



Bottom-Up and Top-Down Processing in Perception

 **Demonstration:** Perceiving a Picture

Recognizing Letters and Objects

Template Matching

Interactive Activation Model

 **Method:** Word Superiority Effect


Feature Integration Theory (FIT)

Recognition-by-Components Theory

Test Yourself 3.1

Perceptual Organization: Putting Together an Organized World

The Gestalt Laws of Perceptual Organization

 **Demonstration:** Finding Faces in a Landscape

The Gestalt Laws Provide “Best Guess” Predictions About What Is Out There

Why Computers Have Trouble Perceiving Objects


The Stimulus on the Receptors Is Ambiguous
Objects Need to Be Distinguished From Their Surroundings and From Each Other

Objects Can Be Hidden or Blurred


The Reasons for Changes of Lightness and Darkness Can Be Unclear

How Experience and Knowledge Create “Perceptual Intelligence”

Heuristics for Perceiving

 **Demonstration:** Shape From Shading

Knowledge Helps Us Perceive Words in Conversational Speech

 **Demonstration:** Organizing Strings of Sounds

Neurons Contain Information About the Environment

Something to Consider: Perception Depends on Attention

 **Demonstration:** Change Detection

Test Yourself 3.2

Chapter Summary

Think About It

If You Want to Know More

Key Terms

CogLab: Change Detection; Apparent Motion; Blind Spot; Metacontrast Masking; Muller-Lyer Illusion; Signal Detection; Visual Search; Garner Interference

Some Questions We Will Consider

- Why does something that is so easy, like looking at a scene and seeing what is out there, become so complicated when we look at the mechanisms involved? (57)
- Why is recognizing an object so easy for humans, but so difficult for computers? (80)
- How is our knowledge of the world, which we use for perceiving, stored in the brain? (90)

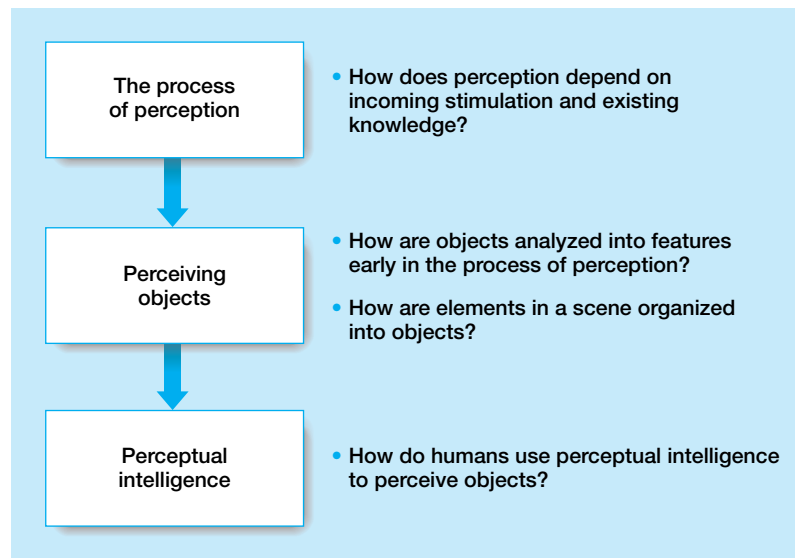
Because of the ease with which we perceive, many people don't see the feats achieved by our senses as complex or amazing. "After all," the skeptic might say, "for vision, a picture of the environment is focused on the back of my eye, and that picture provides all the information my brain needs to duplicate the environment in my consciousness." But the erroneous idea that perception is not that complex is exactly what misled computer scientists in the 1950s and 1960s into proposing that it would take only about a decade or so to create "perceiving machines" that could negotiate the environment with humanlike ease. As it turned out, it took over 50 years to create computer-controlled robots capable of finding their way through the environment, and even these computers fall far short of humans' ability to perceive (Sinha, 2002).

In this chapter, we will explain why **perception** is so complex and why people still outperform computers by a wide margin. We begin by describing how the process of perception depends both on the incoming stimulation and the knowledge we bring to the situation. Following this introduction, we will devote the rest of the chapter to answering the question, "How do we perceive objects?" As we do this, we will see that one reason humans are better at perceiving objects than computers is that humans use perceptual intelligence—knowledge they have gained from their experience in perceiving (Figure 3.1).

One reason we will focus on object perception is that perceiving objects is central to our everyday experience. Consider, for example, what you would say if you were asked to look up and describe what you are perceiving right now. Your answer would, of course, depend on where you are, but it is likely that a large part of your answer would include naming the objects that you see. ("I see a book. There's a chair against the wall. . . .")

We also focus on object perception in this chapter because concentrating on one aspect of perception provides more in-depth understanding of the basic principles of

■ **Figure 3.1**
Flow diagram for
this chapter.



perception than we could achieve by covering a number of different types of perception more superficially. After describing a number of mechanisms of object perception, we will consider “perceptual intelligence”—the idea that the knowledge we bring to a situation plays an important role in perception.

Bottom-Up and Top-Down Processing in Perception

Although perception seems to just “happen,” it is actually the end result of a complex process. We can appreciate the complexity involved in seemingly simple behaviors by returning to our example of Juan and the alarm clock from the beginning of Chapter 2. We saw that one way to describe Juan’s situation was to consider how neurons in his ear and brain respond to the ringing of his alarm. But we also saw that things become more complicated when we consider that Juan’s response to his alarm (hitting the snooze button and going back to sleep) is determined by knowledge that he brings to the situation. His behavior is determined both by the stimulation provided by the ringing alarm clock and his knowledge that he can sleep longer and still get to class on time. We will now consider how behavior is determined both by the energy reaching a person’s receptors and by the knowledge the person brings to a situation.

To illustrate this cooperation between stimulus energy and knowledge, we will consider Ellen, who is taking a walk in the woods. As she walks along the trail she is confronted with a large number of stimuli (Figure 3.2a). When she looks at a particularly distinctive tree off to the right, she doesn’t notice the interesting pattern on the

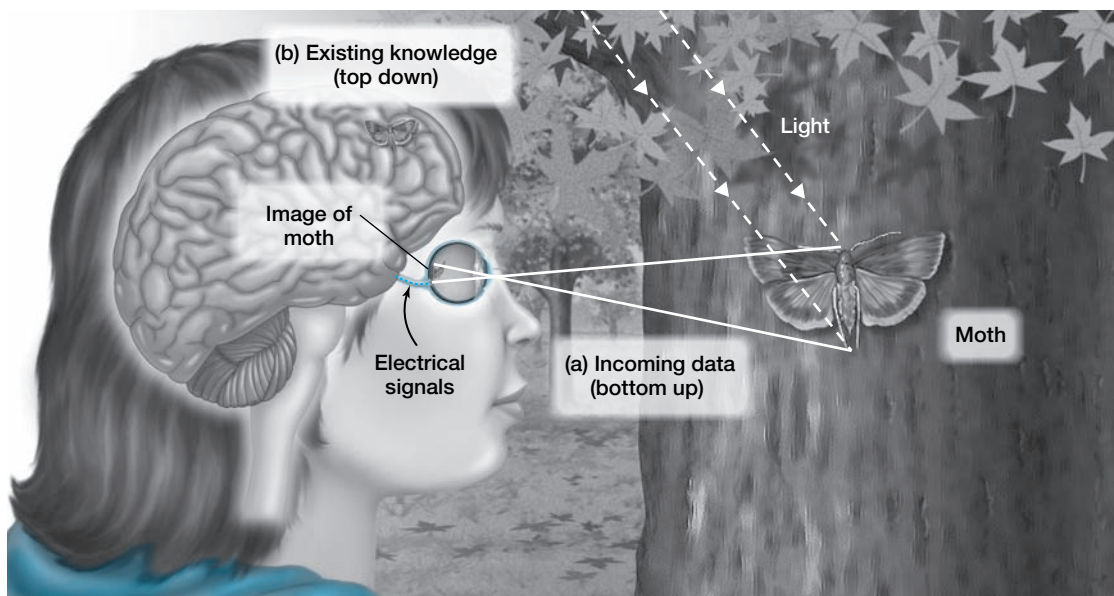


■ **Figure 3.2** (a) Ellen taking a walk in the woods, which contains a large number of stimuli; (b) the moth, which she sees and then recognizes, using a combination of bottom-up and top-down processing.

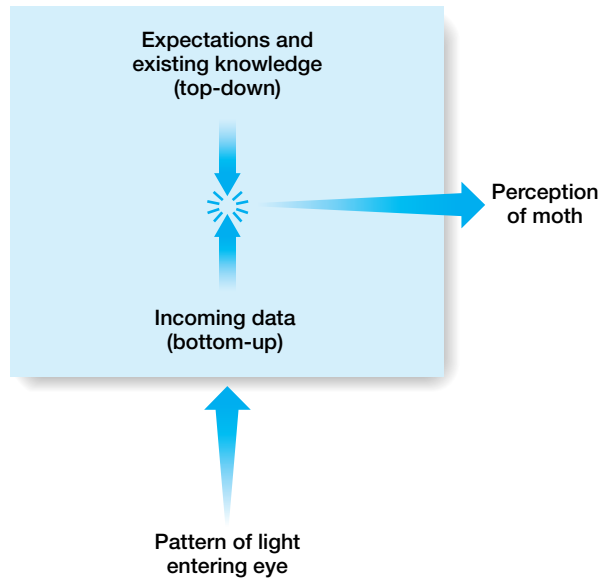
tree trunk at first, but then realizes that what she had at first taken to be a patch of moss was actually a moth (Figure 3.2b).

Let's stop for a moment to consider what has happened. Ellen perceived the moth because light reflected from the moth created an image in her eye (Figure 3.3a). This image triggered the process of transduction we discussed in Chapter 2 (page 31) and resulted in electrical signals, which traveled from the eye to Ellen's brain. This sequence of events, which started with stimulation of the receptors, is called **bottom-up processing**. Bottom-up processing—processing that begins with stimulation of the receptors—is crucial for determining Ellen's experience because if her receptors aren't stimulated, she won't see anything.

But bottom-up processing is not the whole story, because perception involves more than just registering energy on the receptors. We can appreciate this by considering Ellen's problem. Looking at the moth creates a pattern of light and dark on her retina, but it may not be obvious which of the light and dark areas belong to the moth and which belong to the textures of the tree trunk. To help achieve this, Ellen uses her knowledge of moths, not only to detect its presence on the tree, but also to determine that it is a moth, not a butterfly, and to identify what kind of moth it is. Knowledge that Ellen brings to bear on the perceptual problem of seeing and recognizing the moth represents **top-down processing**—processing that involves a person's knowledge (Figure 3.3b). Knowledge doesn't have to be involved in perception but, as we will see, it of-



■ **Figure 3.3** Ellen's perception of the moth is determined by a combination of (a) incoming data (bottom-up information) and (b) existing knowledge (top-down information).



■ **Figure 3.4** Both bottom-up and top-down processing combine to determine perception.

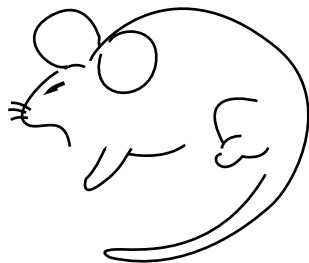
ten is—with bottom-up and top-down processing collaborating to result in perception (Figure 3.4).

In our example, Ellen uses knowledge about moths she had learned much earlier. The following demonstration illustrates that incoming data can be affected by knowledge that has been provided just moments earlier.

Demonstration

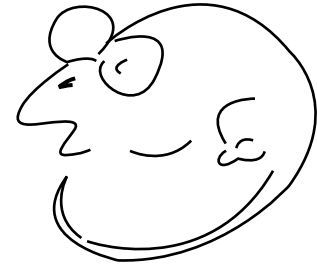
Perceiving a Picture

After looking at the drawing in Figure 3.5, close your eyes, then turn to the next page in the book without looking at the page. Then open and shut your eyes to briefly expose the picture in Figure 3.6 at the top of the page. Decide what the picture is based on this brief exposure. Do this now, before reading further. 🐭



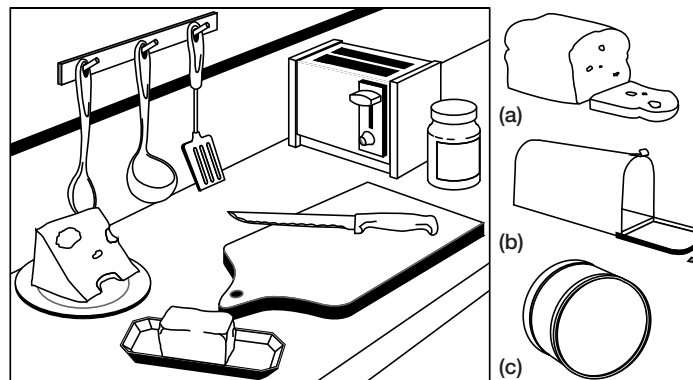
■ **Figure 3.5** Picture for “perceiving a picture” demonstration. (Adapted from “The Role of Frequency in Developing Perceptual Sets,” by B. R. Bugelski and D. A. Alampay, 1961, *Canadian Journal of Psychology*, 15, pp. 205–211, Copyright © 1961 by the Canadian Psychological Association.)

■ **Figure 3.6** (Adapted from “The Role of Frequency in Developing Perceptual Sets,” by B. R. Bugelski et al., 1961, *Canadian Journal of Psychology*, 15, pp. 205–211, Copyright © 1961 by the Canadian Psychological Association.)



What did you see when you looked at Figure 3.6 above? Did it look like a rat (or a mouse)? If it did, you were influenced by the clearly rat- or mouselike figure you saw in Figure 3.5. But people who first observe Figure 3.10 (on page 63) usually identify Figure 3.6 as a man. (Try this demonstration on someone else.) This demonstration, which is called the **rat-man demonstration**, shows how recently acquired knowledge (“that pattern is a rat”) can influence perception.

Another example of an effect of top-down processing is provided by an experiment by Stephen Palmer (1975), in which he presented a context scene such as the one on the left of Figure 3.7 and then briefly flashed one of the target pictures on the right. One of the targets was appropriate to the scene (the loaf of bread), one was inappropriate (the drum), and one was misleading (the mailbox, which was shaped like the loaf of bread). When the participants reported what the target picture was, they were correct 83 percent of the time for the appropriate object, 50 percent for the inappropriate object, and



■ **Figure 3.7** Stimuli like those used in Palmer's (1975) experiment, which showed how context can influence perception. (Reprinted from “The Effects of Contextual Scenes on the Identification of Objects,” by S. E. Palmer, 1975, *Memory and Cognition*, 3, pp. 519–526, Copyright © 1975 with permission of the author and the Psychodynamic Society Publishers.)

40 percent for the misleading object. This experiment shows how a person's knowledge of the context provided by a particular scene can influence perception.

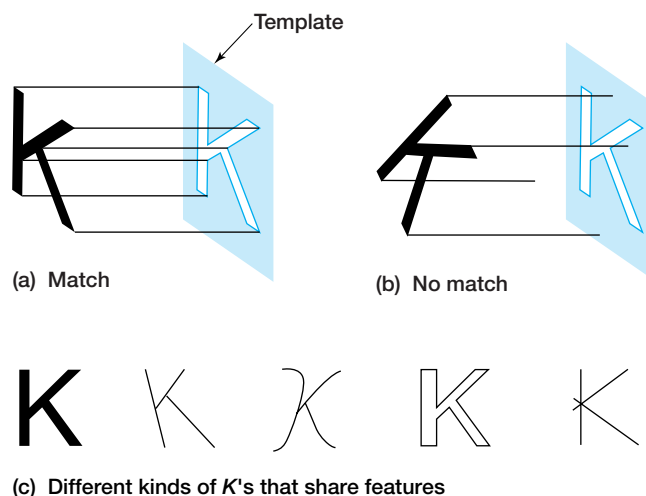
As you will see in later chapters, there are numerous situations in which incoming data interacts with a person's knowledge. This occurs for attention, memory, language, and most of the other types of cognition we will be discussing. In this chapter, we will focus on perception by looking at what cognitive psychologists have discovered about how both bottom-up and top-down processes operate as we perceive objects. We start by describing how incoming stimuli are analyzed by the visual system. This analysis occurs rapidly and without our awareness and provides an example of how bottom-up and top-down processing can interact.

Recognizing Letters and Objects

As a first step in determining how we perceive objects, we will follow the lead of early cognitive psychologists, who focused on the simple case of perceiving letters of the alphabet. We begin with an idea called **template matching**, which turned out to be too simple to explain how we perceive letters, but which led to the idea of perception based on features, which is part of present-day explanations of object perception.

Template Matching

We begin with a simple example—how we recognize the letter *K* in Figure 3.8. One way the perceptual system could achieve this would be to compare the pattern *K* to a model or template of the letter *K* that is stored in the system. According to this idea, when



■ **Figure 3.8** According to the idea of template matching, we can identify an object when it matches a template. Thus, in (a), in which the stimulus matches the template, the perceiver identifies it as a *K*. A problem arises, however, when the stimulus is tilted, as in (b), because then it no longer matches the template, and so the perceiver would not be able to identify it. (c) Each of these *K*'s would require different templates, but because they share features, they can be identified by a mechanism that takes these features into account.

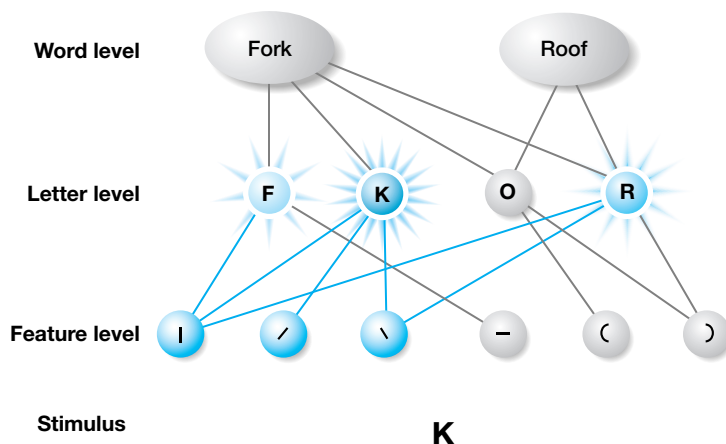
the pattern matches the template, the perceiver recognizes the letter as a *K*. But this idea runs into problems when we consider what happens when the *K* is tilted, as in Figure 3.8b. Tilting the *K* poses no problem for a perceiver, who can still recognize it. However, template-matching theory would require a template for every orientation of the *K*. People also have no trouble identifying different forms of the same letter, like the *K*'s in Figure 3.8c. It is apparent that the template-matching model won't work, because a huge number of different templates would be needed just to recognize one letter. When we multiply this by how many objects there are in the environment, the number becomes astronomical. To deal with this problem, psychologists developed models of letter perception based on the idea that letters can be broken down into features.

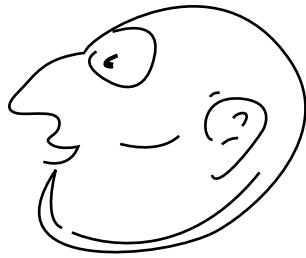
Interactive Activation Model

We saw in Chapter 2 that there are cortical neurons called feature detectors that respond to oriented lines (Hubel & Wiesel, 1965). The discovery of feature detectors in the 1960s suggested that perhaps the perceptual system constructs letters and other objects in the environment from simple features, like oriented lines. Features help solve some of the problems associated with template matching, because although letters like the ones in Figure 3.8c look different, they all have features in common, such as vertical and slanted lines.

This idea led James McClelland and David Rumelhart (1981; also Rumelhart & McClelland, 1982) to propose the model of letter recognition shown in Figure 3.9. This model, which is called the **interactive activation model**, proposes that activation is sent through three levels: The **feature level** contains **feature units**—mainly straight and curved lines; the **letter level** contains **letter units**—one for each letter in the alphabet; and the **word level** contains **word units**—all the words a person knows. The simplified model in Figure 3.9 contains 6 feature units and 4 letter units. The complete model has 12 feature units and 26 letter units. We will use our simplified model to dem-

■ **Figure 3.9** Diagram of McClelland and Rumelhart's (1981) interactive activation model of word recognition. This diagram indicates that feature units at the feature level are activated by the letter *K*, and that these feature units send activation to letter units in the letter level. Color and radiating lines indicate activation.



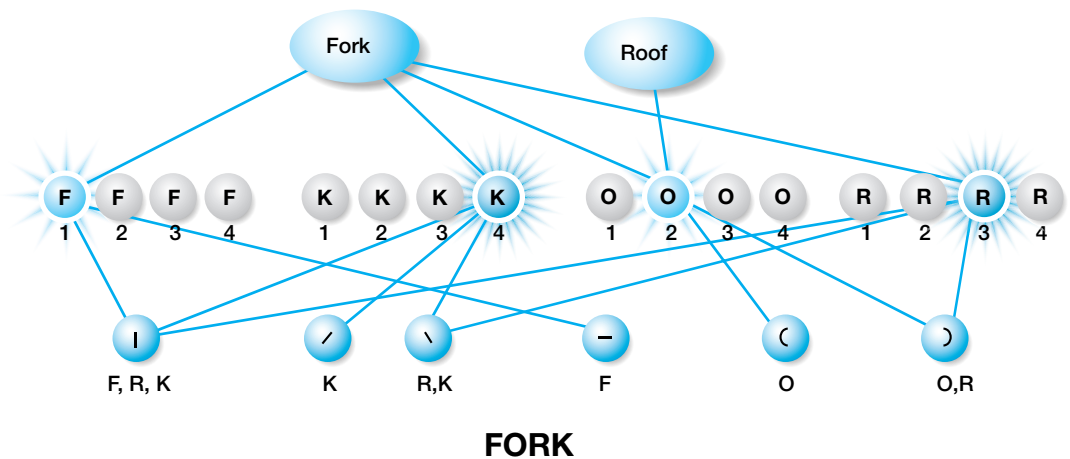


■ **Figure 3.10** The “man” stimulus for the rat–man demonstration. (Adapted from “The Role of Frequency in Developing Perceptual Sets,” by B. R. Bugelski et al., 1961, *Canadian Journal of Psychology*, 15, pp. 205–211, Copyright © 1961 by the Canadian Psychological Association.)

onstrate how interactive activation handles the following three situations: (1) recognizing a single letter; (2) recognizing a single word; and (3) recognizing a letter within a word.

Recognizing a Single Letter Presenting the letter *K* activates feature units for *K*’s features—a straight line and two slanted lines (Figure 3.9). These feature units then send activation to each letter unit that contains these features—the *F*, *K*, *O*, and *R* in our example (in the full model, with all 26 letters, other letters would also be activated). According to the interactive activation model, the letter unit that is activated the most indicates which letter was presented. In our example, the *K* is activated the most, indicating that the *K* was, in fact, the letter that was presented.

Recognizing a Word We now consider how the model responds to the word *FORK*. In Figure 3.11 we have added a characteristic to the model that enables it to deal with words. Now, in the letter stage there is a letter unit for each letter’s *position* in a word. For



■ **Figure 3.11** How the word *FORK* activates the components of the interactive activation model. See text for explanation.

example, presenting *FORK* activates the features for the *F*, and each of these features send activation to letter unit F1—which is for *F* in the first position in a word. Similarly, the features for *O* activate the O2 letter unit (the *O* in the second position), *R*'s features activate the R3 unit, and *K*'s features activate the K4 unit.

These letter units then send activation to all words that contain letters in the correct positions. In our example, the word *FORK* receives signals from the F1, O2, R3, and K4 letter units. Notice that the word *ROOF* receives signals only from the O2 letter unit. In this word, the *R* and the *F* are in the wrong position to receive activation from the *R* and *F* letter units that are activated by *FORK*. Because *FORK* is more highly activated than *ROOF*, the model recognizes *FORK* as the word that was presented. Of course, in the full model, many more words would be involved, but the general result is that the word that is presented causes the most activation.

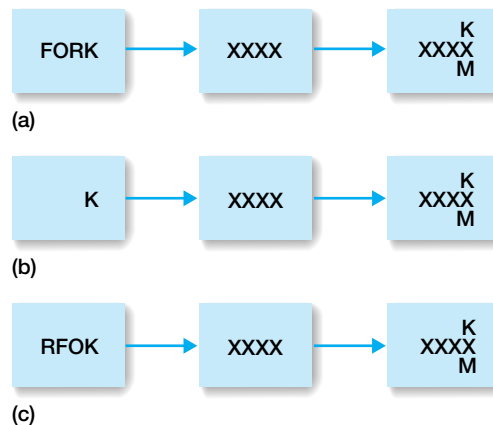
The Word Superiority Effect Next we consider how the model deals with recognizing a letter that is contained in a word, but first we will describe the **word superiority effect**—letters are easier to recognize when they are contained in a word, compared to when they appear alone or are contained in a nonword. This effect was first demonstrated by G. M. Reicher in 1969 using the following procedure.

Method

Word Superiority Effect

A stimulus that is either (a) a word, like *FORK*; (b) a single letter, like *K*; or (c) a nonword, such as *RFOK*, is flashed briefly and is followed immediately by a masking stimulus, indicated in Figure 3.12 by XXXX, that stops further processing of the original stimulus. Following the mask, two letters are briefly presented, one that appeared in the original stimulus, and another that did not. The participants' task is to pick the letter that was presented in the original stimulus. In the example in Figure 3.12a, the word *FORK* was presented, so *K* would

■ **Figure 3.12** Procedure for experiment that demonstrates the word-superiority effect. First the stimulus is presented, then the XXXX's, then the letters. Three types of stimuli are shown: (a) word condition; (b) letter condition; and (c) nonword condition.



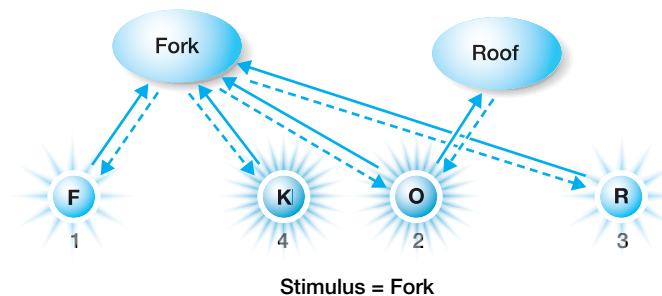
be the correct answer. *K* would also be the correct answer if the *K* were originally presented alone (Figure 3.12b), or if it were presented in a nonword like *RFOK* (Figure 3.12c). 🟡

When Reicher's participants were asked to choose which of the two letters they saw in the original stimulus, they did so more quickly and accurately when the letter was part of the original word, as in Figure 3.12a, than when the letter was presented alone, as in Figure 3.12b, or was part of a nonword, as in Figure 3.12c. This more rapid processing of letters when in a word—the word superiority effect—means that letters in words are not processed letter by letter but that each letter is affected by its surroundings. With this experimental finding in hand, let's consider how the interactive activation model would explain the recognition of a letter within a word.

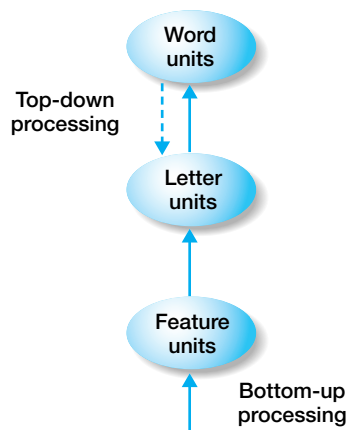
Recognizing a Letter Within a Word Figure 3.13 shows the letter level and word level from Figure 3.11, but with one added feature—**feedback activation**, indicated by the dashed arrows that extend from the word units back to the letter units. Feedback activation is activation that is sent from word units back to each of the letter units for that word. For example, the unit for *FORK* sends activation back to the K4 letter unit. This enhances the activation of the K4 unit.

The enhanced activity of the letter units caused by feedback activation explains the word superiority effect, because feedback activation does not occur when a letter is presented alone (note that the activation for K4 is greater than the activation for the *K* in Figure 3.9). Notice that some feedback activation would occur when a nonword such as *RFOK* is presented (because the K4 letter unit is activated and sends its activation to the *FORK* unit), but much less than for when *FORK* is presented. Thus, the letter *K* and each of the other letters in *FORK* are more highly activated when they appear in the word than when they appear alone or in a nonword.

The model in Figure 3.11 is important for a number of reasons. First, it proposes a mechanism that is consistent with what we know about neural firing. Excitation is sent from one level to another in the model, just as excitation is sent from one neuron to another in the nervous system. The model also contains another characteristic that corresponds to neural firing. It proposes a role for inhibition, which is sent between the letter units and between the word units. We didn't include inhibition in our example



■ **Figure 3.13** The letter and word levels of the interactive activation model, showing how feedback activation from the word level to the letter level (dashed lines) increases activation of the letter units.



■ **Figure 3.14** Summary of how activation flows in the interactive activation model. Activation flowing from the feature units toward the word units represents bottom-up processing. Activation flowing from the word units to the letter units represents top-down processing.

above, but the net effect of inhibition is to enhance the activation of units corresponding to stimulus letters or words, compared to units that do not correspond to other letters or words.

The model is also important because it takes top-down processing into account. Remember that bottom-up processing is initiated by stimulation of the receptors, and top-down processing occurs when a person's knowledge affects processing. Thus, in this model, bottom-up processing occurs when letter or word stimuli activate the receptors, which then activate feature units. Top-down processing is also involved because the existence of word units is based on the person's knowledge of which strings of letters form words, and the feedback activation that is sent back from the words to the letter units reflects top-down processing (Figure 3.14).

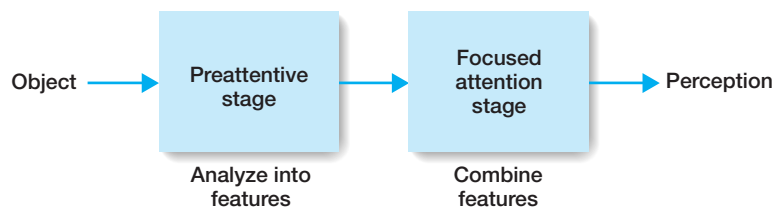
This is an early version of a type of model called a *connectionist model*. Connectionist models involve networks that look like the ones in Figures 3.11 and 3.13. As we will see in Chapter 8, networks like this have been used to explain not only how we recognize letters and words, but how we learn to recognize stimuli we have never experienced before.

Considering how letters are recognized provides a good way to show how bottom-up and top-down processing interact with one another. But we are interested not just in how we recognize letters, but in how we recognize other types of objects as well. This step in our story takes us to Anne Treisman's *feature integration theory* of perception.

Feature Integration Theory (FIT)

Figure 3.15 shows the basic idea behind **feature integration theory (FIT)** (Treisman, 1986). According to this theory, the first stage of perception is the **preattentive stage**, so named because it happens automatically and doesn't require any effort or attention by the perceiver. In this stage, an object is analyzed into its features.

The idea that an object is automatically broken into features may seem counterintuitive because when we look at an object, we see the whole object, not an object that



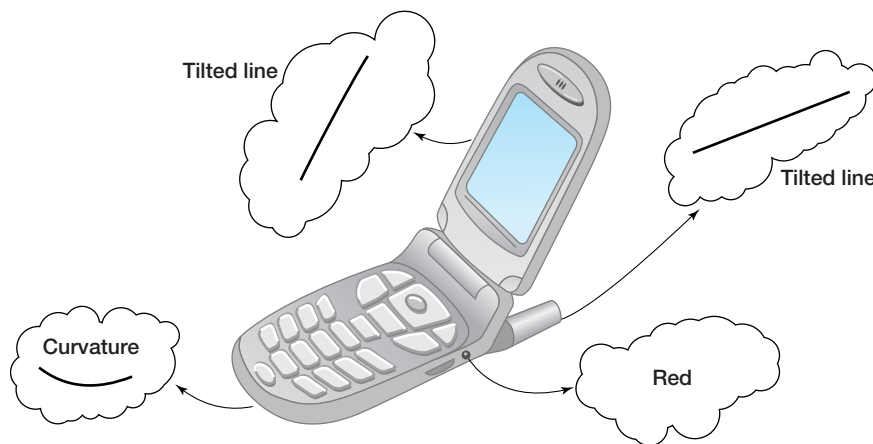
■ **Figure 3.15** Flow diagram for Treisman's (1986) feature integration theory (FIT). According to this theory, objects are first analyzed into features in the preattentive stage, and then these features are combined into an object that can be perceived in the focused attention stage.

has been divided into its individual features. The reason we aren't aware of this process of feature analysis is that it occurs early in the perceptual process, before we have become conscious of the object. Thus, when you see this book, you are conscious of its rectangular shape, but you are not aware that before you saw this rectangular shape, your perceptual system analyzed the book into individual features such as lines with different orientations.

To provide some perceptual evidence that objects are, in fact, analyzed into features, Treisman and H. Schmidt (1982) did an ingenious experiment to show that early in the perceptual process, features may exist independently of one another. Treisman and Schmidt's display consisted of four objects flanked by two black numbers (● Color Plate 3.1). They flashed this display onto a screen for one-fifth of a second, followed by a random-dot masking field designed to eliminate any residual perception that might remain after the stimuli were turned off. Participants were told to report the black numbers first and then to report what they saw at each of the four locations where the shapes had been.

In 18 percent of the trials, participants reported seeing objects that were made up of a combination of features from two different stimuli. For example, after being presented with the display in ● Color Plate 3.1, in which the small triangle was red and the small circle was green, they might report seeing a small red circle and a small green triangle. These combinations of features from different stimuli are called **illusory conjunctions**. Illusory conjunctions can occur even if the stimuli differ greatly in shape and size. For example, a small blue circle and a large green square might be seen as a large blue square and a small green circle.

According to Treisman, these illusory conjunctions occur because at the beginning of the perceptual process each feature exists independently of the others. That is, features such as “redness,” “curvature,” or “tilted line” are, at this early stage of processing, not associated with a specific object (Figure 3.16). They are, in Treisman's (1986) words,



■ **Figure 3.16** The results of the illusory conjunction experiment suggest that very early in the perceptual process, features that make up an object are “free floating.” This is symbolized here by showing some of the features of a cell phone as existing separately from one another at the beginning of the perceptual process.

“free floating” and can therefore be incorrectly combined in laboratory situations when briefly flashed stimuli are followed by a masking field.

You can think about these features as components of a visual “alphabet.” At the very beginning of the process, perceptions of each of these components exist independently of one another, just as the individual letter tiles in a game of Scrabble exist as individual units when the tiles are scattered at the beginning of the game. However, just as the individual Scrabble tiles are combined to form words, the individual features combine to form perceptions of whole objects. According to Treisman’s model, these features are combined in the second stage, which is called the **focused attention stage**. Once the features have been combined in this stage, we perceive the object.

During the focused attention stage, the observer’s attention plays an important role in combining the features to create the perception of whole objects. To illustrate the importance of attention for combining the features, Treisman repeated the illusory conjunction experiment using the stimuli in ● Color Plate 3.1, but she instructed her participants to ignore the black numbers and to focus all of their attention on the four target items. This focusing of attention eliminated illusory conjunctions so that all of the shapes were paired with their correct colors.

When I describe this process in class, some students aren’t convinced. One student said, “I think that when people look at an object, they don’t break it into parts. They just see what they see.” To convince this student (and the many others who, at the beginning of the course, are still not comfortable with the idea that cognition sometimes involves rapid processes we aren’t aware of), I describe the case of R.M., a patient who had parietal lobe damage that resulted in a condition called **Balint’s syndrome**. The crucial characteristic of Balint’s syndrome is an inability to focus attention on individual objects.

According to feature integration theory, lack of focused attention would make it difficult for R.M. to combine features correctly, and this is exactly what happened. When R.M. was presented with two different letters of different colors, such as a red *T* and a blue *O*, he reported illusory conjunctions such as “blue *T*” on 23 percent of the trials, even when he was able to view the letters for as long as 10 seconds (Friedman-Hill et al., 1995; Robertson et al., 1997). The case of R.M. illustrates how a breakdown in the brain can reveal processes that are not obvious when the brain is functioning normally.

The feature analysis approach involves mostly bottom-up processing because knowledge is usually not involved. In some situations, however, top-down processing can come into play. For example, when Treisman did an illusory conjunction experiment using stimuli such as the ones in ● Color Plate 3.2 and asked participants to identify the objects, the usual illusory conjunctions occurred, so the orange triangle would, for example, sometimes be perceived to be black. However, when she told participants that they were being shown a carrot, a lake, and a tire, illusory conjunctions were less likely to occur, so participants were more likely to perceive the triangular “carrot” as being orange. Thus, in this situation, the participants’ knowledge of the usual colors of objects influenced their ability to correctly combine the features of each object. In our every-

day experience, in which we are often perceiving familiar objects, top-down processing combines with feature analysis to help us perceive things accurately.

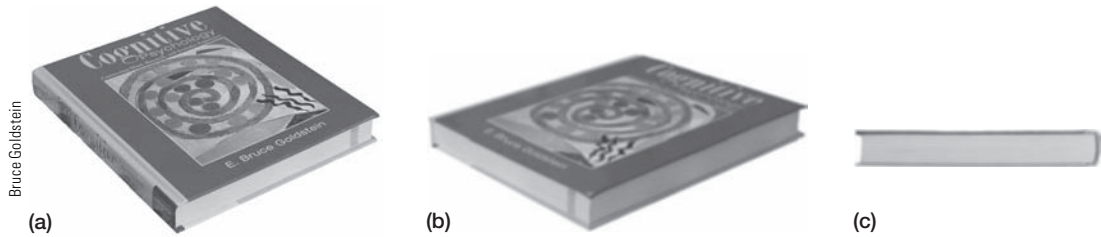
The features in Treisman's model are things like lines, curves, and colors. But these types of features don't explain how we perceive the three-dimensional objects we routinely encounter in our environment. Another feature-based theory, called *recognition-by-components theory*, proposes three-dimensional features to deal with this situation.

Recognition-by-Components Theory

In the **recognition-by-components (RBC) theory** of perception, the features are not lines, curves, or colors, but are three-dimensional volumes called **geons**. Figure 3.17a shows a number of geons, which are shapes such as cylinders, rectangular solids, and pyramids. Irving Biederman (1987), who developed the recognition-by-components theory, has proposed that there are 36 different geons, which is enough to construct a large proportion of the objects that exist in the environment. Figure 3.17b shows a few objects that have been constructed from geons.

An important property of geons is that they can be identified when viewed from different angles. This property, which is called **view invariance**, occurs because geons contain **view invariant properties**—properties such as the three parallel edges of the rectangular solid in Figure 3.17 that remain visible even when the geon is viewed from many different angles.

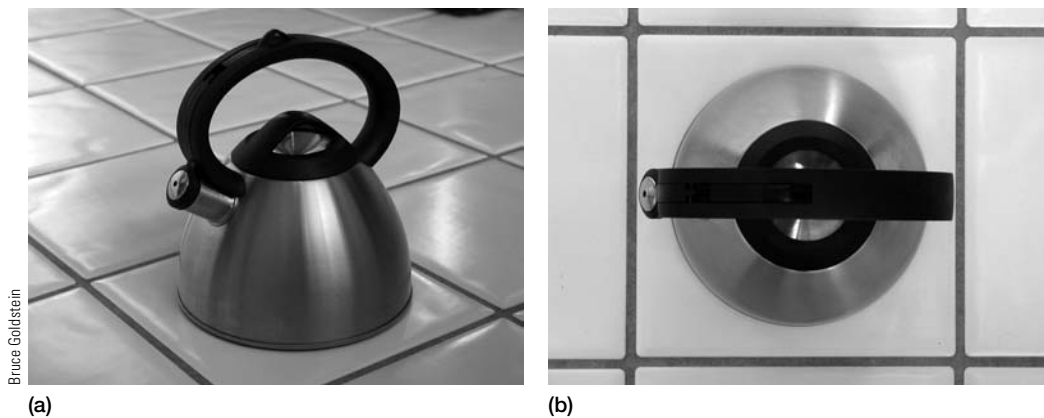
Text not available due to copyright restrictions



■ **Figure 3.18** A view-invariant property of a rectangular object is demonstrated by fact that three parallel edges are present even when we change our viewpoint of the book, as in (a) and (b). In rare cases, such as (c), when the book is viewed from end-on, this invariant property is not perceived.

You can test the view-invariant properties of a rectangular solid yourself by picking up a book and moving it around, so you are looking at it from many different viewpoints. As you do this, notice what percentage of the time you are seeing the three parallel edges. Also notice that occasionally, as when you look at the book end-on, you do not see all three edges (Figure 3.18c). However, these situations occur only rarely, and when they do occur, it becomes more difficult to recognize the object. For example, when we view the object in Figure 3.19a from the rarely encountered unusual perspective in Figure 3.19b, we see fewer basic geons and therefore have difficulty identifying it.

Two other properties of geons are *discriminability* and *resistance to visual noise*. **Discriminability** means that each geon can be distinguished from the others from almost all viewpoints. **Resistance to visual noise** means we can still perceive geons under “noisy” conditions such as might occur under conditions of low light or fog. For example, look at Figure 3.20. The reason you can identify this object (what is it?)—even though over half of its contour is obscured—is because you can still identify its geons.



■ **Figure 3.19** (a) A familiar object; (b) the same object seen from a viewpoint that obscures most of its geons. This makes it harder to recognize the object.



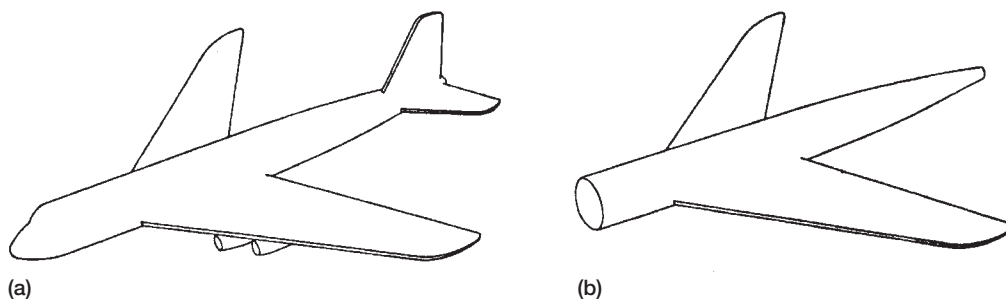
■ **Figure 3.20** What is the object behind the mask? (Adapted from “Recognition-by-Components: A Theory of Human Image Understanding,” by I. Biederman, 1987, *Psychological Review*, 24, 2, pp. 115–147, Figure 26, Copyright © 1987 with permission from the author and the American Psychological Association.)



■ **Figure 3.21** The same object as in Figure 3.20 (a flashlight) with the geons obscured. (Adapted from “Recognition-by-Components: A Theory of Human Image Understanding,” by I. Biederman, 1987, *Psychological Review*, 24, 2, pp. 115–147, Figure 25, Copyright © 1987 with permission from the author and the American Psychological Association.)

However, in Figure 3.21, in which the visual noise is arranged so the geons cannot be identified, it becomes impossible to recognize that the object is a flashlight.

The basic message of recognition-by-components theory is that if enough information is available to enable us to identify an object’s basic geons, we will be able to identify the object (also see Biederman, 2001; Biederman & Cooper, 1991; Biederman et al., 1993). A strength of Biederman’s theory is that it shows that we can recognize objects based on a relatively small number of basic shapes. For example, we easily recognize Figure 3.22a, which has nine geons, as an airplane, but even when only three geons are present, as in Figure 3.22b, we can still identify an airplane.



■ **Figure 3.22** An airplane, as represented by (a) nine geons; (b) three geons. (Adapted from “Recognition-by-Components: A Theory of Human Image Understanding,” by I. Biederman, 1987, *Psychological Review*, 24, 2, pp. 115–147, Figure 13, Copyright © 1987 with permission from the author and the American Psychological Association.)

Both feature integration theory and recognition-by-components theory are based on the idea of early analysis of objects into parts. These two theories explain different facets of object perception. Feature integration theory is more concerned with very basic features like lines, curves, colors, and with how attention is involved in combining them, whereas recognition-by-components theory is more about how we perceive three-dimensional shapes. Thus, both theories explain how objects are analyzed into parts early in the perceptual process.

There is, however, more to perceiving objects than analyzing them into parts. We will now consider another aspect of object perception, which focuses not on analysis that occurs early in the perceptual process, but on how we organize elements of the environment into separate objects.



Test Yourself 3.1

1. Describe the role of bottom-up and top-down processing as applied to Ellen seeing the moth on the tree, to the rat-man demonstration, and to Palmer's kitchen experiment.
2. What is the basic idea behind the feature analysis approach to perception? Describe the integrative activation model for recognizing letters. How do parts of this model relate to what we know about physiology? How do the word units help explain the word superiority effect?
3. Describe Treisman's feature integration theory. How do her experiments on illusory conjunctions support the idea that features are "free floating" in the preattentive stage? What is the focused attention stage, and what is the evidence that attention is important for combining the features?
4. Describe Biederman's recognition-by-components theory. How is it similar to Treisman's theory, and how is it different?

Perceptual Organization: Putting Together an Organized World

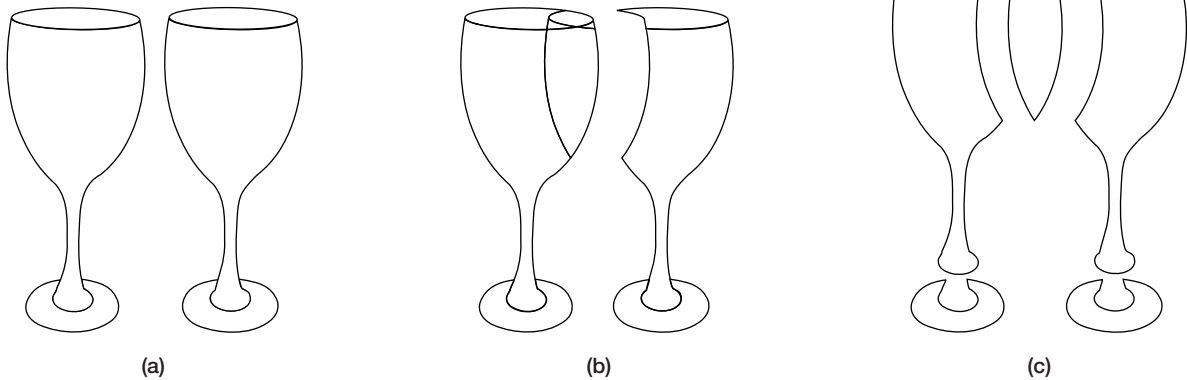
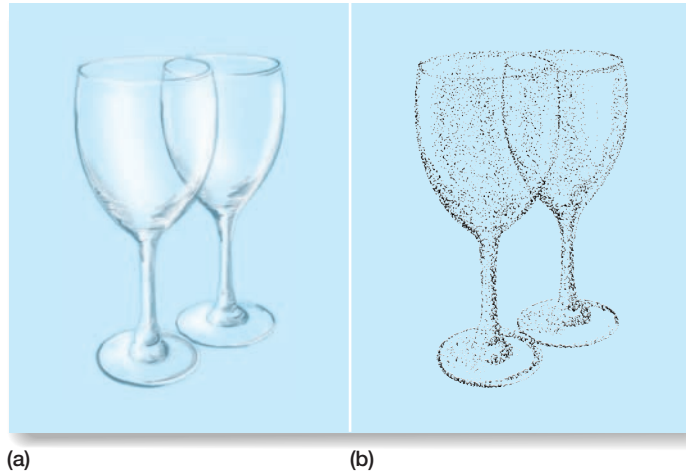
What do you see in Figure 3.23? Take a moment and decide before reading further.

If you have never seen this picture before, you may just see a bunch of black splotches on a white background. However, if you look closely you can see that the picture is a Dalmatian facing to the left, with its nose to the ground. Once you have seen the Dalmatian, it is hard to not to see it. Your mind has achieved **perceptual organization**—the organization of elements of the environment into objects—and has perceptually organized the black areas into a Dalmatian. But what is behind this process? The first psychologists to study this question were the **Gestalt psychologists**, who were active in Europe beginning in the 1920s.

Image not available due to copyright restrictions

In Chapter 1, we described how, early in the 1900s, perception was explained by an approach called **structuralism**, which involved adding up small, elementary units called **sensations**. According to this idea, we see the two glasses in Figure 3.24a because hundreds of tiny sensations, indicated by the dots in Figure 3.24b, add up to create our perception of the glasses. But the Gestalt psychologists took a different approach. Instead of looking at the glasses as a collection of tiny sensations, they considered the overall pattern created by the glasses. According to the Gestalt approach, the pattern in Figure 3.24a can potentially be perceived as representing a number of different objects, as shown in Figures 3.25a, b, and c. But even though many different objects could have created the pattern in Figure 3.24a, the fact that we automatically see the picture as two separate glasses, as in Figure 3.25a, caused the Gestalt psychologists to ask what causes us to organize our perception in this way. They answered this question by proposing that the mind groups patterns according to rules that they called the *laws of perceptual organization*.

■ **Figure 3.24** (a) Two overlapping wine glasses; (b) each dot represents a sensation. According to the structuralist approach, these individual sensations are combined to result in our perception of the glasses.

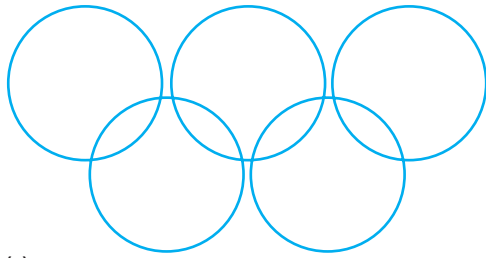


■ **Figure 3.25** Each of the objects in (a), (b), and (c) could have resulted in the perception in Figure 3.24a if arranged appropriately in relation to one another. The Gestalt psychologists pointed out that we see the pattern as two glasses, as in (a), and proposed “laws of perceptual organization” to explain why certain perceptions are more likely than others.

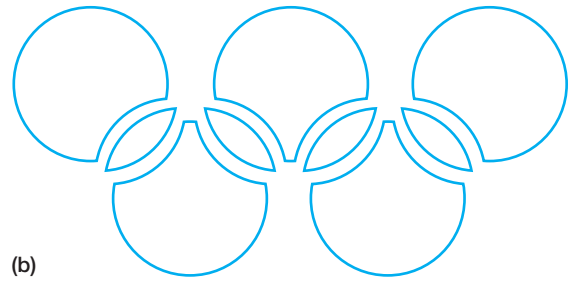
The Gestalt Laws of Perceptual Organization

The **laws of perceptual organization** are a series of rules that specify how we perceptually organize parts into wholes. Let’s look at six of the Gestalt laws.

Pragnanz *Pragnanz*, roughly translated from the German, means “good figure.” The **law of Pragnanz**, the central law of Gestalt psychology, which is also called the **law of good figure** or the **law of simplicity**, states: *Every stimulus pattern is seen in such a*



(a)

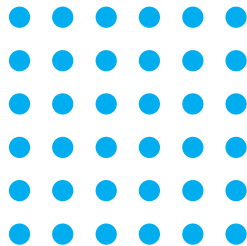


(b)

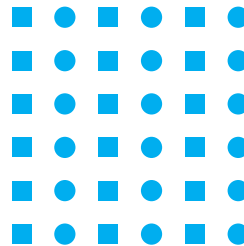
■ **Figure 3.26** Law of simplicity. We see five circles, as in (a), not the more complex array of nine objects, as in (b).

way that the resulting structure is as simple as possible. The familiar Olympic symbol in Figure 3.26a is an example of the law of simplicity at work. We see this display as five circles and not as other, more complicated shapes such as the ones in Figure 3.26b. We can also apply this law to the wine glasses in Figure 3.25. Seeing the pattern as two glasses as in Figure 3.25a is much simpler than seeing it as the more complex objects in Figures 3.25b and c.

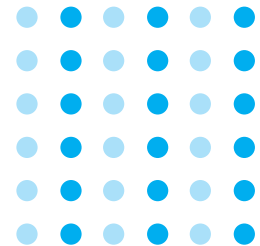
Similarity Most people perceive Figure 3.27a as either horizontal rows of circles, vertical columns of circles, or both. But when we change some of the circles to squares, as in Figure 3.27b, most people perceive vertical columns of squares and circles. This perception illustrates the **law of similarity**: *Similar things appear to be grouped together*. This law causes the circles to be grouped with other circles and the squares to be grouped with other squares. Grouping can also occur because of similarity of lightness (Figure 3.27c), hue, size, or orientation.



(a)



(b)

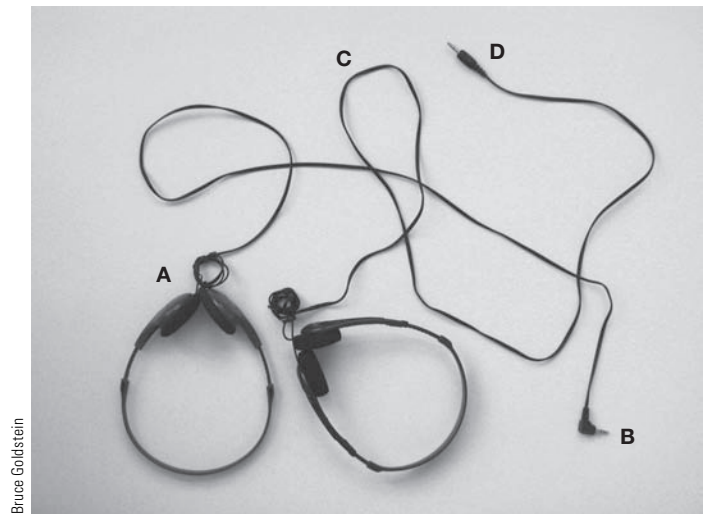


(c)

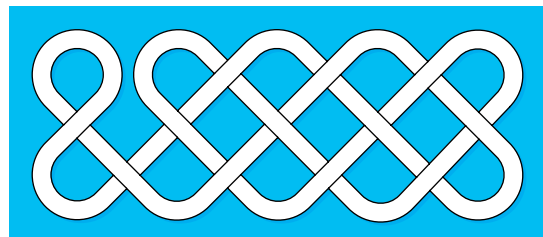
■ **Figure 3.27** Law of similarity. (a) This display can be perceived as either vertical columns or horizontal rows; (b) this is more likely perceived as columns of squares alternating with columns of circles, due to similarity of shape; (c) this is perceived as columns because of similarity of lightness.

Good Continuation We see wire starting at A in Figure 3.28 as flowing smoothly to B. It does not go to C or D because that path would involve making sharp turns and would violate the **law of good continuation**: *Points that, when connected, result in straight or smoothly curving lines, are seen as belonging together, and the lines tend to be seen as following the smoothest path.* Another effect of good continuation is shown in the Celtic knot pattern in Figure 3.29. In this case, good continuation assures that we see a continuous interweaved pattern that does not appear to be broken into little pieces every time one strand overlaps another strand. Good continuation also helped us to perceive the smoothly curving Olympic circles in Figure 3.26.

■ **Figure 3.28** Good continuation helps us perceive two separate wires, even though they overlap.

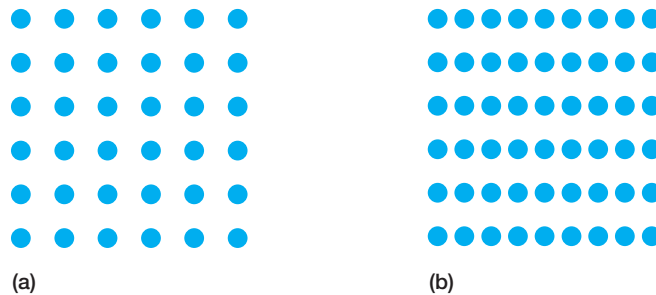


■ **Figure 3.29** We perceive this pattern as continuous interwoven strands because of good continuation.



Proximity or Nearness Figure 3.30a is the pattern from Figure 3.27a that can be seen as either horizontal rows or vertical columns or both. By moving the circles closer together, as in Figure 3.30b, we increase the likelihood that the circles will be seen in horizontal rows. This illustrates the **law of proximity** or **law of nearness**: *Things that are near to each other appear to be grouped together.*

Common Fate The **law of common fate** states: *Things that are moving in the same direction appear to be grouped together.* Thus, when you see a flock of hundreds of birds all flying together, you tend to see the flock as a unit, and if some birds start flying in another direction, this creates a new unit (Figure 3.31).



■ **Figure 3.30** Grouping by nearness. The pattern in (a) is perceived as vertical columns or horizontal rows, but when the dots are near each other, as in (b), the perception changes to horizontal rows.



■ **Figure 3.31** A flock of birds that are moving in the same direction are seen as grouped together. When a portion of the flock changes direction, their movement creates a new group. This illustrates the law of common fate.

Familiarity According to the **law of familiarity**, *things are more likely to form groups if the groups appear familiar or meaningful* (Helson, 1933; Hochberg, 1971). You can appreciate how meaningfulness determines perceptual organization by doing the following demonstration.



Demonstration

Finding Faces in a Landscape

Consider the picture in ● Color Plate 3.3. At first glance this scene appears to contain mainly trees, rocks, and water. But on closer inspection you can see some faces in the trees in the background, and if you look more closely, you can see that a number of faces are formed by various groups of rocks. See if you can find all 12 faces that are hidden in this picture. 🌊

In this demonstration some people find it difficult to perceive the faces at first, but then suddenly they succeed. (Hint: The group of rocks at the bottom of the picture, just slightly to the right of center, forms a face.) The change in perception from “rocks in a stream” or “trees in a forest” into “faces” is a change in the perceptual organization of the rocks and the trees. The two shapes that you at first perceive as two separate rocks in the stream become perceptually grouped together when they become the left and right eyes of a face. In fact, once you perceive a particular grouping of rocks as a face, it is often difficult *not* to perceive them in this way—they have become permanently organized into a face. This effect of meaning on perceptual organization is an example of the operation of top-down processing in perception.

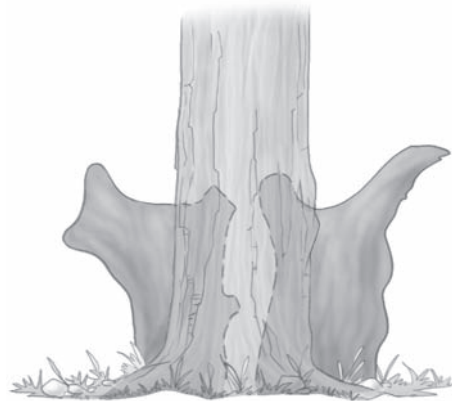
The Gestalt Laws Provide “Best Guess” Predictions About What Is Out There

The purpose of perception is to provide accurate information about the properties of the environment. The Gestalt laws help provide this information because they reflect things we know from long experience in our environment and because we are using them unconsciously all the time. For example, the law of good continuation reflects our understanding that many objects in the environment have straight or smoothly curving contours, so when we see smoothly curving contours, such as the wires in Figure 3.28, we correctly perceive the two wires.

The Gestalt laws usually result in accurate perceptions of the environment, but not always. We can illustrate a situation in which the Gestalt laws might cause an incorrect perception by imagining the following: As you are hiking in the woods, you stop cold in your tracks because, not too far ahead, you see what appears to be an animal lurking behind a tree (Figure 3.32a). The Gestalt laws of organization play a role in creating this perception. You see the two dark shapes to the left and right of the tree as a single object because of the Gestalt law of similarity (because both shapes are dark, it is likely that



(a)



(b)

■ **Figure 3.32** (a) What lurks behind the tree? (b) It is two strangely shaped tree stumps, not an animal!

they are part of the same object). Also, good continuation links these two parts into one, because the line along the top of the object extends smoothly from one side of the tree to another. Finally, the image resembles animals you've seen before. For all of these reasons, it is not surprising that you perceive the two dark objects as part of one animal.

Because you fear that the animal might be dangerous, you take a different path, and as your detour takes you around the tree, you notice that the dark shapes aren't an animal after all, but are two oddly shaped tree stumps (Figure 3.32b). So in this case, the Gestalt laws have misled you.

Because the Gestalt laws do not always result in accurate perceptions of the environment, it is more correct to call them *heuristics* rather than *laws*. A **heuristic** is a “rule of thumb” that provides a best-guess solution to a problem. Another way of solving a problem, an **algorithm**, is a procedure guaranteed to solve a problem. An example of an algorithm is the procedures we learn for addition, subtraction, and long division. If we apply these procedures correctly, we get the right answer every time. In contrast, a heuristic may not result in a correct solution every time.

To illustrate the difference between a heuristic and an algorithm, let's consider two different ways of finding a cat hiding somewhere in the house. An algorithm for doing this would be to systematically search every room in the house (being careful not to let the cat sneak past you!). If you do this, you will eventually find the cat, although it may take a while. A heuristic for finding the cat would be to first look in the places where the cat likes to hide. So you check under the bed and in the hall closet. This may not always lead to finding the cat, but if it does, it has the advantage of being faster than the algorithm.

The fact that heuristics are usually faster than algorithms helps explain why the perceptual system is designed in a way that sometimes produces errors. Consider, for example, what the algorithm would be for determining what the shape in Figure 3.32a really is. The algorithm would involve walking around the tree so you can see the shape from different angles, perhaps taking a more close-up look at the objects behind the tree and maybe even poking them to see if they move. Although this may result in an accurate determination of what the shapes are, it is potentially risky (what if the shape actually *is* a dangerous animal?), and slow. The advantage of our Gestalt-based heuristics is that they are fast and are correct most of the time.

The influence of knowledge and the top-down processing that accompanies knowledge means that it is accurate to describe perception as being “intelligent.” This intel-

ligence becomes apparent when we bring our knowledge of faces to bear on the creation of faces in the rocks and trees of ● Color Plate 3.3. However, we could argue that there is a certain intelligence behind even simpler processes, such as grouping by similarity and nearness. The idea that these simple grouping processes could involve intelligence is perhaps not obvious because they seem so automatic. In fact, people often react to some of the Gestalt laws as if they are simply common sense. Our skeptic from the beginning of the chapter, who thought perception was simple, might say, “Of course things that are close to each other will become grouped. I don’t think there’s much intelligence involved in that.”

It is easy to understand why someone might say this because these groupings usually happen so easily and naturally that it doesn’t appear that much of anything is going on. It is a case of perception appearing to just “happen.” But in reality there is a lot going on, because the Gestalt laws are based on characteristics of our environment. Grouping is easy because our perceptual system is tuned to respond, so when we encounter things that commonly occur in the environment, we will be likely to perceive them accurately. The need for perceptual intelligence becomes more obvious when we consider some of the problems both computers and humans must solve in order to perceive objects.

Why Computers Have Trouble Perceiving Objects

At the beginning of the chapter, we noted that in the 1960s computer scientists predicted the problem of perception would be easily solved. As it turned out it wasn’t easy at all, and although a chess-playing computer beat the world chess champion in 1997, it wasn’t until 2005 that computer-controlled vehicles were able to successfully navigate a course that involved avoiding obstacles while traveling over varied types of terrain (Figure 3.33).

Image not available due to copyright restrictions