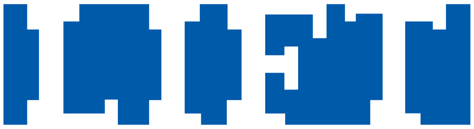
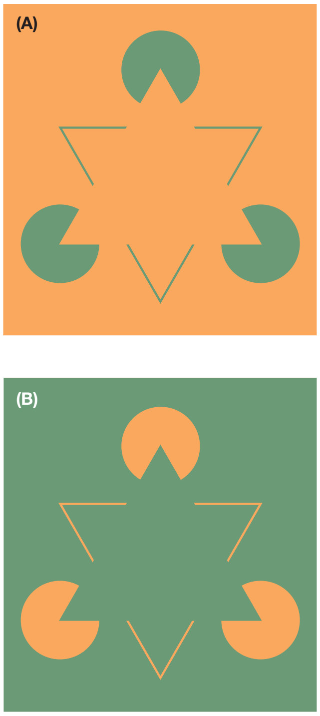


5.4 A hidden figure At first the figure seems not to contain the features needed to identify the various letters. But once the figure is reorganized with the white parts forming the figure and not the dark parts, its features are easily detected. So it seems that the analysis of features depends on a preliminary step in which the viewer must organize the figure.



5.5 Subjective contours In (A) we see an orange triangle whose vertices lie on top of the three green circles. The three sides of this orange triangle (which looks brighter than the orange background) are clearly visible, even though they don't exist physically. In (B) we see the same effect with green and orange reversed. Here, the green triangle—which looks darker than the green background—has subjective green contours.

Gestalt psychology A theoretical approach that emphasizes the role of organized wholes in perception and other psychological processes.

Figure 5.4 illustrates a different type of interpretation. These dark shapes seem meaningless at first, but after a moment most people find a way to reorganize the figure so that the familiar letters come into view. But let's be clear about what this means. At the start, the form doesn't seem to contain the features we need to identify the *L,* the *I,* and so on—and so we don't detect these letters. Once we've reorganized the form, though, it does contain the relevant features and so we immediately recognize the letters. Apparently, therefore, the catalog of features present in this figure depends on how we interpret its overall form. Based on one interpretation, the features defining these letters are absent—and so we can't detect the letters or the word *LIFT.* With a different interpretation, the features are easily visible and we can immediately read the word. It seems, then, that features are as much "in the eye of the beholder" as they are in the figure itself.

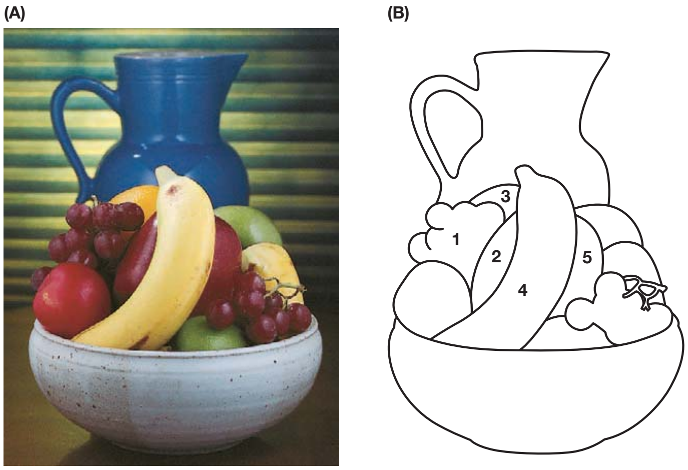
As a related example, consider Figure 5.5. Here, most people easily perceive two complete triangles—one on the orange background, one on the green. But again, the features of these triangles aren't present on the page; specifically, the *sides* of the triangle aren't marked in the figure at all. However, the perceiver organizes the overall form so that the missing sides are filled in—she's essentially creating the features for herself. Once that's done, she can clearly perceive the triangles.

PERCEPTUAL PARSING

The previous examples powerfully suggest that the perception of form depends both on feature input—that is, what's actually in front of your eyes—and on how you organize and interpret the form. But what exactly does it mean for a perceiver to "interpret" a form, or to find an "organization" within a figure? And why do you end up with one interpretation and not another? Why, for example, do most people decide that Figure 5.5 should be organized in a way that connects the angles so that they become parts of a single form rather than treating them as separate forms?

Questions like these were crucial for **Gestalt psychology,** a school of psychology that emphasized that organization is an essential feature of all mental activity: We understand the elements of the visual input as linked to each other in a certain way, and the identity of these elements depends on the linkage. (That's why in Figure 5.5, we perceive the round elements as intact circles, each partially hidden by another form, rather than as a series of "Pac-Man" figures.) Likewise, we appreciate a work of music because we perceive the individual notes as forming a cohesive whole. Similarly, our thoughts have meaning only in relationship to each other. In all cases, the Gestalt psychologists wanted to ask how this organization was achieved, and how it influenced us. (The word *Gestalt* is derived from a German word meaning "form" or "appearance.")

Gestalt psychologists described several aspects of this organization and identified several principles that guided it. Some of the principles are concerned with the way you *parse* the input—that is, how you separate a scene into individual objects, linking together the parts of each object but not linking one object's parts to some other object. To make this idea concrete, consider the still life in Figure 5.6. To make sense of this picture, your perception must somehow group the elements of the scene appropriately.

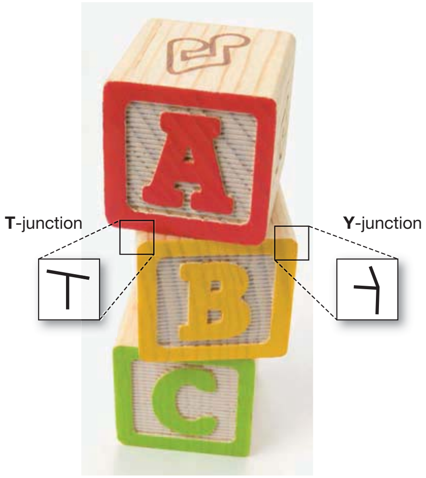


5.6 Perceptual parsing (A) A still life. (B) An overlay designating five different segments of the scene shown in (A). To determine what an object is, the perceptual system must first decide what goes with what: Does portion 2 go with 1 or with 3, 4, or 5? Or does it go with none of them?

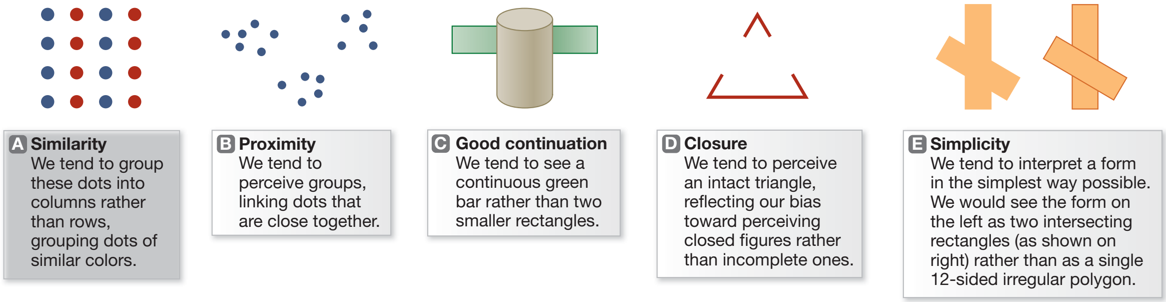
Portion 2 (part of the apple) must be united with portion 5 (more of the apple), even though they're separated by portion 4 (a banana). Portion 2 should not be united with portion 1 (a bunch of grapes), even though they're adjacent and about the same color. The bit of the apple hidden from view by the banana must somehow be filled in, so that you perceive an intact apple rather than two apple slices. All of these steps involved in deciding which bits go with which other bits fall under the label of parsing.

similarity In perception, a principle by which we tend to group like figures, especially by color and orientation.

What cues guide you toward parsing a stimulus pattern one way rather than another? The answer involves both feature information and information about the larger-scale pattern. For example, we tend to interpret certain features (such as a T-junction—Figure 5.7) as indicating that one edge has disappeared behind another; we interpret other features differently. In this way, "local" information—information contained in one small part of the scene—helps guide our parsing. But "global" information—information about the whole scene—is also crucial. For example, perceivers tend to group things together according to a principle of **similarity**—meaning that, all other things being equal, they group together figures that



5.7 How features guide parsing We noted earlier that feature analysis depends on a preliminary step in which the viewer organizes the overall figure. But it turns out that the opposite is also true: The features determine how the viewer organizes the figure. For example, viewers usually interpret a T-junction as one surface dropping from view behind another. They usually interpret a Y-junction as a corner pointing toward them.



A Similarity We tend to group these dots into columns rather than rows, grouping dots of similar colors.

B Proximity We tend to perceive groups, linking dots that are close together.

C Good continuation We tend to see a continuous green bar rather than two smaller rectangles.

D Closure We tend to perceive an intact triangle, reflecting our bias toward perceiving closed figures rather than incomplete ones.

E Simplicity We tend to interpret a form in the simplest way possible. We would see the form on the left as two intersecting rectangles (as shown on right) rather than as a single 12-sided irregular polygon.

5.8 Other Gestalt principles

resemble each other. So in Figure 5.8A, we group blue dots with blue dots, red with red. Perceivers are also influenced by **proximity**—the closer two figures are to each other, the more we tend to group them together perceptually (for more on these principles of perceptual organization, see Figure 5.8; Palmer, 2002; Wertheimer, 1923).

Parsing is also guided by several other principles, including **good continuation**—a preference for organizations in which contours continue smoothly along their original course. This helps us understand why portions 2 and 5 in Figure 5.6 are grouped together as parts of a single object; but good continuation can also be documented in much simpler stimuli (Figure 5.8C). Good continuation is also relevant to Figure 5.5. Some theorists interpret the **subjective contours** visible in this figure as a special case of good continuation. In their view, the contour is seen to continue along its original path—even, if necessary, jumping a gap or two to achieve this continuation (Kellman & Shipley, 1991; Kellman, Garrigan, & Shipley, 2005).

proximity In perception, the closeness of two figures. The closer together they are, the more we tend to group them together perceptually.

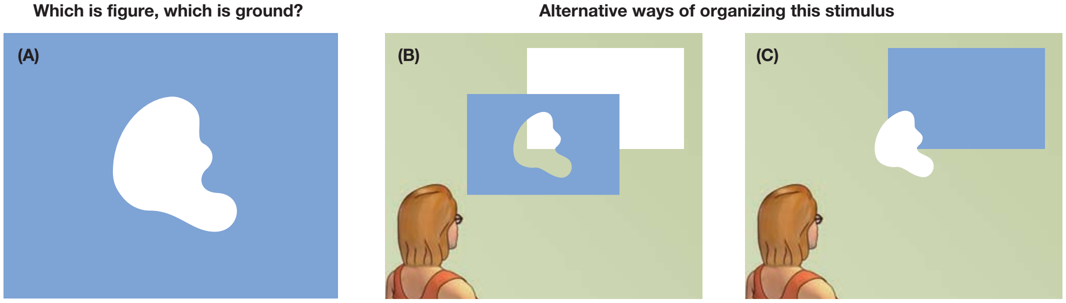
good continuation A factor in visual grouping; we tend to perceive contours in a way that alters their direction as little as possible.

subjective contours Perceived contours that do not exist physically. We tend to complete figures that have gaps in them by perceiving a contour as continuing along its original path.

FIGURE AND GROUND

Another part of visual organization is the separation of the object from its setting, so that the object is seen as a coherent whole, separate from its background. This separation of *figure* and *ground* allows you to focus on just the banana in Figure 5.6, treating everything else in the scene as merely the backdrop for the banana. But the separation of figure and ground is just as important with simpler and entirely unfamiliar figures. In Figure 5.9A, the white splotch appears to most people as the figure and is

5.9 Figure and ground (A) One of the early steps in seeing a form is to segregate it from its background. If we perceive the figure in part (A) as a blue rectangle with a hole in it (B), the edge marks the contour of the hole. The situation is reversed in (C). Now the edge marks the white blob, not a break in the blue background. In this sense, the edge belongs to the figure, not the ground. As it turns out, the perception in (C) is much more likely.



perceived as closer to the viewer than the blue region (which is seen as the ground) as shown in 5.9C. The edge between the blue and white regions is perceived as part of the figure, defining its shape. The same edge does not mark a contour for the blue region but merely marks the point at which the blue region drops from view.

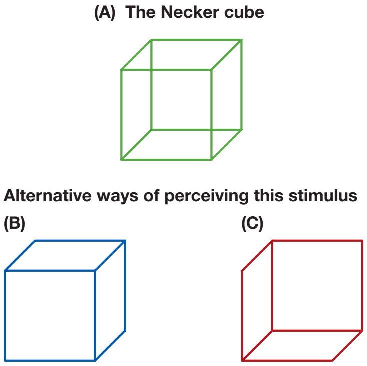
Of course, you can usually identify a figure so quickly and easily that it feels like this specification is somehow specified by the stimulus itself and is not an element of your interpretation. But the fact remains that identifying the figure, like all aspects of perceptual organization, is up to you. This is most evident whenever you realize there's more than one way to interpret a given stimulus—as in Figure 5.10, which can be seen either as a white vase or as two blue faces in profile. This **reversible figure** makes it clear that the stimulus itself is neutral in its organization. What is figure and what is ground, it seems, depends on how we look at it.

Other examples point to the same broad conclusion and highlight the perceiver's active role in interpreting the input. Is Figure 5.11A—the Necker cube—aligned with the solid cube shown in Figure 5.11B, so we're viewing it from above? Or is it aligned with the cube shown in Figure 5.11C, so we're viewing it from below? Most people can organize the Necker cube in either way, so they first perceive it to have one orientation and then the other. Apparently, then, the organization is not specified by the figure itself but is instead up to the perceiver.

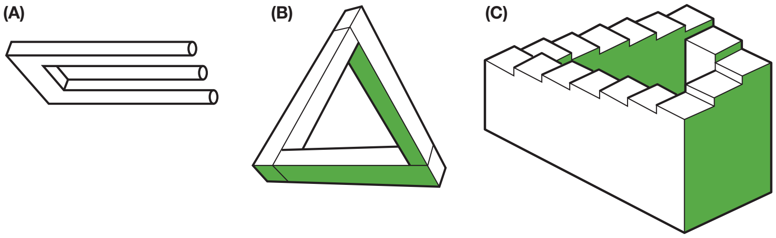
All of these observations suggest that perception is less "objective" than one might suppose, because what we perceive is, it seems, often determined by how we interpret or organize the input. At the same time, it's important to realize that perceivers' inferences and interpretations tend to be neither foolish nor random. Quite the contrary: Our interpretations of the sensory input are, first of all, shaped by our experience; and they're correct far more often than not (Enns, 2004). Likewise, the interpretations themselves tend to be quite logical, as if our visual system always follows certain rules. We've already mentioned some of these rules—a preference for grouping *similar* things together, for example, or a preference for parsing the input so that it creates smooth contours. But other rules also guide us: For example, we seem to prefer perceptual interpretations that explain *all* the information contained within the stimulus, and so we avoid interpretations that would explain only bits and pieces of the stimulus. We also seem to avoid interpretations that would involve some contradiction, such as perceiving a surface to be both opaque and transparent. What's more, we seem to avoid interpretations that depend on accident or coincidence. ("This is what the form would look like if viewed from exactly the right position.") Of course, no one claims that the perceptual apparatus is literally proceeding through a sequence of logical steps, weighing each of these rules in turn. Still, our perception does seem guided by these principles, so that our interpretations of the input will be logical and usually correct (Figure 5.12).

5.10 Reversible figure-ground pattern This figure can be seen as either a pair of silhouetted faces or a white vase.

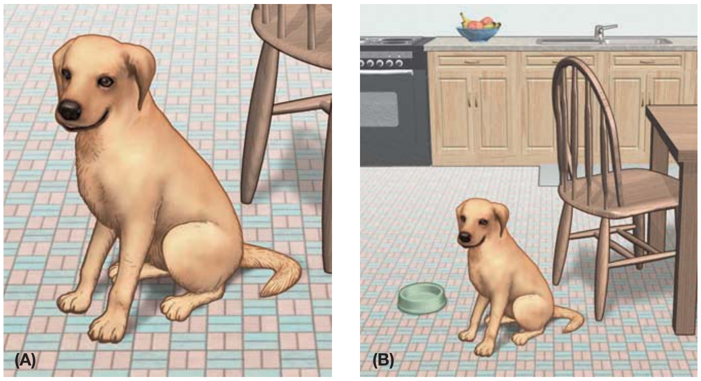
reversible figure A visual pattern that easily allows more than one interpretation, in some cases changing the specification of figure and ground, in other cases changing the perceived organization in depth.



5.11 The Necker cube The ambiguous Necker cube, shown in (A), can be perceived as aligned with either the cube shown in (B) or the one in (C).



5.12 Impossible figures We've mentioned how "logical" the perceptual system seems to be, but it's important to realize that this logic has limits. As an example, consider these so-called impossible figures. We perceive them as if they show three-dimensional objects, although contradictions within each figure guarantee that they can't be three-dimensional.



5.21 An invariant relationship that provides information about size (A) and (B) show a dog at different distances from the observer. The retinal size of the dog varies with distance, but the ratio between the retinal size of the dog and the retinal size of the textural elements (e.g., the floor tiles) is constant.

Unconscious Inference

How do we achieve each of these forms of constancy? One hypothesis focuses on *relationships* within the retinal image. In judging size, for example, we might be helped by the fact that we generally see objects against some background, and various elements in the background can provide a basis for comparison with the target object. Thus the dog sitting nearby on the kitchen floor is half as tall as the chair and hides a number of the kitchen's floor tiles from view. If we take several steps back from the dog, none of these relationships changes, even though the sizes of all the retinal images are reduced (Figure 5.21). Size constancy, therefore, might be achieved by focusing not on the images themselves but on these unchanging relationships.

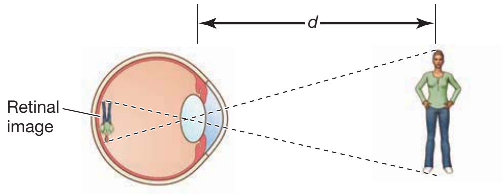
Relationships do contribute to size constancy, and that's why we are better able to judge size when comparison objects are in view or when the target we're judging sits on a surface that has a uniform visual texture (like the floor tiles in the example). But these relationships don't tell the whole story. Size constancy is found even when the visual scene offers no basis for comparison—if, for example, the object to be judged is the only object in view—provided that other cues signal the *distance* of the target object (Chevrier & Delorme, 1983; Harvey & Leibowitz, 1967; Holway & Boring, 1947).

How might our visual system use this distance information? More than a century ago, the German physicist Hermann von Helmholtz developed an influential hypothesis regarding this question. Helmholtz started with the fact that there's a simple inverse relationship between distance and retinal image size: If an object doubles its distance from the viewer, the size of its image is reduced by half. If an object triples its distance, the size of its image is reduced to a third of its initial size. This relationship is guaranteed to hold true because of the principles of optics, and the relationship makes it possible for perceivers to achieve size constancy by means of a simple calculation. Of course, Helmholtz knew that we don't run through a conscious calculation every time we perceive an object's size; but he believed we were calculating nonetheless—and so he referred to the process as an **unconscious inference** (Helmholtz, 1909).

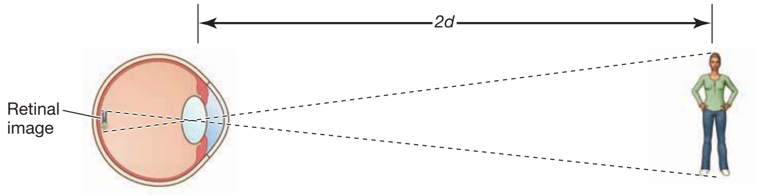
What is the calculation that allows someone to perceive size correctly? It's simply multiplication: the size of the image on the retina, multiplied by the distance between you and the object. (We'll have more to say about how you know this distance in a later section.) Thus, imagine an object that, at a distance of 10 feet, casts an image on the

unconscious inference A process postulated by Hermann von Helmholtz to explain certain perceptual phenomena such as size constancy. For example, an object is perceived to be at a certain distance and this is unconsciously taken into account in assessing its retinal image size, with the result that size constancy is maintained.

(A) Closer objects cast larger retinal images



(B) Father objects cast smaller retinal images



5.22 The relationship between image size and distance If an object moves to a new distance, the size of the retinal image cast by that object changes. A doubling of the distance reduces the retinal image by half. If the distance is tripled, the retinal image is cut to one-third of its initial size.

retina that's 4 millimeters across (Figure 5.22). The same object, at a distance of 20 feet, casts an image of 2 millimeters. In both cases, the product—10 × 4 or 20 × 2—is the same. If, therefore, your size estimate depends on that product, your size estimate won't be thrown off by viewing distance—and of course, that's exactly what we want.

What's the evidence that size constancy does depend on this sort of inference? In many experiments, researchers have shown people some object and, without changing the object's retinal image, changed the apparent distance of the object. (There are many ways to do this—lenses that change how the eye has to focus to bring the object into sharp view, or mirrors that change how the two eyes have to angle inward so that the object's image is centered on both foveas.) If people are—as Helmholtz proposed—using distance information to judge size, then these manipulations should affect size perception. Any manipulation that makes an object seem farther away (without changing retinal image size) should make that object seem bigger. Any manipulation that makes the object seem closer should make it look smaller. And, in fact, these predictions are correct—a powerful confirmation that we do use distance to judge size.

A similar proposal explains how people achieve shape constancy. Here, we take the slant of the surface into account and make appropriate adjustments—again, an unconscious inference—in our interpretation of the retinal image's shape. Likewise for brightness constancy: We seem to be quite sensitive to how a surface is oriented relative to the available light sources, and we take this information into account in estimating how much light is reaching the surface. Then we use this assessment of lighting to judge the surface's brightness (e.g., whether it's black or gray or white). In all these cases, therefore, it appears that our perceptual system does draw some sort of unconscious inference, taking our viewing circumstances into account in a way that allows us to perceive the constant properties of the visual world.

Illusions

This process of taking information into account—no matter whether we're taking viewing distance into account, or viewing angle, or illumination—is crucial for achieving constancy. More than that, it's yet another indication that we don't just "receive"

DISTANCE PERCEPTION: WHERE IS IT?

So far in this chapter, we've emphasized how you recognize the objects you encounter. This focus has led us to consider how you manage to perceive forms as well as how you cope with variations in viewing circumstances in order to perceive an object's shape and size correctly. And once again, this discussion leads to a new question: To perceive *what* something is, you need to achieve constancy. But, to achieve constancy, you need to perceive *where* something is—how far it is from you (so that you can achieve size constancy) and how it is angled relative to your line of view (so that you can achieve shape constancy).

Of course, information about where things are in your world is also valuable for its own sake. If you want to walk down a hallway without bumping into things, you need to know which obstacles are close to you and which ones are far off. If you wish to caress a loved one, you need to know where he or she is; otherwise, you're likely to poke him or her in the eye. Plainly, then, you need to know where objects in your world are located.

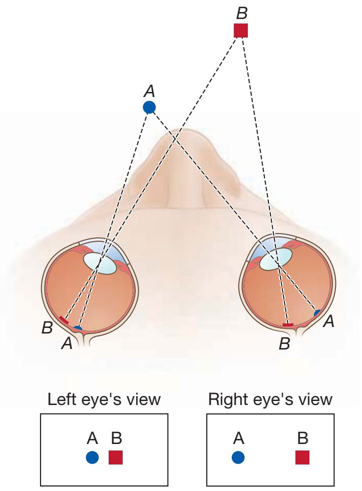
How, therefore, do you manage to perceive a three-dimensional world, judging which objects are close and which are far? The answer centers on **depth cues**—features of the stimulus that indicate an object's position. What are these cues?

depth cues Sources of information that signal the distance from the observer to the distal stimulus.

binocular disparity A depth cue based on the differences between the two eyes' views of the world. This difference becomes less pronounced the farther an object is from the observer.

Binocular Cues

One important cue for distance comes from the fact that our two eyes look out onto the world from slightly different positions; as a result, each eye has a slightly different view. This difference between the two eyes' views is called **binocular disparity**, and it gives us important information about distance relationships in the world (Figure 5.26).



5.26 Binocular disparity Two images at different distances from the observer will present somewhat different retinal images. In the left eye's view, these images are close together on the retina; in the right eye's view, the images are farther apart. This disparity between the views serves as a powerful cue for depth.

Binocular disparity can induce the perception of depth even when no other distance cues are present. For example, the bottom panels of Figure 5.26 show the views that each eye would receive while looking at a pair of nearby objects. If we present each of these views to the appropriate eye (e.g., by drawing the views on two cards and placing one card in front of each eye), we can obtain a striking impression of depth.

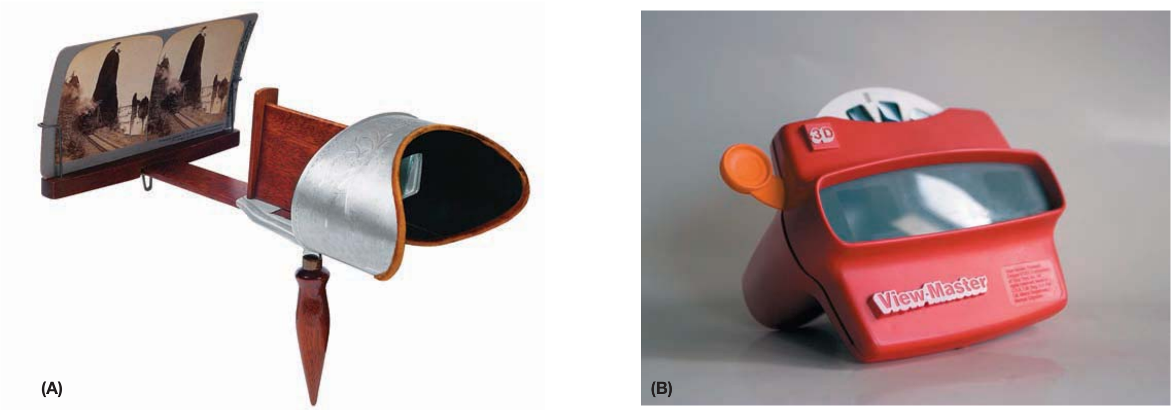
Disparity was the principle behind the stereoscope, a device popular in the 19th century (Figure 5.27), which presented a slightly different photograph to each eye and so created a vivid sense of depth. The same principle is used in 3-D movies, in which two different movies—presenting two slightly different views of each scene—are projected simultaneously onto the theatre's screen. For these movies, viewers wear special glasses to ensure that their left eye sees one of the movies and their right eye sees the other. In this way, each eye gets the appropriate input and creates the binocular disparity that in turn produces a compelling perception of depth.

Monocular Cues

Binocular disparity has a powerful effect on the way we perceive depth. But we can also perceive depth with one eye closed; so, clearly, there must be cues for depth that depend only on what each eye sees by itself. These are the **monocular depth cues.**

monocular depth cues Features of the visual stimulus that indicate distance even if the stimulus is viewed with only one eye.

One of the monocular depth cues depends on the adjustment that the eye must make to see the world clearly. Specifically, we've already mentioned that in each eye, muscles adjust the shape of the lens to produce a sharply focused image on the retina. The amount of adjustment depends on how far away the viewed object is—there's a lot of adjustment for nearby objects, less for those a few steps away, and virtually no adjustment at all for objects more than a few meters away. It turns out that perceivers



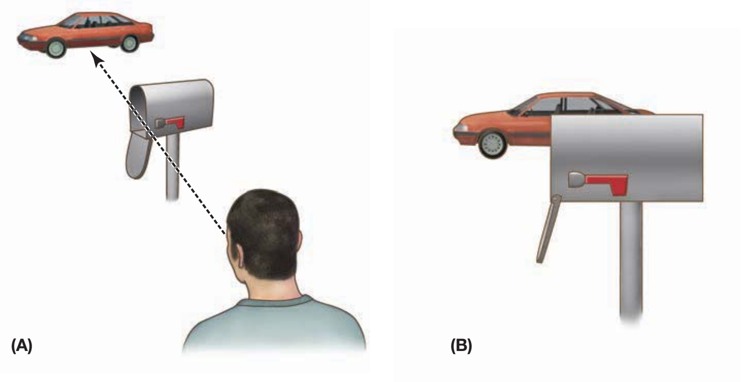
5.27 Stereoscope and View-Master After their invention in 1833, stereoscopes were popular for many years. They work by presenting one picture to the left eye and another to the right; the disparity between the pictures creates a vivid sense of depth. The View-Master, a popular children's toy, works exactly the same way. The photos on the wheel are actually pairs—at any rotation, the left eye views the leftmost photo (the one at 9 o'clock on the wheel) and the right eye views the rightmost photo (the one at 3 o'clock).

are sensitive to the amount of adjustment and use it as a cue indicating how far away the object is.

Another set of monocular cues have been exploited for centuries by artists to create an impression of depth on a flat surface—that is, within a picture—which is why these cues are often called **pictorial cues.** In each case, these cues rely on straightforward principles of physics. For example, imagine a situation in which a man is trying to admire a sports car, but a mailbox is in the way (Figure 5.28A). In this case, the mailbox will inevitably block the view simply because light can't travel through an opaque object. This fact about the physical world provides a cue we can use in judging distance. The cue is known as **interposition** (Figure 5.28B)—the blocking of our view of one object by some other object. In this example, interposition tells the man that the mailbox is closer than the car.

pictorial cues Patterns that can be represented on a flat surface in order to create a sense of a three-dimensional object or scene.

interposition A monocular cue to distance that relies on the fact that objects farther away are blocked from view by closer objects.



5.28 Pictorial cues (A) This man is looking at the sports car, but the mailbox blocks part of his view. (B) Here's how this scene looks from the man's point of view. Because the mailbox blocks the view, we get a simple but powerful cue that the mailbox must be closer to the man than the sports car is.

5.29 Linear perspective as a cue for depth

In the same way, distant objects necessarily produce a smaller retinal image than do nearby objects of the same size; this is a fact about optics. But this physical fact again gives us perceptual information we can use. In particular, it's the basis for the cue of **linear perspective**, the name for the pattern in which parallel lines seem to converge as they get farther and farther from the viewer (Figure 5.29).

One more pictorial cue is provided by *texture gradients*. Consider what meets the eye when we look at cobblestones on a street or patterns of sand on a beach. The retinal projection of the sand or the cobblestones shows a pattern of continuous change in which the elements of the texture grow smaller and smaller as they become more distant. This pattern of change by itself can reveal the spatial layout of the relevant surfaces (Figure 5.30). If these textures also have discontinuities, they can tell us even more about how the surfaces are laid out (Figure 5.31; Gibson, 1950, 1966).

linear perspective A cue for distance based on the fact that parallel lines seem to converge as they get farther away from the viewer.

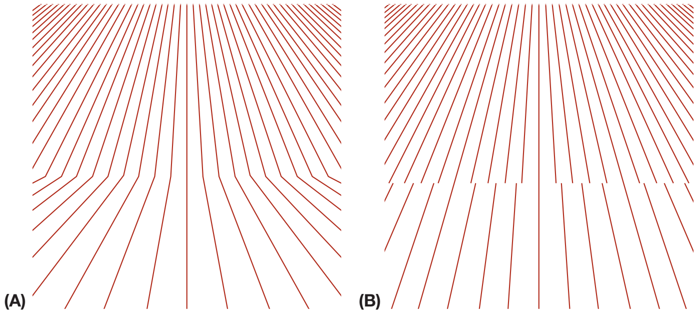
motion parallax A depth cue based on the fact that, as an observer moves, the retinal images of nearby objects move more rapidly than do the retinal images of objects farther away.

The Perception of Depth through Motion

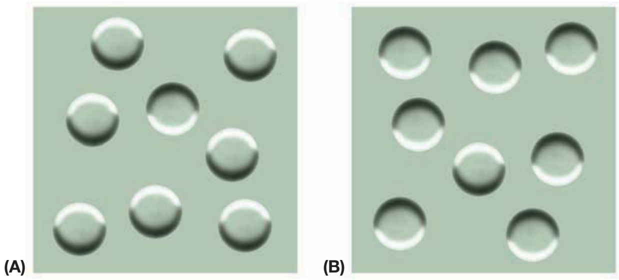
Whenever you move your head, the images projected by the objects in your world necessarily move across your retinas. For reasons of geometry, the projected images of nearby objects move more than those of distant ones; this pattern of motion in the retinal images gives us yet another distance cue, called **motion parallax** (Helmholtz, 1909).

A different motion cue is produced when we move toward or away from objects. As we approach an object, its image gets larger and larger; as we move away, it gets smaller. The pattern of stimulation across the entire visual field also changes as we move toward an object, resulting in a pattern of change in the retinal stimulation that's called *optic flow*. This flow gives us crucial information about depth and plays a large role in the coordination of our movements (Gibson, 1950, 1979).

5.30 Texture gradients as cues for depth Uniformly textured surfaces produce texture gradients that give us information about depth: as the surface recedes, the size of the texture elements decreases, and the density of these elements increases.



5.31 The effect of changes in texture gradients Such changes provide important information about spatial arrangements in the world. Examples are (A) an upward tilt at a corner; and (B) a sudden drop.



5.32 Monocular cues to depth: light and shadow Observers are sensitive to many different depth cues, including depth from shading. (A) Eight circular objects. Most viewers will say the object in the middle looks concave (indented), and the other seven look like they're bulging out. (B) The same figure rotated 180 degrees. Now the middle object looks convex, while the other seven seem concave. The reason is the location of the shadows. When the shadow is at the bottom, the object looks convex; when it's at the top, the object looks concave. This makes sense because light almost always comes from above.

The Role of Redundancy

You might think that the various distance cues all end up providing the same information—each one tells us which objects are close by and which are far. On that basis, it might be efficient for the visual system to focus on just one or two cues and ignore the others. The fact is, however, that we make use of all these cues as well as several others we haven't described (e.g., Figure 5.32).

Why did natural selection favor a system influenced by so many cues, especially since these cues often provide redundant information? It's because different distance cues become important in different circumstances. For example, binocular disparity is a powerful cue, but it's informative only when objects are relatively close by. (For targets farther than 30 feet away, the two eyes receive virtually the same image.) Likewise, motion parallax tells us a great deal about the spatial layout of our world, but only if we're moving. Texture gradients are informative only if there's a suitably uniform texture in view. So while these various cues are often redundant, each type of cue can give us information when the others cannot. By being sensitive to them all, we're able to judge distance in nearly any situation we encounter.

MOTION PERCEPTION: WHAT IS IT DOING?

We obviously want to know what objects are in view and where they're located, but we also want to know what these objects are doing. Are they moving or standing still, approaching slowly or rapidly, racing toward the food we wanted for ourselves, or heading off in some altogether different direction? These questions bring us to a different aspect of perception—namely, how we perceive *motion.*

Retinal Motion

One might think that the perception of motion is extremely simple: If an object in our world moves, then the image cast by that object moves across our retinas. We detect that image motion, and thus we perceive movement.

As we'll soon see, however, this account is way too simplistic. Still, it contains a key element of truth: We do detect an image's motion on the retina, and this is one aspect of the overall process of motion perception. More specifically, some cells in the visual cortex respond to image movements on the retina by firing at an increased rate whenever movement is present. However, these cells don't respond to just any kind of movement, because the cells are *direction specific*. Thus, the cells fire if a stimulus moves across their receptive field from, say, left to right; but not if the stimulus moves from right to left. (Other cells, of course, show the reverse pattern.) These cells are therefore well suited to act as **motion detectors** (see, for example, Vaultin & Berkeley, 1977).

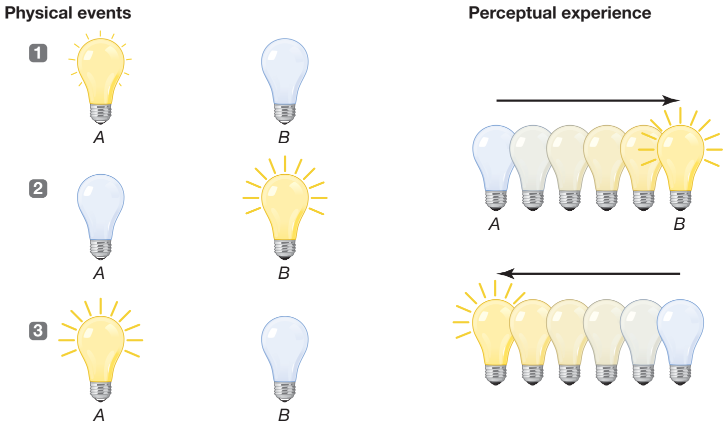
motion detectors Cells in the visual cortex that are sensitive to an image moving in a particular direction across the retina.

apparent movement The perception of movement produced by stimuli that are stationary but are presented first at one positions and then, at an appropriate time interval, presented at a different position.

Apparent Movement

It's clear, however, that retinal motion is only part of the story. Suppose we turn on a light in one location in the visual field, then quickly turn it off, and after an appropriate interval (somewhere between 30 and 200 milliseconds) turn on a second light in a different location. The result is **apparent movement.** The light appears to travel from one point to another, even though there was no motion and, indeed, no stimulation whatsoever in the locations between the two lights (Figure 5.33). This phenomenon is perceptually quite compelling; given the right timing, apparent movement is indistinguishable from real movement (Wertheimer, 1912). This is why the images in movies seem to move, even though movies actually consist of a sequence of appropriately timed still pictures (Figure 5.34).

Apparent movement might seem like an artificial phenomenon because the objects in our world tend to move continuously —they don't blink out of existence *here* and then reappear a moment later *there.* It turns out, however, that, the motion we encounter in the world is often so fast that it's essentially just a blur across the retina, and so triggers no response from the retinal motion detectors. Even so, we do perceive the motion by perceiving the object first to be in one place and then, soon after, to be somewhere else. In this way, the phenomenon of apparent movement actually mirrors a process that we rely on all the time, thanks to the fact that our eyes often need to work with brief "samples" taken from the stream of continuous motion (Adelson & Bergen, 1985).



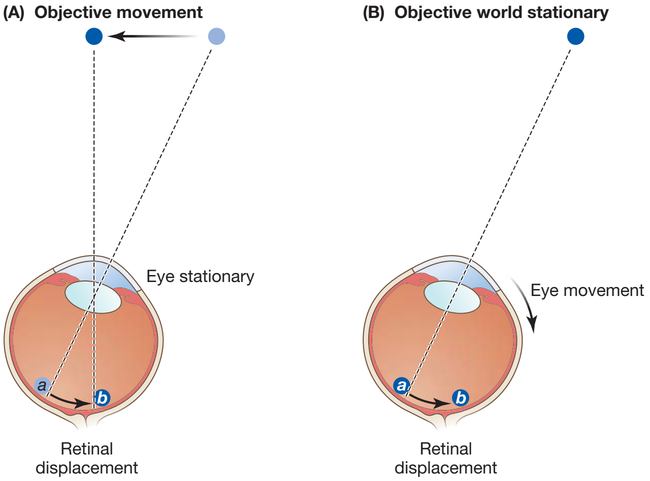
5.33 Apparent movement The sequence of optical events that produces apparent movement. Light A flashes at time 1, followed by light B at time 2, then back to light A at time 3. If the time intervals are appropriately chosen, the viewer will perceive a light moving from left to right and back.

5.34 Apparent movement created by a series of stills A sequence of stills showing a gymnast doing a flip. If the stills are shown in succession, with proper timing, the viewer will perceive smooth movement—even though there's nothing actually moving in the stimulus input.

Eye Movements

As you look around the world, you're constantly moving your head and eyes. This activity creates another complication for motion perception. Each movement brings you a somewhat different view, and so each movement necessarily causes a change in the retinal image. But, despite all this retinal motion, the world doesn't seem to move each time you shift your viewing position. Clearly, it takes more than motion across the retina to produce a perception of motion in the world.

But how do you avoid becoming confused about this retinal motion? How do you manage to separate the retinal motion that's caused by movement in the world from the retinal motion produced by a change in your viewing position? The answer parallels our earlier discussion of constancy. As we've seen, people take *viewing distance* into account when judging *size,* and that's how they achieve size constancy. In the same way, you seem to take *your own movements* into account when judging the *position* of objects in the world, and so you perceive the objects as having *position constancy*. How does this work? Whenever you move your eyes or turn your head, you unconsciously compute the shift in the retinal image that your own motion will produce, and you cancel out this amount of movement in interpreting the visual input (Figure 5.35). The result is constancy.



5.35 Compensation for eye movements In (A), an object has moved from right to left, so its retinal image has shifted from location *a* to location *b*. In (B), there's no motion in the world; instead, the eye has moved from left to right. But here, too, the object's retinal image shifts from location *a* to location *b*. Based only on the retinal information, the displacements in (A) and (B) seem identical. But our brains allow for the displacements caused by changes in eye position. So in (B), the brain would decide that there had been no movement because the motion of the eye was precisely equal (and opposite) to the displacement on the retina.