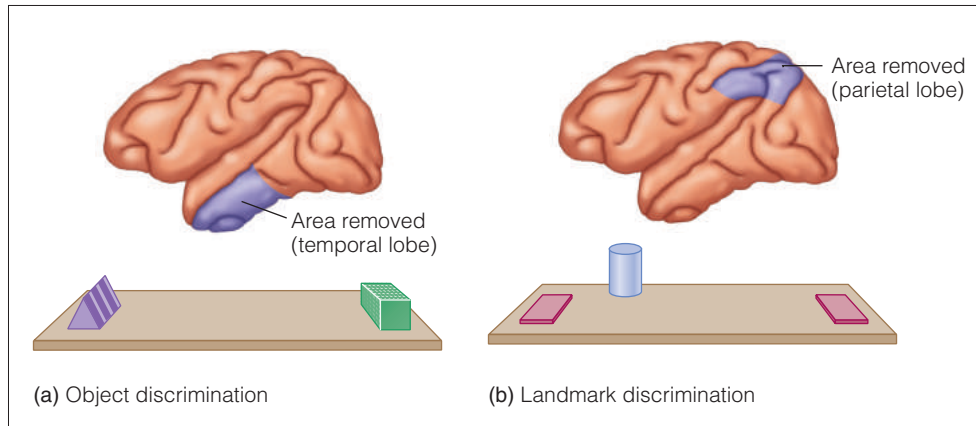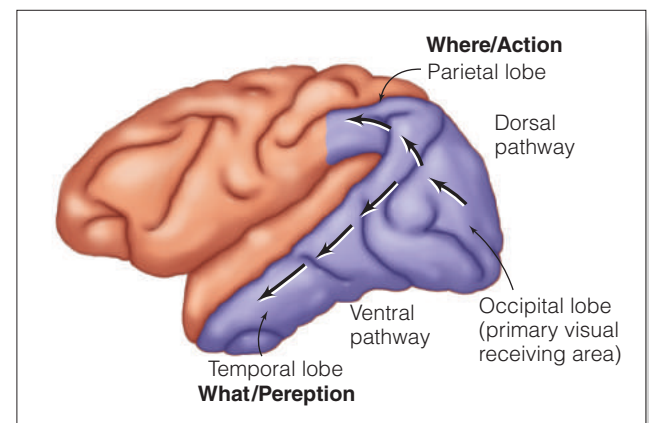
Other monkeys, which had their parietal lobes removed, had difficulty solving the landmark discrimination problem. This result indicates that the pathway that leads to the parietal lobe is responsible for determining an object's location. Ungerleider and Mishkin therefore called the pathway leading from the striate cortex to the parietal lobe the **where pathway** (Figure 3.31).

The *what* and *where* pathways are also called the **ventral pathway** (what) and the **dorsal pathway** (where), because the lower part of the brain, where the temporal lobe is located, is the ventral part of the brain, and the upper part of the brain, where the parietal lobe is located, is the dorsal part of the brain. The term *dorsal* refers to the back or the upper surface of an organism; thus, the dorsal fin of a shark or dolphin is the fin on the back that sticks out of the water. **Figure 3.32** shows that for upright, walking animals such as humans, the dorsal part of the brain is the top of the brain. (Picture a person with a dorsal fin sticking out of the top of his or her head!) *Ventral* is the opposite of dorsal, hence it refers to the lower part of the brain.

Applying this idea of *what* and *where* pathways to our example of a person picking up a cup of coffee, the *what* pathway would be involved in the initial perception of the cup and the *where* pathway in determining its location—important information if we are going to carry out the action of reaching for the cup. In the next section, we consider another physiological approach to studying perception and action by describing how studying the behavior of a person with brain damage provides further insights into what is happening in the brain as a person reaches for an object.

Perception and Action Streams

David Milner and Melvyn Goodale (1995) used the neuropsychological approach (studying the behavior of people with brain damage) to reveal two streams, one involving the temporal lobe and the other involving the parietal lobe. The researchers studied D.F., a 34-year-old woman who suffered damage to her temporal lobe from carbon monoxide poisoning caused by a gas leak in her home. One result of the brain damage was revealed when D.F.
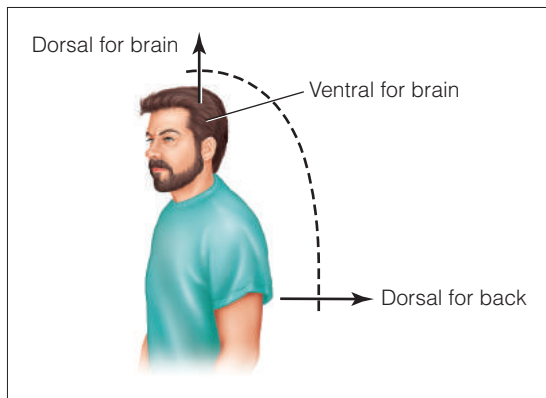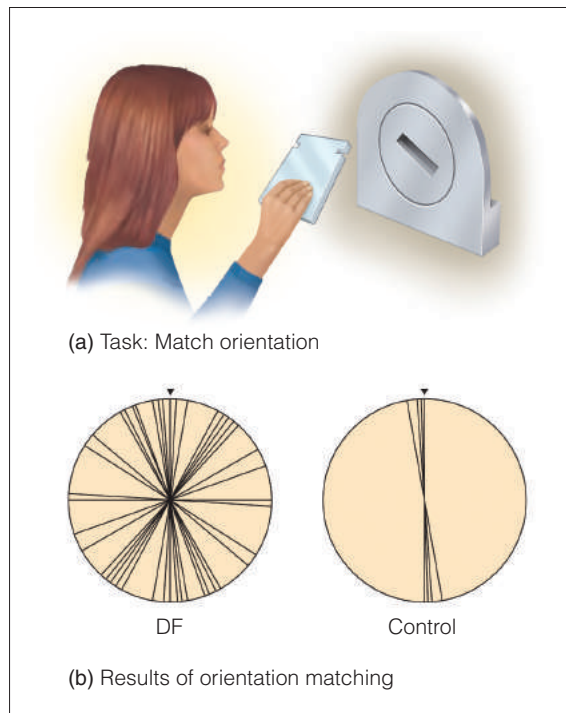
➤ **Figure 3.32** *Dorsal* refers to the back surface of an organism. In upright standing animals such as humans, dorsal refers to the back of the body *and* to the top of the head, as indicated by the arrows and the curved dashed line. Ventral is the opposite of dorsal.

was asked to rotate a card held in her hand to match different orientations of a slot (**Figure 3.33a**). She was unable to do this, as shown in the left circle in **Figure 3.33b**. Each line in the circle indicates how D.F. adjusted the card's orientation. Perfect matching performance would be indicated by a vertical line for each trial, but D.F.'s responses are widely scattered. The right circle shows the accurate performance of the normal controls.

Because D.F. had trouble rotating a card to match the orientation of the slot, it would seem reasonable that she would also have trouble *placing* the card through the slot because to do this she would have to turn the card so that it was lined up with the slot. But when D.F. was asked to "mail" the card through the slot (**Figure 3.34a**), she could do it, as indicated by the results in **Figure 3.34b**. Even though D.F. could not turn the card to match the slot's orientation, *once she started moving the card toward the slot*, she was able to rotate it to match the orientation of the slot. Thus, D.F. performed poorly in the static orientation matching task but did well as soon as *action* was involved (Murphy, Racicot, & Goodale, 1996). Milner and Goodale interpreted D.F.'s behavior as showing that there is one mechanism for judging orientation and another for coordinating vision and action.

Based on these results, Milner and Goodale suggested that the pathway from the visual cortex to the temporal lobe (which was damaged in D.F.'s brain) be called the **perception pathway** and the pathway from the visual cortex to the parietal lobe (which was intact in D.F.'s brain) be called the **action pathway** (also called the *how pathway* because it is associated with how the person takes action). The perception pathway corresponds to the *what* pathway we described in conjunction with the monkey experiments, and the action pathway

(a) Task: Match orientation

DF   Control

(b) Results of orientation matching

(a) Task: "Mail" card in slot

DF   Control

(b) Results of active mailing

➤ **Figure 3.33** (a) D.F.'s orientation task. A number of different orientations were presented. D.F.'s task was to rotate the card to match each orientation. (b) Results for the orientation task. Correct matches are indicated by vertical lines.

(Source: Based on A. D. Milner & M. A. Goodale, 1995)
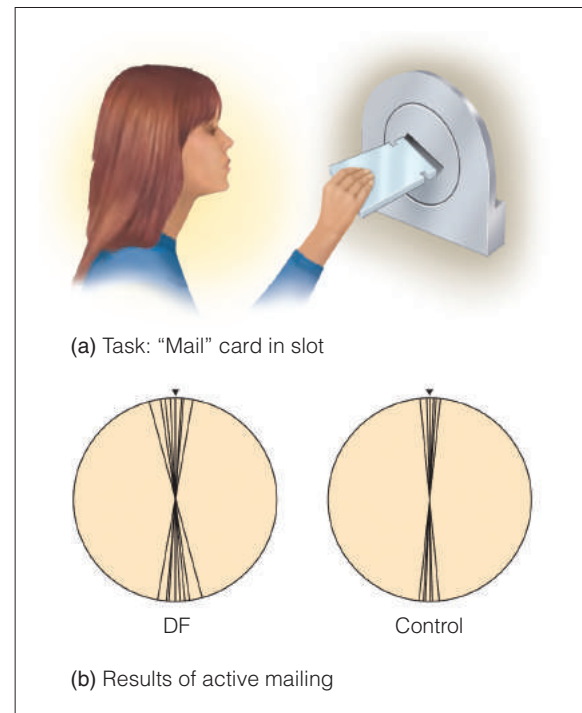
➤ **Figure 3.34** (a) D.F.'s "mailing" task. A number of different orientations were presented. D.F.'s task was to "mail" the card through the slot. (b) Results for the mailing task. Correct orientations are indicated by vertical lines.

(Based on A. D. Milner & M. A. Goodale, 1995)

corresponds to the *where* pathway. Thus, some researchers refer to *what* and *where* pathways and some to *perception* and *action* pathways. Whatever the terminology, the research shows that perception and action are processed in two separate pathways in the brain.

With our knowledge that perception and action involve two separate mechanisms, we can add physiological notations to our description of picking up the coffee cup (Figure 3.29) as follows: The first step is to identify the coffee cup among the vase of flowers and the glass of orange juice on the table (*perception* or *what pathway*). Once the coffee cup is perceived, we reach for the cup (*action* or *where pathway*), taking into account its location on the table. As we reach, avoiding the flowers and orange juice, we position our fingers to grasp the cup (*action pathway*), taking into account our perception of the cup's handle (*perception pathway*), and we lift the cup with just the right amount of force (*action pathway*), taking into account our estimate of how heavy it is based on our perception of the fullness of the cup (*perception pathway*).

Thus, even a simple action like picking up a coffee cup involves a number of areas of the brain, which coordinate their activity to create perceptions and behaviors. A similar coordination between different areas of the brain also occurs for the sense of hearing, so hearing someone call your name and then turning to see who it is activates two separate pathways in the auditory system—one that enables you to hear and identify the sound (the auditory *what* pathway) and another that helps you locate where the sound is coming from (the auditory *where* pathway) (Lomber & Malhotra, 2008).

The discovery of different pathways for perceiving, determining location, and taking action illustrates how studying the physiology of perception has helped broaden our conception far beyond the old "sitting in the chair" approach. Another physiological discovery that has extended our conception of visual perception beyond simply "seeing" is the discovery of mirror neurons.

Mirror Neurons

In 1992, G. di Pelligrino and coworkers were investigating how neurons in the monkey's premotor cortex (**Figure 3.35a**) fired as the monkey performed an action like picking up a piece of food. **Figure 3.35b** shows how a neuron responded when the monkey picked up food from a tray—a result the experimenters had expected. But as sometimes happens in science, they observed something they didn't expect. When one of the experimenters picked up a piece of food while the monkey was watching, the same neuron fired (**Figure 3.35c**). What was so unexpected was that the neurons that fired to observing the experimenter pick up the food were the same ones that had fired earlier when the monkey had picked up the food.
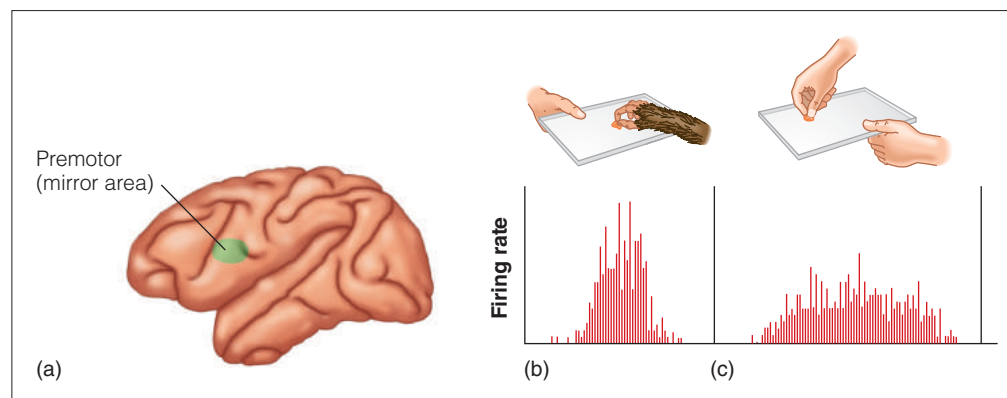
➤ **Figure 3.35** (a) Location of the monkey's premotor cortex. (b) Responses of a mirror neuron when the monkey grasps food on a tray and (c) when the monkey watches the experimenter grasp the food.
(Source: Rizzolatti et al., 2000)
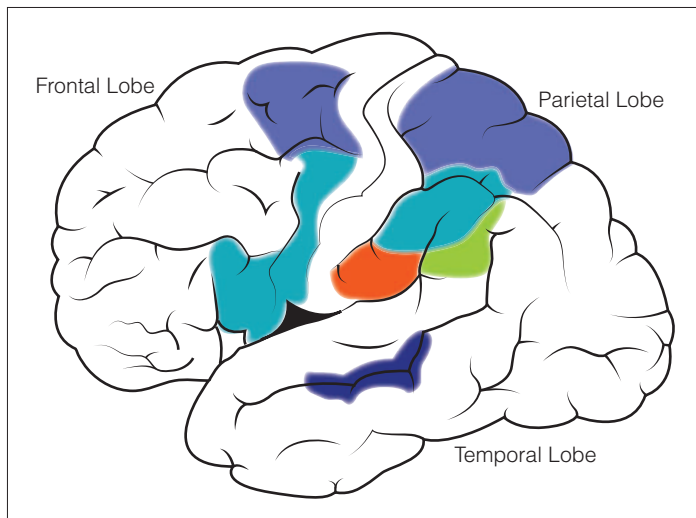
➤ **Figure 3.36** Cortical areas in the human brain associated with the mirror neuron system. Colors indicate the type of actions processed in each region. Turquoise: movement directed toward objects; purple: reaching movements; orange: tool use; green: movements not directed toward objects; blue: upper limb movements.

(Source: Adapted from Cattaneo & Rizzolatti, 2009)

This initial observation, followed by many additional experiments, led to the discovery of **mirror neurons**—neurons that respond both when a monkey observes someone else grasping an object such as food on a tray and when the monkey itself grasps the food (Gallese et al., 1996; Rizzolatti et al., 2006; Rizzolatti & Sinigaglia, 2016). They are called mirror neurons because the neuron's response to watching the experimenter grasp an object is similar to the response that would occur if the monkey were performing the same action. Although you might think that the monkey may have been responding to the anticipation of receiving food, the type of object made little difference. The neurons responded just as well when the monkey observed the experimenter pick up an object that was not food.

At this point, you might be wondering whether mirror neurons are present in the human brain. Some research with humans does suggest that our brains contain mirror neurons. For example, researchers who were using electrodes to record the brain activity in people with epilepsy in order to determine which part of their brains was generating their seizures have recorded activity from neurons with the same mirror properties as those identified in monkeys (Mukamel et al., 2010). Additional work done using fMRI in neurologically normal people has further suggested that these neurons are distributed throughout the brain in a network that has been called the **mirror neuron system** (**Figure 3.36**) (Caspers et al., 2010; Cattaneo & Rizzolatti, 2009; Molenbergs et al., 2012).

What is the purpose of these mirror neurons? One suggestion is that they are involved in determining the *goal* or *intention* behind an action. To understand what this means, let's return to Crystal reaching for her coffee cup. She could be reaching for the cup for a number of reasons. Maybe she intends to drink some coffee, although if we notice that the cup is empty, we might instead decide that she is going to take the cup back to the counter of the coffee shop to get a refill, or if we know that she never drinks more than one cup, we might decide that she is going to place the cup in the used cup bin. Thus, a number of different intentions can be associated with perception of the same action.

What is the evidence that the response of mirror neurons can be influenced by different intentions? Mario Iacoboni and coworkers (2005) did an experiment in which they measured participants' brain activity as they watched short film clips. There were three versions of the film, all showing the same motion of a hand picking up a cup, but in different contexts. Version 1 showed a hand reaching to pick up a full cup of coffee from a neatly set up table, with food on a plate. Version 2 showed the same motion but the cup was on a messy table, the food was eaten, and the cup was empty. Version 3 showed the hand picking up an isolated cup. Iacoboni hypothesized that viewing film clip 1 would lead the viewer to infer that the person picking up the cup intends to drink from it, that viewing film clip 2 would lead the viewer to infer that the person is cleaning up, and that viewing film clip 3 would lead to no particular inference.

When Iacoboni compared the brain activity from viewing the two intention films to the activity from the non-intention film, he found that the intention films caused greater activity than the non-intention film in areas of the brain known to have mirror neuron properties. The amount of activity was least for the non-intention film, higher for the cleaning-up film, and was highest for the drinking film. Based on the increased activity for the two intention films, Iacoboni concluded that the mirror neuron area is involved with understanding the intentions behind the actions shown in the films. He reasoned that if the mirror neurons were just signaling the action of picking up the cup, then a similar response would occur regardless of whether a context surrounding the cup was present.

Mirror neurons, according to Iacoboni, code the "why" of actions and respond differently to different intentions (also see Fogassi et al., 2005 for a similar experiment on a monkey).

If mirror neurons do, in fact, signal intentions, how do they do it? One possibility is that the response of these neurons is determined by the sequence of motor activities that could be *expected* to happen in a particular context (Fogassi et al., 2005; Gallese, 2007). For example, when a person picks up a cup with the intention of drinking, the next expected actions would be to bring the cup to the mouth and then to drink some coffee. However, if the intention is to clean up, the expected action might be to carry the cup over to the sink. According to this idea, mirror neurons that respond to different intentions are responding to the action that is happening *plus* the sequence of actions that is most likely to follow, given the context.

When considered in this way, the operation of mirror neurons shares something with perception in general. Remember Helmholtz's likelihood principle—we perceive the object that is *most likely* to have caused the pattern of stimuli we have received. In the case of mirror neurons, the neuron's firing may be based on the sequence of actions that are *most likely* to occur in a particular context. In both cases the outcome—either a perception or firing of a mirror neuron—depends on knowledge that we bring to a particular situation.

The exact functions of mirror neurons in humans are still being debated, with some researchers assigning mirror neurons a central place in determining intentions (Caggiano et al., 2011; Gazzola et al., 2007; Kilner, 2011; Rizzolatti & Sinigaglia, 2016) and others questioning this idea (Cook et al., 2014; Hickock, 2009). But whatever the exact role of mirror neurons in humans, there is no question that there is some mechanism that extends the role of perception beyond providing information that enables us to take action, to yet another role—inferring why other people are doing what they are doing.

▶ SOMETHING TO CONSIDER: KNOWLEDGE, INFERENCE, AND PREDICTION

"Brains, it has recently been argued, are essentially prediction machines" Clark (2013)

Two terms that have appeared throughout this chapter are *knowledge* and *inference*. Knowledge was the foundation of Helmholtz's theory of unconscious inference, and the basis of the likelihood principle. Inference depends on knowledge. For example, we saw how inference based on knowledge helps resolve the ambiguity of the retinal image and how knowledge of transitional probabilities helps us infer where one word in a conversation ends and the other begins. Knowledge and the inferences that follow are the basis of top-down processing (p. 67).

Another way to think about knowledge and inference is in terms of *prediction*. After all, when we say that a particular retinal image is caused by a book (Figure 3.7), we are making a prediction of what is probably out there. When we say that a briefly presented shape on a kitchen counter is probably a loaf of bread (Figure 1.13), we are making a prediction based on what is likely to be sitting on a kitchen counter. We are making predictions about what is out there constantly, which is the basis of the assertion that "brains . . . . are essentially prediction machines" at the beginning of this section (Clark, 2013).

A hint that prediction extends beyond simply seeing is provided by the **size-weight illusion**: When a person is presented with two similar objects, such as two cubes, that are the same weight but different sizes, the larger one seems lighter when they are lifted together. One explanation for this is that we predict that larger objects will be heavier than smaller objects, because objects of the same type typically get heavier as they get larger (Buckingham et al., 2016; Plaisier & Smeets, 2015). We are therefore surprised when the larger one is lighter than predicted. Just as perception is guided by predictions, so are the actions associated with perceptions.

As it turns out, <mark>prediction is a central principle that operates throughout cognition</mark>. Here is a preview of a few of the predictions we will encounter in the chapters that follow:

➤ Chapter 4 (Attention)—Prediction guides where we direct our eyes as we scan a scene.

➤ Chapter 7 (Memory)—Our ability to predict what might happen in the future is based on our ability to remember events from the past.

➤ Chapter 11 (Language)—Not only does prediction help us perceive individual words in the speech stream, as we saw in this chapter, but it also helps us understand the meanings of sentences, follow conversations, and understand stories.

➤ Chapter 13 (Thinking)—People sometimes use "rules of thumb" called heuristics to make predictions that help them make decisions or determine solutions to problems.

Although the idea of prediction is nothing new, having been proposed by Helmholtz in the 19th century, it has become an important topic across a number of different areas of cognition.

TEST YOURSELF 3.3

1. What is the oblique effect? Describe how this effect could be caused by evolution and by experience.

2. Describe the interaction between perceiving and taking action, giving a specific example from everyday perception.

3. Describe the Ungerleider and Mishkin experiment. How did they use the procedure of brain ablation to demonstrate *what* and *where* streams in the cortex?

4. Describe how Milner and Goodale's testing of D.F. demonstrated pathways for matching orientation and for combining vision and action. Describe the perception pathway and the action pathway. How do these pathways correspond to Ungerleider and Mishkin's *what* and *where* streams?

5. Describe how the perception and action pathways both play a role in an action such as picking up a cup of coffee.

6. What are mirror neurons? What have some researchers proposed about how mirror neurons might link perception and action?

7. What is the connection among knowledge, inference, and prediction?


CHAPTER SUMMARY

1. The example of Crystal running on the beach and having coffee later illustrates how perception can change based on new information, how perception can be based on principles that are related to past experiences, how perception is a process, and how perception and action are connected.

2. We can easily describe the relation between parts of a city scene, but it is often challenging to indicate the reasoning that led to the description. This illustrates the need to go beyond the pattern of light and dark in a scene to describe the process of perception.

3. Attempts to program computers to recognize objects have shown how difficult it is to program computers to perceive at a level comparable to humans. A few of the difficulties facing computers are (1) <mark>the stimulus on the receptors is ambiguous, as demonstrated by the inverse projection problem; (2) objects in a scene can be hidden or blurred; (3) objects look different from different viewpoints; and (4) scenes contain high-level information</mark>.

4. Perception starts with bottom-up processing, which involves stimulation of the receptors, creating electrical signals that

reach the visual receiving area of the brain. Perception also involves top-down processing, which is associated with knowledge stored in the brain.

5. Examples of top-down processing are the multiple personalities of a blob and how knowledge of a language makes it possible to perceive individual words. Saffran's experiment has shown that 8-month-old infants are sensitive to transitional probabilities in language.

6. The idea that perception depends on knowledge was proposed by Helmholtz's theory of unconscious inference.

7. The Gestalt approach to perception proposed a number of laws of perceptual organization, which were based on how stimuli usually occur in the environment.

8. Regularities of the environment are characteristics of the environment that occur frequently. We take both physical regularities and semantic regularities into account when perceiving.

9. Bayesian inference is a mathematical procedure for determining what is likely to be "out there"; it takes into account a person's prior beliefs about a perceptual outcome and the likelihood of that outcome based on additional evidence.

10. Of the four approaches to object perception—unconscious inference, Gestalt, regularities, and Bayesian—the Gestalt approach relies more on bottom-up processing than the others. Modern psychologists have suggested a connection between the Gestalt principles and past experience.

11. One of the basic operating principles of the brain is that it contains some neurons that respond best to things that occur regularly in the environment.

12. Experience-dependent plasticity is one of the mechanisms responsible for creating neurons that are tuned to respond to specific things in the environment. The experiments in which people's brain activity was measured as they learned about Greebles supports this idea. This was also illustrated in the experiment described in Chapter 2 in which kittens were reared in vertical or horizontal environments.

13. Perceiving and taking action are linked. Movement of an observer relative to an object provides information about the object. Also, there is a constant coordination between perceiving an object (such as a cup) and taking action toward the object (such as picking up the cup).

14. Research involving brain ablation in monkeys and neuropsychological studies of the behavior of people with brain damage have revealed two processing pathways in the cortex—a pathway from the occipital lobe to the temporal lobe responsible for perceiving objects, and a pathway from the occipital lobe to the parietal lobe responsible for controlling actions toward objects. These pathways work together to coordinate perception and action.

15. Mirror neurons are neurons that fire both when a monkey or person takes an action, like picking up a piece of food, and when they observe the same action being carried out by someone else. It has been proposed that one function of mirror neurons is to provide information about the goals or intentions behind other people's actions.

16. Prediction, which is closely related to knowledge and inference, is a mechanism that is involved in perception, attention, understanding language, making predictions about future events, and thinking.

THINK ABOUT IT

1. Describe a situation in which you initially thought you saw or heard something but then realized that your initial perception was in error. (Two examples: misperceiving an object under low-visibility conditions; mishearing song lyrics.) What were the roles of bottom-up and top-down processing in this situation of first having an incorrect perception and then realizing what was actually there?

2. Look at the picture in **Figure 3.37**. Is this a huge giant's hand getting ready to pick up a horse, a normal-size hand picking up a tiny plastic horse, or something else? Explain, based on some of the things we take into account in

addition to the image that this scene creates on the retina, why it is unlikely that this picture shows either a giant hand or a tiny horse. How does your answer relate to top-down processing?

3. In the section on experience-dependent plasticity, it was stated that neurons can reflect knowledge about properties of the environment. Would it be valid to suggest that the response of these neurons represents top-down processing? Why or why not?

4. Try observing the world as though there were no such thing as top-down processing. For example, without the aid of
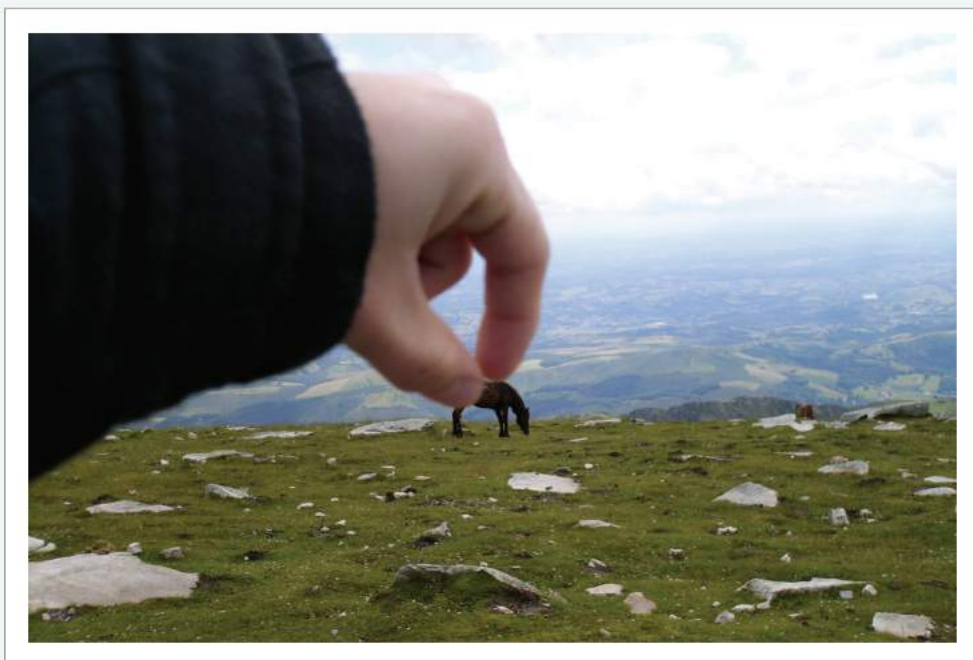
➤ **Figure 3.37** Is a giant hand about to pick up the horse?

top-down processing, seeing a restaurant's restroom sign that says "Employees must wash hands" could be taken to mean that we should wait for an employee to wash our hands! If you try this exercise, be warned that it is extremely difficult because top-down processing is so pervasive in our environment that we usually take it for granted.

KEY TERMS

Action pathway, 84

Apparent movement, 71

Bayesian inference, 76

Bottom-up processing, 67

Brain ablation, 82

Direct pathway model, 00

Dorsal pathway, 83

Gestalt psychologists, 71

Inverse projection problem, 65

Law of pragnanz, 72

Landmark discrimination problem, 82

Light-from-above assumption, 74

Likelihood, 76

Likelihood principle, 70

Mirror neuron, 85

Mirror neuron system, 86

Object discrimination problem, 82

Oblique effect, 74

Perception, 60

Perception pathway, 84

Physical regularities, 74

Placebo, 00

Placebo effect, 00

Principle of good continuation, 72

Principle of good figure, 72

Principle of similarity, 73

Principle of simplicity, 72

Principles of perceptual organization, 71

Prior, 76

Prior probability, 76

Regularities in the environment, 74

Scene schema, 76

Semantic regularities, 74

Size-weight illusion, 87

Speech segmentation, 68

Statistical learning, 68

Theory of natural selection, 79

Top-down processing, 67

Transitional probabilities, 68

Unconscious inference, 70

Ventral pathway, 83

Viewpoint invariance, 66

What **pathway,** 82

Where **pathway,** 83