Prof. Noureddine Mouhadjer

LMD3

Cognitive Psychology

Visual Perception

These lectures are totally based on the book "Cognitive Psychology" written by ROBERT J. STERNBERG et al (2012). I have opted for this book to teach Cognitive Psychology to third year LMD students, for its richness with texts and illustrations and for its very technical and easy to follow approach of introducing them.

Visual Perception

If we look out of the window the buildings that are less than a mile away look about as small as our computer screen. Yet we know that they are actually much bigger than our screen—they only appear to be small. This is just one example of the complex process of perception.

Perception is the set of processes by which we recognize, organize, and make sense of the sensations we receive from environmental stimuli (Goodale, 2000a, 2000b; Kosslyn & Osherson, 1995; Marr, 1982; Pomerantz, 2003). Perception encompasses many psychological phenomena. In this chapter, we focus primarily on visual perception. It is the most widely recognized and the most widely studied perceptual modality (i.e., system for a particular sense, such as touch or smell).

First, we will get to know a few basic terms and concepts of perception. We will then consider optical illusions that illustrate some of the intricacies of human perception. Next, we will have a look at the biology of the visual system. We will consider some approaches to explain perception, and afterward have a closer look at some details of the perceptual process, namely the perception of objects and forms, and how the environment provides cues to help you perceive your surroundings. We will also explore what happens when people have difficulties in perception.

1. Interpretation of Sensation

Our brain actively tries to make sense of the many stimuli that enter our eyes and fall on our retina. Take a look at Figure 1. You can see two high-rise buildings. In the right photo, the right tower seems to be substantially higher than the left one. The left picture, however, shows that the towers actually are in fact exactly the same height. Depending on your viewpoint, objects can look quite different, revealing different details. Thus, perception does not consist of just seeing what is being projected onto your retina; the process is much more complex. Your brain processes the visual stimuli, giving the stimuli meaning and interpreting them.

Figure 1:Visual Interpretation

How difficult it is to interpret what we see has become clear in recent years as researchers have tried to teach computers to "see"; but computers are still lagging behind humans in object recognition as computers cannot figure out what is depicted in the photo if it is a reflection of another picture. So, while it may not take you a lot of effort to identify the objects in a photo reflected on a mirror, it does take a lot of processing to perceive them, as the stimuli are very ambiguous.

2. Some Basic Concepts of Perception

In his influential and controversial work, James Gibson (1966, 1979) provided a useful framework for studying perception. He introduced the concepts of distal (external) object,

informational medium, proximal stimulation, and perceptual object. Let's examine each of these.

The distal (far) object is the object in the external world (e.g., Thunder). The event of thunder creates a pattern on an informational medium. The informational medium could be sound waves, as in the sound of the thunder. The informational medium might also be reflected light, chemical molecules, or tactile information coming from the environment. For example, when the information from light waves come into contact with the appropriate sensory receptors of the eyes, proximal (near) stimulation occurs (i.e., the cells in your retina absorb the light waves). Perception occurs when a perceptual object (i.e., what you see) is created in you that reflects the properties of the external world. That is, an image of a falling tree is created on your retina that reflects the falling tree that is in front of you.

Table 1 lists the various properties of distal objects, informational media, proximal stimuli, and perceptual objects for five different senses (sight, sound, smell, taste, and touch). The processes of perception vary tremendously across the different senses.

Perception occurs when the informational medium carries information about a distal object to a person. When the person's sense receptors pick up on the information, proximal stimulation occurs, which results in the person's perceiving an object.

Modality	Distal Object	Informational Medium	Proximal Stimulation	Perceptual Object
Vision—sight	Grandma's face	Reflected light from Grandma's face (visible electromagnetic waves)	Photon absorption in the rod and cone cells of the retina, the receptor surface in the back of the eye	Grandma's face
Audition—sound	A falling tree	Sound waves generated by the tree's fall	Sound-wave conduction to the basilar membrane, the receptor surface within the cochlea of the inner ear	A falling tree
Olfaction—smell	Bacon being fried	Molecules released by frying bacon	Molecular absorption in the cells of the olfactory epithelium, the receptor surface in the nasal cavity	Bacon
Gustation—taste	Ice cream	Molecules of ice cream both released into the air and dissolved in water	Molecular contact with taste buds, the receptor cells on the tongue and soft palate, combined with olfactory stimulation	Ice cream
Touch	A computer keyboard	Mechanical pressure and vibration at the point of contact between the surface of the skin and the keyboard	Stimulation of various receptor cells within the dermis, the innermost layer of skin	Computer keys

Table 1: Perceptual Continuum

3. Seeing Things That Aren't There

To find out about some of the phenomena of perception, psychologists often study situations that pose problems in making sense of our sensations. Consider, for example, the image displayed in Figure 2. To most people, the figure initially looks like a blur of meaningless shadings. A recognizable creature is staring them in the face, but they may not see it. When people finally realize what is in the figure, they rightfully feel "cowed." The figure of the cow is hidden within the continuous gradations of shading that constitute the picture. Before you recognized the figure as a cow, you correctly sensed all aspects of the figure. But you had not yet organized those sensations to form a mental percept—that is, a mental representation of a stimulus that is perceived. Without such a percept of the cow, you could not meaningfully grasp what you previously had sensed.

Figure 2: Dallenbach's Cow.

Source: From Dallenbach, K. M. (1951). A puzzle-picture with a new principle of concealment. American Journal of Psychology, 54, 431–433.

The preceding examples show that sometimes we cannot perceive what does exist. At other times, however, we perceive things that do not exist. For example, notice the black triangle in the center of the left panel of following Figure. Also note the white triangle in the center of the right panel of the Figure. They jump right out at you.

Figure 3: Elusive Triangles: Real or Illusions?

Now look very closely at each of the panels. You will see that the triangles are not really all there. The black that constitutes the center triangle in the left panel looks darker, or blacker, than the surrounding black. But it is not. Nor is the white central triangle in the right panel any brighter, or whiter, than the surrounding white. Both central triangles are optical illusions. They involve the perception of visual information not physically present in the visual sensory stimulus.

So, sometimes we perceive what is not there. Other times, we do not perceive what is there. And at still other times, we perceive what cannot be there. The existence of perceptual illusions suggests that what we sense (in our sensory organs) is not necessarily what we perceive (in our minds). Our minds must be taking the available sensory information and manipulating that information somehow to create mental representations of objects, properties, and spatial relationships within our environments (Peterson, 1999).

Architects are not the only ones to have recognized some fundamental principles of perception. For centuries, artists have known how to lead us to perceive 3-D percepts when viewing two-dimensional (2-D) images. What are some of the principles that guide our perceptions of both real and illusory percepts? We will explore the answer to this question as we move through the chapter. We begin with examining our visual system.

4. How Does Our Visual System Work?

The precondition for vision is the existence of light. Light is electromagnetic radiation that can be described in terms of wavelength. Humans can perceive only a small range of the wavelengths that exist; the visible wavelengths are from 380 to 750 nanometers (Figure 4; Starr, Evers, & Starr, 2007).
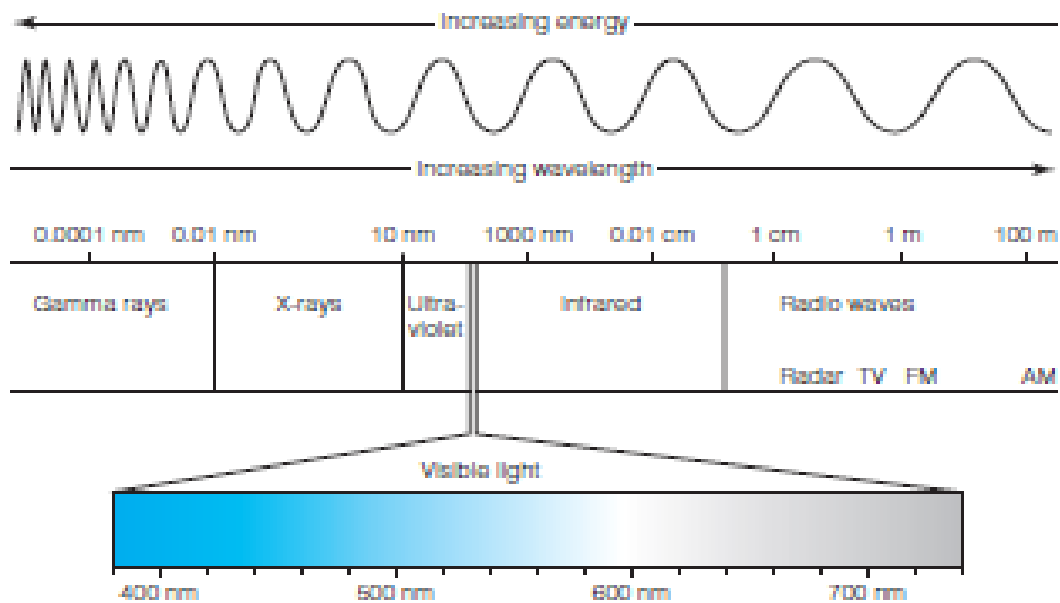
Figure 4: The Electromagnetic Spectrum.

Vision begins when light passes through the protective covering of the eye (Figure 5). This covering, the cornea, is a clear dome that protects the eye. The light then passes through the pupil, the opening in the center of the iris. It continues through the crystalline lens and the vitreous humor. The vitreous humor is a gel-like substance that comprises the majority of the eye.

Eventually, the light focuses on the retina where electromagnetic light energy is transduced—that is, converted—into neural electrochemical impulses (Blake, 2000).
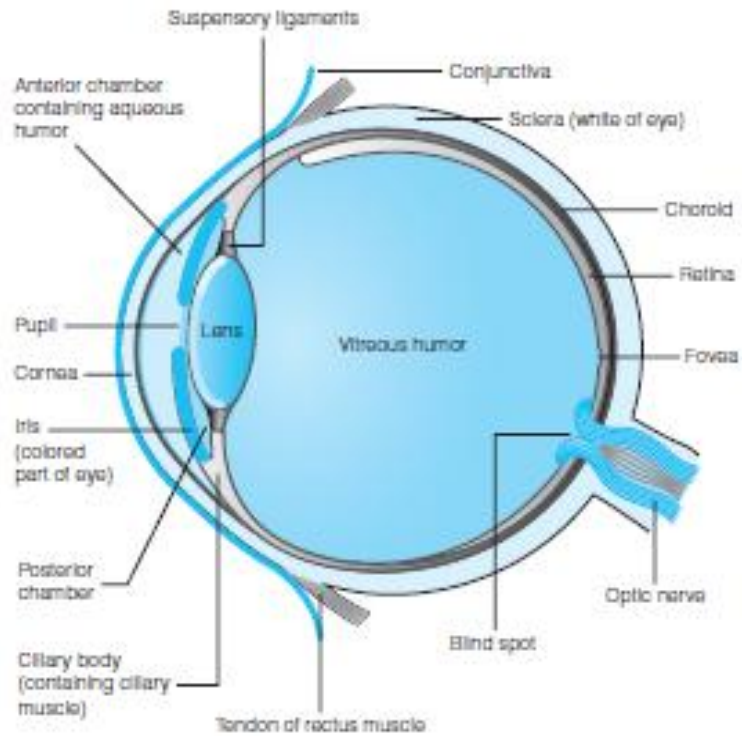
Figure 5: The Human Eye.

Vision is most acute in the fovea, which is a small, thin region of the retina, the size of the head of a pin. When you look straight at an object, your eyes rotate so that the image falls directly onto the fovea. Although the retina is only about as thick as a single page in this book, it consists of three main layers of neuronal tissue (Figure 6).
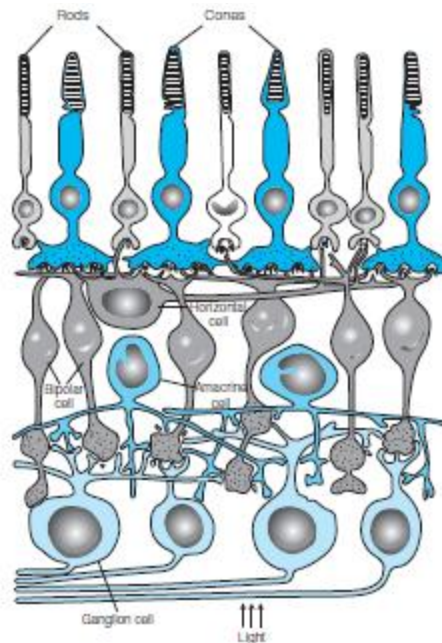

Figure 6: The Retina

The first layer of neuronal tissue—closest to the front, outward-facing surface of the eye—is the layer of ganglion cells, whose axons constitute the optic nerve. The second layer consists of three kinds of interneuron cells. Amacrine cells and horizontal cells make single lateral (i.e., horizontal) connections among adjacent areas of the retina in the middle layer of cells. Bipolar cells make dual connections forward and outward to the ganglion cells, as well as backward and inward to the third layer of retinal cells.

The third layer of the retina contains the photoreceptors, which convert light energy into electrochemical energy that is transmitted by neurons to the brain. There are two kinds of photoreceptors—rods and cones. Each eye contains roughly 120 million rods and 8 million cones. Rods and cones differ not only in shape but also in their compositions, locations, and responses to light. Within the rods and cones are photopigments, chemical substances that react to light and transform physical electromagnetic energy into an electrochemical neural impulse that can be understood by the brain. The rods are long and thin photoreceptors. They are more highly concentrated in the periphery of the retina than in the foveal region. The rods are responsible for night vision and are sensitive to light and dark stimuli. The cones are short and thick photoreceptors and allow for the perception of color. They are more highly concentrated in the foveal region than in the periphery of the retina (Durgin, 2000).

The rods, cones, and photopigments could not do their work were they not somehow hooked up to the brain. The neurochemical messages processed by the rods and cones of the retina travel via the bipolar cells to the ganglion cells (see Goodale, 2000a, 2000b). The axons of the ganglion cells in the eye collectively form the optic nerve for that eye. The optic nerves of the two eyes join at the base of the brain to form the optic chiasma. At this point, the ganglion cells from the inward, or nasal, part of the retina—the part closer to your nose—cross through the optic chiasma and extend to the opposite hemisphere of the brain. The ganglion cells from the outward, or temporal area of the retina closer to your temple go to the hemisphere on the same side of the body. The lens of each eye naturally inverts the image of the world as it projects the image onto the retina. In this way, the message sent to your brain is literally upside-down and backward.

After being routed via the optic chiasma, about 90% of the ganglion cells then go to the lateral geniculate nucleus of the thalamus. From the thalamus, neurons carry information to the primary visual cortex (V1 or striate cortex) in the occipital lobe of the brain. The visual cortex

contains several processing areas. Each area handles different kinds of visual information relating to intensity and quality, including color, location, depth, pattern, and form.

5. Pathways to Perceive the What and the Where

What are the visual pathways in the brain? A pathway in general is the path the visual information takes from its entering the human perceptual system through the eyes to its being completely processed. Generally, researchers agree that there are two pathways. Work on visual perception has identified separate neural pathways in the cerebral cortex for processing different aspects of the same stimuli (De Yoe & Van Essen, 1988; Köhler et al., 1995). Perception deficits like ataxia and agnosia thatare covered later in this chapter also point toward the existence of different pathways.

Why are there two pathways? It is because the information from the primary visual cortex in the occipital lobe is forwarded through two fasciculi (fiber bundles): One ascends toward the parietal lobe (along the dorsal pathway), and one descends to the temporal lobe (along the ventral pathway). The dorsal pathway is also called the where pathway and is responsible for processing location and motion information; the ventral pathway is called the what pathway because it is mainly responsible for processing the color, shape, and identity of visual stimuli (Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982).

This general view is referred to as the what/where hypothesis. Most of the research in this area has been carried out with monkeys. In particular, a group of monkeys with lesions in the temporal lobe were able to indicate where things were but seemed unable to recognize what they were. In contrast, monkeys with lesions in the parietal lobe were able to recognize what things were but not where they were.

An alternative interpretation of the visual pathways has been suggested. This interpretation is that the two pathways refer not to what things are and to where they are, but rather, to what they are and to how they function. This view is known as the what/how hypothesis (Goodale & Milner, 2004; Goodale & Westwood, 2004).This hypothesis argues that spatial information about where something is located in space is always present in visual information processing. What differs between the two pathways is whether the emphasis is on identifying what an object is or, instead, on how we can situate ourselves so as to grasp the object.

The what pathway can be found in the ventral stream and is responsible for the identification of objects. The how pathway is located in the dorsal stream and controls movements in relation to the objects that have been identified through the "what" pathway. Ventral and dorsal streams both arise from the same early visual areas (Milner & Goodale, 2008).

The what/how hypothesis is best supported by evidence of processing deficits: There are deficits that impair people's ability to recognize what they see and there are distinct deficits that impair people's ability to reach for what they see (how).

CHECK YOUR KNOWLEDGE

1. What is the difference between sensation and perception?
2. What is the difference between the distal and the perceptual object?
3. How are rods and cones both similar to and different from each other?
4. What are some of the major parts of the eye and what are their functions?
5. What is the "what/where" hypothesis?

6. **How Do We Make Sense of What We See?**

   Approaches to Perception

Now that we know how a light stimulus that enters our eye is processed and routed to the brain, the question still remains as to how we actually perceive what we see. Do we just perceive whatever is being projected on our retina, or is there more to perception? Does our knowledge, and other rules we have learned throughout our life, maybe influence our perception of the world? Going back to our view out of the window, the image on our retina suggests that the buildings we see in the distance are very small.

However, we do see other buildings, trees, and streets in front of them that suggest that those buildings are in fact quite large and just appear small because they are far away from our office. In this case, our experience and knowledge about perception and the world allows us to perceive those buildings as tall ones even though they do not look larger than does our hand in front of us on our desk.

There are different views on how we perceive the world. These views can be summarized as bottom-up theories and top-down theories. Bottom-up theories describe approaches where perception starts with the stimuli whose appearance you take in through your eye. You look out onto the cityscape, and perception happens when the light information is transported to your brain. Therefore, they are datadriven (i.e., stimulus-driven) theories.

Not all theorists focus on the sensory data of the perceptual stimulus. Many theorists prefer top-down theories, according to which perception is driven by high-level cognitive processes, existing knowledge, and the prior expectations that influence perception (Clark, 2003). These theories then work their way down to considering the sensory data, such as the perceptual stimulus. You perceive buildings as big in the background of the city scene because you know these buildings are far away and therefore must be bigger than they appear. From this viewpoint, expectations

are important. When people expect to see something, they may see it even if it is not there or is no longer there. For example, suppose people expect to see a certainperson in a certain location.

They may think they see that person, even if they are actually seeing someone else who looks only vaguely similar (Simons, 1996). Top-down and bottom-up approaches have been applied to virtually every aspect of cognition. Bottom-up and top-down approaches usually are presented as

being in opposition to each other. But to some extent, they deal with different aspects of the same phenomenon. Ultimately, a complete theory of perception will need to encompass both bottom-up and top-down processes.

6.1 Bottom-Up Theories

The four main bottom-up theories of form and pattern perception are direct perception, template theories, feature theories, and recognition-by-components theory.

Direct Perception

How do you know the letter A when you see it? Easy to ask, hard to answer. Of course, it's an A because it looks like an A. What makes it look like an A, though, instead of like an H? Just how difficult it is to answer this question becomes apparent when you look at Figure 7. You

probably will see the image in Figure 7 as the words "THE CAT." Yet the H of "THE" is identical to the A of "CAT." What subjectively feels like a simple process of pattern recognition is almost certainly quite complex.

Figure 7: Can read these words?

Gibson's Theory of Direct Perception How do we connect what we perceive to what we have stored in our minds? Gestalt psychologists referred to this problem as the Hoffding function (Köhler, 1940). It was named after 19th-century Danish psychologist Harald Hoffding. He questioned whether perception is such a simple process that all it takes is to associate what is seen with what is remembered (associationism). An influential and controversial theorist who questioned associationism is James J. Gibson (1904–1980).

According to Gibson's theory of direct perception, the information in our sensory receptors, including the sensory context, is all we need to perceive anything. As the environment supplies us with all the information we need for perception, this view is sometimes also called ecological perception. In other words, we do not need higher cognitive processes or anything else to mediate between our sensory experiences and our perceptions. Existing beliefs or higher-level inferential thought processes are not necessary for perception.

Gibson believed that, in the real world, sufficient contextual information usually exists to make perceptual judgments. He claimed that we need not appeal to higher level intelligent processes to explain perception. Gibson (1979) believed that we use this contextual information directly. In essence, we are biologically tuned to respond to it. According to Gibson, we use texture gradients as cues for depth and distance. Those cues aid us to perceive directly the relative proximity or distance of objects and of parts of objects.
need the aid of complex thought processes.

Direct perception may also play a role in interpersonal situations when we try to make sense of others' emotions and intentions (Gallagher, 2008). After all, we can recognize emotion in faces as such; we do not see facial expressions that we then try to piece together to result in the perception of an emotion (Wittgenstein, 1980).

Neuroscience and Direct Perception Neuroscience also indicates that direct perception may be involved in person perception. About 30 to 100 milliseconds after a visual stimulus, mirror neurons start firing. Mirror neurons are active both when a person acts and when he or she observes that same act performed by somebody else. So before we even have time to form hypotheses about what we are perceiving, we may already be able to understand the expressions, emotions, and movements of the person we observe (Gallagher, 2008).

Furthermore, studies indicate that there are separate neural pathways (what pathways) in the lateral occipital area for the processing of form, color, and texture in objects. When asked to judge the length of an object, for example, people cannot ignore the width. However, they can judge the color, form, and texture of an object independently of the other qualities (Cant & Goodale, 2007; Cant, Large, McCall, & Goodale, 2008).

Template Theories

Template theories suggest that we have stored in our minds myriad sets of templates. Templates are highly detailed models for patterns we potentially might recognize. We recognize a pattern by comparing it with our set of templates. We then choose the exact template that perfectly matches what we observe (Selfridge & Neisser, 1960). We see examples of template matching in our everyday lives. Fingerprints are matched in this way. Machines rapidly process imprinted numerals on checks by comparing them to templates. Increasingly, products of all kinds are identified with universal product codes (UPCs or "bar codes"). They can be scanned and identified by computers at the time of purchase. Chess players who have knowledge of many games use a matching strategy in line with template theory to recall previous games (Gobet & Jackson, 2002).

Template matching theories belong to the group of chunk-based theories that suggest that expertise is attained by acquiring chunks of knowledge in long-term memory that can later be accessed for fast recognition. Studies with chess players have shown that the temporal lobe is

indeed activated when the players access the stored chunks in their long-term memory (Campitelli, Gobet, Head, Buckley, & Parker, 2007).

In each of the aforementioned instances, the goal of finding one perfect match and disregarding imperfect matches suits the task. You would be alarmed to find that your bank's numeral-recognition system failed to register a deposit to your account. Such failure might occur because it was programmed to accept an ambiguous character according to what seemed to be a best guess. For template matching, only an exact match will do. This is exactly what you want from a bank computer. However, consider your perceptual system at work in everyday situations. It rarely would work if you required exact matches for every stimulus you were to recognize. Imagine, for example, needing mental templates for every possible percept of the face of someone you love. Imagine one for each facial expression, each angle of viewing, each addition or removal of makeup, each hairdo, and so on.

Template-matching theories fail to explain some aspects of the perception of letters. For one thing, such theories cannot easily account for our perception of the letters and words in Figure 7. We identify two different letters (A and H) from only one physical form. Hoffding (1891) noted other problems. We can recognize an A as an A despite variations in the size, orientation, and form in which the letter is written. Are we to believe that we have mental templates for each possible size, orientation, and form of a letter? Storing, organizing, and retrieving so many templates in memory would be unwieldy. How could we possibly anticipate and create so many templates for every conceivable object of perception (Figure 8)? Neuroscience and Template Theories Letters of the alphabet are simpler than faces and other complex stimuli. But how do we recognize letters? And does it make a difference to our brain whether we perceive letters or digits? Experiments suggest that there is indeed a difference between letters and digits.

There is an area on or near the left fusiform gyrus that is activated significantly more when a person is presented with letters than with digits. It is not clear if this "letter area" only processes letters or if it also plays a more minor role in the processing of digits (Polk et al., 2002). The notion of the visual cortex specializing in different stimuli is not new; other areas have been found that specialize in faces, for example (see Kanwisher et al., 1997; McCarthy et al., 1997).

Why Computers Have Trouble Reading Handwriting Think about how easy it is for you to perceive and understand someone's handwriting. In handwriting, everybody's numbers and letters look a bit different. You can still distinguish them without any problems (at least in most cases). This is something computers do not do very well at all. For computers, the reading of handwriting is an incredibly difficult process that's prone to mistakes. When you deposit a check at an ATM machine, it "reads" your check automatically. In fact, the numbers at the bottom of your check that are written in a strange-looking font are so distinct that a machine cannot mistake them for one another. However, it is much harder for a machine to decipher handwriting. Similarly, a machine also will have trouble determining that all the letters in the right of Figure 8 are As (unless it has a template for each one of the As). Therefore, some computers work with algorithms that consider the context in which the word is presented, the angular positions of the written letters (e.g., upright or tilted), and other factors.

Given the sophistication of current-day robots, what is the source of human superiority?
There may be several, but one is certainly knowledge. We simply know
much more about the environment and sources of regularity in the environment
than do robots. Our knowledge gives us a great advantage that robots, at least of
the current day, are still unable to bridge.

Figure 8: Templates matching in Barcodes and Letters

Feature-Matching Theories

Yet another alternative explanation of pattern and form perception may be found in feature-matching theories. According to these theories, we attempt to match features of a pattern to features stored in memory, rather than to match a whole pattern to a template or a prototype (Stankiewicz, 2003).

The Pandemonium Model One such feature-matching model has been called Pandemonium ("pandemonium" refers to a very noisy, chaotic place and hell). In it, metaphorical "demons" with specific duties receive and analyze the features of a stimulus (Selfridge, 1959). In Oliver Selfridge's Pandemonium Model, there are four kinds of demons: image demons, feature demons, cognitive demons, and decision demons. Figure 9 shows this model.
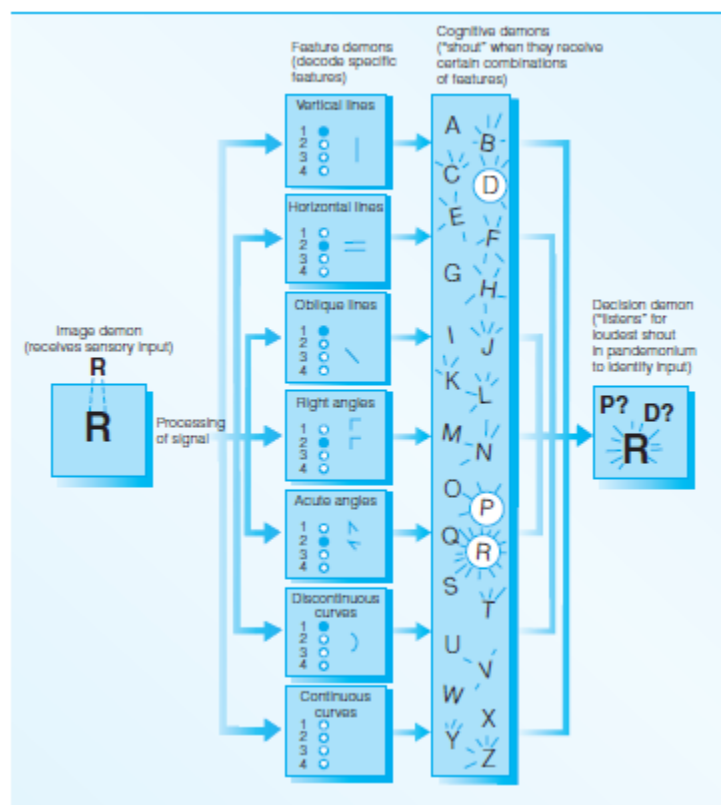

Figure 9: Selfridge's Feature-Matching Model

The "image demons" receive a retinal image and pass it on to "feature demons." Each feature demon calls out when there are matches between the stimulus and the given feature. These matches are yelled out at demons at the next level of the hierarchy, the "cognitive (thinking) demons." The cognitive demons in turn shout out possible patterns stored in memory that conform to one or more of the features noticed by the feature demons.
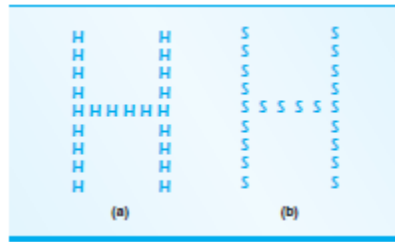
Figure 10: The Global Precedence Effect.

A "decision demon" listens to the pandemonium of the cognitive demons. It decides on what has been seen, based on which cognitive demon is shouting the most frequently (i.e., which has the most matching features).

Although Selfridge's model is one of the most widely known, other feature models have been proposed. Most also distinguish not only different features but also different kinds of features, such as global versus local features. Local features constitute the small-scale or detailed aspects of a given pattern. There is no consensus as to what exactly constitutes a local feature. Nevertheless, we generally can distinguish such features from global features, the features that give a form its overall shape.

Consider, for example, the stimuli depicted in Figure 10 (a) and (b). These are of the type used in some research on pattern perception (see for example Navon, 1977, or Olesen et al., 2007). Globally, the stimuli in panels (a) and (b) form the letter H. In panel (a), the local features (small Hs) correspond to the global ones. In panel (b), comprising many local letter Ss, they do not.

In one study, participants were asked to identify the stimuli at either the global or the local level (Navon, 1977). When the local letters were small and positioned close together, participants could identify stimuli at the global level (the "big" letter) more quickly than at the local level. When participants were required to identify stimuli at the global level, whether the local features (small letters) matched the global one (big letter) did not matter. They responded equally rapidly whether the global H was made up of local Hs or of local Ss. However, when the participants were asked to identify the "small" local letters, they responded more quickly if the global features agreed with the local ones. In other words, they were slowed down if they had to identify local (small) Ss combining to form a global (big) H instead of identifying local (small) Hs combining to form a global (big) H. This pattern of

results is called the global precedence effect (see also Kimchi, 1992). Experiments have showed that global information dominates over local information even in infants (Cassia, Simion, Milani, & Umiltà, 2002).

In contrast, when letters are more widely spaced, as in panels (a) and (b) of Figure 11, the effect is reversed. Then a local precedence effect appears. That is, the participants more quickly identify the local features of the individual letters than the global ones, and the local features interfere with the global recognition in cases of contradictory stimuli (Martin, 1979).
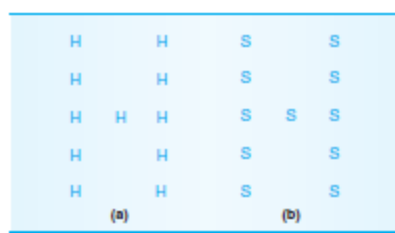

Figure 11: The Local Precedence Effect.

So when the letters are close together at the local level, people have problems identifying the local stimuli (small letters) if they are not concordant with the global stimulus (big letter). When the letters on the local level are relatively far apart from each other, it is harder for people to identify the global stimulus (big letter) if it is not concordant with the local stimuli (small letters). Other limitations (e.g., the size of the stimuli) besides special proximity of the local stimuli hold as well, and other kinds of features also influence perception.

Neuroscience and Feature-Matching Theories Some support for feature theories comes from neurological and physiological research. Researchers used single-cell recording techniques with animals (Hubel & Wiesel, 1963, 1968, 1979). They carefully measured the responses of individual neurons to visual stimuli in the visual cortex. Then they mapped those neurons to corresponding visual stimuli for particular locations in the visual field (see Chapter 2). Their research showed that the visual cortex contains specific neurons that respond only to a particular kind of stimulus (e.g., a horizontal line), and only if that stimulus fell onto a specific region of the retina. Each individual cortical neuron, therefore, can be mapped to a specific receptive field on the retina. A disproportionately large amount of the visual cortex is

devoted to neurons mapped to receptive fields in the foveal region of the retina, which is the area of the most acute vision.

Most of the cells in the cortex do not respond simply to spots of light. Rather, they respond to "specifically oriented line segments" (Hubel & Wiesel, 1979, p. 9). What's more, these cells seem to show a hierarchical structure in the degree of complexity of the stimuli to which they respond, somewhat in line with the ideas behind the Pandemonium Model. That means that the outputs of the cells are combined to create higher-order detectors that can identify increasingly more complex features.

At the lowest level, cells respond to lines, at a higher level they respond to corners and edges, then to shapes, and so forth. Neurons that can recognize a complex object are called gnostic units or "grandmother cells" because they imply that there is a neuron that is capable of recognizing your grandmother. None of those neurons are quite so specific, however, that they respond to just one person's head. Even at such a high level there is still some selectivity involved that allows cells to generally fire when a human face comes into view.

Consider what happens as the stimulus proceeds through the visual system to higher levels in the cortex. In general, the size of the receptive field increases, as does the complexity of the stimulus required to prompt a response. As evidence of this hierarchy, there were once believed to be just two kinds of visual cortex neurons, simple cells and complex cells (Hubel & Wiesel, 1979), which were believed to differ in the complexity of the information about stimuli they processed.

Recognition-by-Components Theory

How do we form stable 3-D mental representations of objects? The recognition-by-components theory explains our ability to perceive 3-D objects with the help of simple geometric shapes.

Seeing with the Help of Geons Irving Biederman (1987) suggested that we achieve this by manipulating a number of simple 3-D geometric shapes called geons (for geometrical ions). They include objects such as bricks, cylinders, wedges, cones, and their curved axis counterparts (Biederman, 1990/1993b). According to Biederman's recognition-by-components (RBC) theory, we quickly recognize objects by observing the edges of them and

then decomposing the objects into geons. The geons also can be recomposed into alternative arrangements. You know that a small set of letters can be manipulated to compose countless words and sentences. Similarly, a small number of geons can be used to build up many basic shapes and then myriad basic objects (Figure 12).

Biederman's RBC theory explains how we may recognize general instances of chairs, lamps, and faces, but it does not adequately explain how we recognize particular chairs or particular faces. An example would be your own face or your best friend's face. They are both made up of geons that constitute your mouth, eyes, nose, eyebrows, and so forth. But these geons are the same for both your and your friend's faces.

So RBC theory cannot explain how we can distinguish one face from the next. Biederman recognized that aspects of his theory require further work, such as how the relations among the parts of an object can be described (Biederman, 1990/1993b). Another problem with Biederman's approach, and the bottom-up approach in general, is how to account for the effects of prior expectations and environmental context on some phenomena of pattern perception.

Neuroscience and Recognition-by-Components Theory What results would we expect if we were to confirm Biederman's theory? Geons are viewpoint-invariant, so studies should show that neurons exist that react to properties of an object that stay the same, no matter whether you look at them from the front or the side. And indeed, there are studies that have found neurons in the inferior temporal cortex that are sensitive to just those viewpoint-invariant properties (Vogels et al., 2001).

However, many neurons respond primarily to one view of an object and decrease their response gradually the more the object is rotated (Logothetis, Pauls, & Poggio, 1995). This finding contradicts the notion of Biederman's theory that we recognize objects by means of viewpoint-invariant geons. As a result, it is not clear at this point whether Biederman's theory is correct.
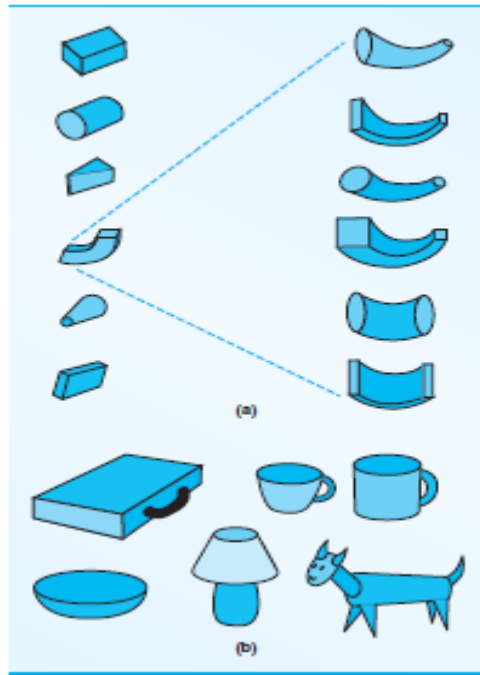
Figure 12: Geons

6.2 Top-Down Theories

In contrast to the bottom-up approach to perception is the top-down, constructive approach (Bruner, 1957; Gregory, 1980; Rock, 1983; von Helmholtz, 1909/1962). In constructive perception, the perceiver builds (constructs) a cognitive understanding (perception) of a stimulus. The concepts of the perceiver and his or her cognitive processes influence what he or she sees. The perceiver uses sensory information as the foundation for the structure but also uses other sources of information to build the perception. This viewpoint also is known as intelligent perception because it states that higher-order thinking plays an important role in perception. It also emphasizes the role of learning in perception (Fahle, 2003). Some investigators have pointed out that not only does the world affect our perception but also the world we experience is actually formed by our perception (Goldstone, 2003). In other words, perception is reciprocal with the world we experience. Perception both affects and is affected by the world as we experience it.

An interesting feature of the theory of constructive perception is that it links human intelligence even to fairly basic processes of perception. According to this theory, perception comprises not merely a low-level set of cognitive processes, but actually a quite sophisticated set of processes that interact with and are guided by human intelligence. When you look out your window, you "see" many things, but what you recognize yourself as seeing is highly

processed by your intelligence. Interestingly, Titchener's structuralist approach (described in Chapter 1) ultimately failed because despite the efforts of Titchener and his followers to engage in introspection independently of their prior knowledge, they and others found this, in the end, to be impossible.

What you perceive is shaped, at some level, by what you know and what you think. For example, picture yourself driving down a road you have never traveled before. As you approach a blind intersection, you see an octagonal red sign with white lettering. It bears the letters "ST_P." An overgrown vine cuts between the T and the P. Chances are, you will construct from your sensations a perception of a stop sign. You thus will respond appropriately.

Perceptual constancies are another example (see below). When you see a car approaching you on the street, its image on your retina gets bigger as the car comes closer. And yet, you perceive the car to stay the same size. This suggests that high-level constructive processes are at work during perception. In color constancy, we perceive that the color of an object remains the same despite changes in lighting that alter the hue. Even in lighting that becomes so dim that color sensations are virtually absent, we still perceive bananas as yellow, plums as purple, and so on.

According to constructivists, during perception we quickly form and test various hypotheses regarding percepts. The percepts are based on three things:
• what we sense (the sensory data),
• what we know (knowledge stored in memory), and
• what we can infer (using high-level cognitive processes).

In perception, we consider prior expectations. You'll be fast to recognize your friend from far away on the street when you have arranged a meeting. We also use what we know about the context. When you see something approaching on rail tracks you infer that it must be a train. And we also may use what we reasonably can infer, based both on what the data are and on what we know about the data.

According to constructivists, we usually make the correct attributions regarding our visual sensations. The reason is that we perform unconscious inference, the process by which we

unconsciously assimilate information from a number of sources to create a perception (Snow & Mattingley, 2003). In other words, using more than one source of information, we make judgments that we are not even aware of making.

In the stop-sign example, sensory information implies that the sign is a meaningless assortment of oddly spaced consonants. However, your prior learning tells you something important—that a sign of this shape and color posted at an intersection of roadways and containing these three letters in this sequence probably means that you should stop thinking about the odd letters. Instead, you should start slamming on the brakes. Successful constructive perception requires intelligence and thought in combining sensory information with knowledge gained from previous experience.

One reason for favoring the constructive approach is that bottom-up (datadriven) theories of perception do not fully explain context effects. Context effects are the influences of the surrounding environment on perception (e.g., our perception of "THE CAT" in Figure 7). Fairly dramatic context effects can be demonstrated experimentally (Biederman, 1972; Biederman et al., 1974; Biederman, Glass, & Stacy, 1973; De Graef, Christiaens, & D'Ydewalle, 1990). In one study, people were asked to identify objects after they had viewed the objects in either an appropriate or an inappropriate context for the items (Palmer, 1975). For example,participants might see a scene of a kitchen followed by stimuli such as a loaf of bread, a mailbox, and a drum. Objects that were appropriate to the established context, such as the loaf of bread in this example, were recognized more rapidly than were objects that were inappropriate to the established context.

The strength of the context also plays a role in object recognition (Bar, 2004). Perhaps even more striking is a context effect known as the configural-superiority effect (Bar, 2004; Pomerantz, 1981), by which objects presented in certain configurations are easier to recognize than the objects presented in isolation, even if the objects in the configurations are more complex than those in isolation. Suppose you show a participant four stimuli, all of them diagonal lines [see Figure 13 (a)]. Three of the lines are slanting one way, and one line is slanting the other way. The participant's task is to identify which stimulus is unlike the others. Now suppose that you show participants four stimuli that are comprised of three lines each [Figure 13 (c)]. Three of the stimuli are shaped like triangles, and one is not. In each case, the stimulus is a diagonal line [Figure 13 (a)] plus other lines [Figure 13 (b)]. Thus, the stimuli in this second condition are more complex variations of the stimuli in the first condition.

However, participants can more quickly spot which of the three-sided, more complicated figures is different from the others than they can spot which of the lines is different from the others.
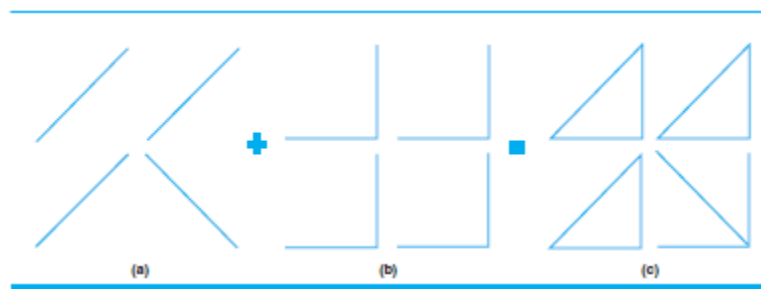

Figure 13: The Configural-Superiority Effect.

In a similar vein, there is an object-superiority effect, in which a target line that forms a part of a drawing of a 3-D object is identified more accurately than a target that forms a part of a disconnected 2-D pattern (Lanze, Weisstein, & Harris, 1982; Weisstein & Harris, 1974). These findings parallel findings in the study of letter and word recognition: The word-superiority effect indicates that when people are presented with strings of letters, it is easier for them to identify a single letter if the string makes sense and forms a word instead of being just a nonsense sequel of letters. For example, it is easier to recognize the letter "o" in the word "house" than in the word "huseo" (Reicher, 1969).

The viewpoint of constructive or intelligent perception shows the central relation between perception and intelligence. According to this viewpoint, intelligence is an integral part of our perceptual processing. We do not perceive simply in terms of what is "out there in the world." Rather, we perceive in terms of the expectations and other cognitions we bring to our interaction with the world. In this view, intelligence and perceptual processes interact in the formation of our beliefs about what it is that we are encountering in our everyday contacts with the world at large.

An extreme top-down position would drastically underestimate the importance of sensory data. If it were correct, we would be susceptible to gross inaccuracies of perception. We frequently would form hypotheses and expectancies that inadequately evaluated the sensory data available. For example, if we expected to see a friend and someone else came into view, we might inadequately consider the perceptible differences between the friend and a stranger

and mistake the stranger for the friend. Thus, an extreme constructivist view of perception would be highly errorprone and inefficient. However, an extreme bottom-up position would not allow for any influence of past experience or knowledge on perception. Why store knowledge that has no use for the perceiver? Neither extreme is ideal for explaining perception. It is more fruitful to consider ways in which bottom-up and top-down processes interact to form meaningful percepts.

6.3 How Do Bottom-Up Theories and Top-Down Theories Go Together?

Both theoretical approaches have garnered empirical support (cf. Cutting & Kozlowski, 1977, vs. Palmer, 1975). So how do we decide between the two? On one level, the constructive-perception theory, which is more top-down, seems to contradict direct-perception theory, which is more bottom-up. Constructivists emphasize the importance of prior knowledge in combination with relatively simple and ambiguous information from the sensory receptors. In contrast, directperception theorists emphasize the completeness of the information in the receptors themselves. They suggest that perception occurs simply and directly. Thus, there is little need for complex information processing. Instead of viewing these theoretical approaches as incompatible, we may gain deeper insight into perception by considering the approaches to be complementary.

Sensory information may be more richly informative and less ambiguous in interpreting experiences than the constructivists would suggest. But it may be less informative than the direct-perception theorists would assert. Similarly, perceptual processes may be more complex than hypothesized by Gibsonian theorists. This would be particularly true under conditions in which the sensory stimuli appear only briefly or are degraded. Degraded stimuli are less informative for various reasons. For example, the stimuli may be partially obscured or weakened by poor lighting. Or they may be incomplete, or distorted by illusory cues or other visual "noise" (distracting visual stimulation analogous to audible noise). We likely use a combination of information from the sensory receptors and our past knowledge to make sense of what we perceive.Some experimental evidence supports this integrated view (Treue, 2003; van Zoest & Donk, 2004; Wolfe et al., 2003).

Recent work suggests that, whereas the very first stage of the visual pathway represents only what is in the retinal image of an object, very soon, color, orientation, motion, depth, spatial frequency, and temporal frequency are represented. Later-stage representations emphasize the

viewer's current interest or attention. In other words, the later-stage representations are not independent of our attentional focus. On the contrary, they are directly affected by it (Maunsell, 1995). Moreover, vision for different things can take different forms. Visual control of action is mediated by cortical pathways that are different from those involved in visual control of perception (Ganel & Goodale, 2003). In other words, when we merely see an object, such as a cell phone, we process it differently than if we intend also to pick up the object. In general, according to Ganel and Goodale (2003), we perceive objects holistically. But if we plan to act on them, we perceive them more analytically so that we can act in an effective way.

To summarize, current theories concerning the ways we perceive patterns explain some, but not all, of the phenomena we encounter in the study of form and pattern perception. Given the complexity of the process, it is impressive that we understand as much as we do. At the same time, clearly a comprehensive theory is still forthcoming. Such a theory would need to account fully for the kinds of context effects, such as the configural-superiority effect, described in this section.

7. Perception of Objects and Forms

Do we perceive objects in a viewer-centered or in an object-centered way? When we gaze at any object in the space around us, do we perceive it in relation to us rather than its actual structure, or do we perceive it in a more objective way that is independent of how it appears to us right this moment? We'll examine this question in the next section. Then, we look at Gestalt principles for perception, which explain why we perceive some objects as in groups but others as not so grouped (what is it that makes some birds flying in the afternoon sky appear to be in a group whereas others do not?). Finally, we will consider the question of how we perceive patterns, for example faces.

Right now one of your authors is looking at the computer on which he is typing this text. He depicts the results of what he sees as a mental representation. What form does this mental representation take? There are two common positions regarding the answer to this question. One position, viewer-centered representation, is that the individual stores the way the object looks to him or her. Thus, what matters is the appearance of the object to the viewer (in this case, the appearance of the computer to the author), not the actual structure of the object. The shape of the object changes, depending on the angle from which we look at it. A number of

views of the object are stored, and when we try to recognize an object, we have to rotate that object in our mind until it fits one of the stored images.

The second position, object-centered representation, is that the individual stores a representation of the object, independent of its appearance to the viewer. In this case, the shape of the object will stay stable across different orientations (McMullen & Farah, 1991). This stability can be achieved by means of establishing the major and minor axes of the object, which then serve as a basis for defining further properties of the object.

Both positions can account for how the author represents a given object and its parts. The key difference is in whether he represents the object and its parts in relation to him (viewer-centered) or in relation to the entirety of the object itself, independent of his own position (object-centered). Consider, for example, the computer on which this text is being written. It has different parts: a screen, a keyboard, a mouse, and so forth. Suppose the author represents the computer in terms of viewer-centered representation. Then its various parts are stored in terms of their relation to him. He sees the screen as facing him at perhaps a 20-degree angle. He sees the keyboard facing him horizontally. He sees the mouse off to the right side and in front of him. Suppose, instead, that he uses an object-centered representation. Then he would see the screen at a 70-degree angle relative to the keyboard. And the mouse is directly to the right side of the keyboard, neither in front of it nor in back of it.

One potential reconciliation of these two approaches to mental representation suggests that people may use both kinds of representations. According to this approach, recognition of objects occurs on a continuum (Burgund & Marsolek, 2000; Tarr, 2000; Tarr & Bülthoff, 1995). At one end of this continuum are cognitive mechanisms that are more viewpoint-centered. At the other end of the continuum are cognitive mechanisms that are more object-centered. For example, suppose you see a picture of a car that is inverted. How do you know it is a car? Object-centered mechanisms would recognize the object as a car, but viewpoint-centered mechanisms would recognize the car as inverted.

A third orientation in representation is landmark-centered. In landmark-centered representation, information is characterized by its relation to a well-known or prominent item. Imagine visiting a new city. Each day you leave your hotel and go on short trips. It is easy to imagine that you would represent the area you explore in relation to your hotel. Evidence

indicates that, in the laboratory, participants can switch between these three strategies. There are, however, differences in brain activation among these strategies (Committeri et al., 2004).

CHECK YOUR KNOWLEDGE

1. What are the major Gestalt principles?

2. What is the "recognition by components" theory?

3. What is the difference between top-down and bottom-up theories of perception?

4. What is the difference between viewer-centered and object-centered perception?

8. The Environment Helps You See

As we have seen, perceptual processes are not so easily completed that the image on your retina can be taken as is without further interpretation. Our brain needs to interpret the stimuli it receives and make sense of them. The environment provides cues that aid in the analysis of the retinal image and facilitate the construction of a perception that is as close as possible to what is out there in the world—at least, to the extent we can ascertain what is out there! The following part of this chapter explains how we use environmental cues to perceive the world.

8.1 Perceptual Constancies

Picture yourself walking to your cognitive psychology class. Two students are standing outside the classroom door. They are chatting as you approach. As you get closer to the door, the amount of space on your retina devoted to images of those students becomes increasingly large. On the one hand, this proximal sensory evidence suggests that the students are becoming larger. On the other hand, you perceive that the students have remained the same size. Why?

The perceptual system deals with variability by performing a rather remarkable analysis regarding the objects in the perceptual field. Your classmates' perceived constancy in size is an example of perceptual constancy. Perceptual constancy occurs when our perception of an object remains the same even when our proximal sensation of the distal object changes (Gillam, 2000). The physical characteristics of the external distal object are probably not changing. But because we must be able to deal effectively with the external world, our perceptual system has mechanisms that adjust our perception of the proximal stimulus. Thus, the perception remains constant although the proximal sensation changes. Here we consider two of the main constancies: size and shape constancies.

Size constancy is the perception that an object maintains the same size despite changes in the size of the proximal stimulus. The size of an image on the retina depends directly on the distance of that object from the eye. The same object at two different distances projects different-sized images on the retina. Some striking illusions can be achieved when our sensory and perceptual systems are misled by the very same information that usually helps us to achieve size constancy.

An example of size constancy is the Müller-Lyer illusion (Figure 14). Here, two line segments that are of the same length appear to be of different lengths. We use shapes and angles from our everyday experience to draw conclusions about the relative sizes of objects. Equivalent image sizes at different depths usually indicate different-sized objects.

Studies indicate that the right posterior parietal cortex (involved in the manipulation of mental images) and the right temporo-occipital cortex are activated when people are asked to judge the length of the lines in the Müller-Lyer illusion. The strength of the illusion can be changed by adjusting the angles of the arrows that delimit the horizontal line—the sharper the angles, the more pronounced the illusion. The strength of the illusion is associated with bilateral (on both sides) activation in the lateral (i.e., located on the side of) occipital cortex and the right superior parietal cortex. As the right intraparietal sulcus (furrow) is activated as well, it seems like there is an interaction of the illusory information with the top-down processes in the right parietal cortex that are responsible for visuo-spatial judgments (Weidner & Fink, 2007).
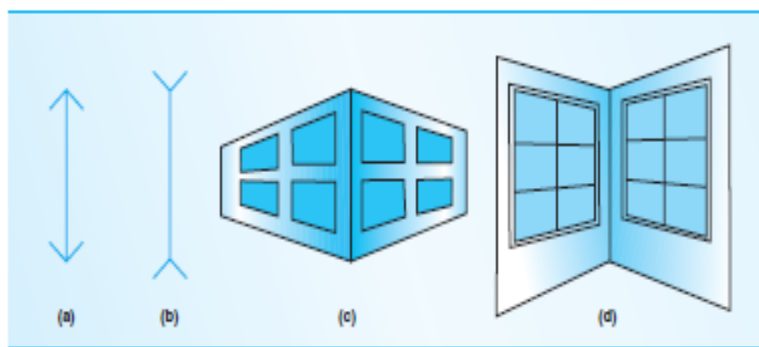

Figure 14: The Müller-Lyer Illusion

Like size constancy, shape constancy relates to the perception of distances but in a different way. Shape constancy is the perception that an object maintains the same shape despite changes in the shape of the proximal stimulus (Figure 15). An object's perceived shape remains the same despite changes in its orientation and hence in the shape of its retinal image. As the actual shape of the pictured door changes, some parts of the door seem to be changing differentially in their distance from us.

It is possible to use neuropsychological imaging to localize parts of the brain that are used in this shape analysis. They are in the extrastriate cortex (Kanwisher et al., 1996, 1997). Points near the outer edge of the door seem to move more quickly toward us than do points near the inner edge. Nonetheless, we perceive that the door remains the same shape.

Figure 15: Shape Constancy

8.2. Depth Perception

Consider what happens when you reach for a cup of tea, or throw a baseball. You must use information regarding depth. Depth is the distance from a surface, usually using your own body as a reference surface when speaking in terms of depth perception.

This use of depth information extends beyond the range of your body's reach. When you drive, you use depth to assess the distance of an approaching automobile. When you decide to

call out to a friend walking down the street, you determine how loudly to call. Your decision is based on how far away you perceive your friend to be. How do you manage to perceive 3-D space when the proximal stimuli on your retinas comprise only a 2-D projection of what you see? You have to rely on depth cues. The next section explores what depth cues are and how we use them.

Depth Cues

Look at the impossible configurations in Figure 3.25. They are confusing because there is contradictory depth information in different sections of the picture. Small segments of these impossible figures look reasonable to us because there is no inconsistency in their individual depth cues (Hochberg, 1978). However, it is difficult to make sense of the figure as a whole. The reason is that the cues providing depth information in various segments of the picture are in conflict.

Generally, depth cues are either monocular (mon-, "one"; ocular, "related to the eyes") or binocular (bin-, "both," "two"). Monocular depth cues can be represented in just two dimensions and observed with just one eye. Figure 16 illustrates several of the monocular depth cues defined in Table 2. They include texture gradients, relative size, interposition, linear perspective, aerial perspective, location in the picture plane, and motion parallax. Before you read about the cues in either the table or the figure caption, look just at the figure. See how many depth cues you can decipher simply by observing the figure carefully.
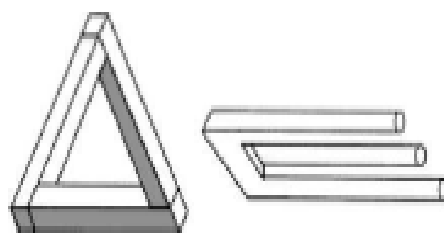

Figure 16 : Impossible Figures.

Motion parallax requires movement. It thus cannot be used to judge depth within a stationary image, such as a picture. Another means of judging depth involves binocular depth cues, based on the receipt of sensory information in three dimensions from both eyes (Parker, Cumming, & Dodd, 2000). Table 3.3 also summarizes some of the binocular cues used in perceiving depth.

Binocular depth cues use the relative positioning of your eyes. Your two eyes are positioned far enough apart to provide two kinds of information to your brain: binocular disparity and binocular convergence. In binocular disparity, your two eyes send increasingly disparate (differing) images to your brain as objects approach you. Your brain interprets the degree of disparity as an indication of distance from you. In addition, for objects we view at relatively close locations, we use depth cues based on binocular convergence.

Various perceptual cues aid in our perception of the 3-D world. Some of these cues can be observed by one eye alone; other cues require the use of both eyes.

Cues for Depth Perception	Appears Closer	Appears Farther Away
Monocular Depth Cues		
Texture gradients	Larger grains, farther apart	Smaller grains, closer together
Relative size	Bigger	Smaller
Interposition	Partially obscures other object	Is partially obscured by other object
Linear perspective	Apparently parallel lines seem to diverge as they move away from the horizon	Apparently parallel lines seem to converge as they approach the horizon
Aerial perspective	Images seem crisper, more clearly delineated	Images seem fuzzier, less clearly delineated
Location in the picture plane	Above the horizon, objects are higher in the picture plane; below the horizon, objects are lower in the picture plane	Above the horizon, objects are lower in the picture plane; below the horizon, objects are higher in the picture plane
Motion parallax	Objects approaching get larger at an ever-increasing speed (i.e., big and moving quickly closer)	Objects departing get smaller at an ever-decreasing speed (i.e., small and moving slowly farther away)
Binocular Depth Cues		
Binocular convergence	Eyes feel tug inward toward nose	Eyes relax outward toward ears
Binocular disparity	Huge discrepancy between image seen by left eye and image seen by right eye	Minuscule discrepancy between image seen by left eye and image seen by right eye

Table 2: Monocular and Binocular Cues for Depth Perception

In binocular convergence, your two eyes increasingly turn inward as objects approach you. Your brain interprets these muscular movements as indications of distance from you. In about 8% of people whose eyes are not aligned properly (strabismic eyes), depth perception can occur even with just one eye. Usually people with strabismic eyes have a sensitive zone in their retina other than the fovea that captures a part of the space that should have been captured were the eyes properly aligned. This capacity normally goes along with a partial inhibition of signals from the fovea.

If the fovea stays sensitive, however, those people produce double images, which can be fused and result in stereoscopic vision with just one eye (Rychkova & Ninio, 2009). Depth

perception may depend upon more than just the distance or depth at which an object is located relative to oneself. The perceived distance to a target is influenced by the effort required to walk to the location of the target (Proffitt et al., 2003, 2006). People with a heavy backpack perceive the distance to a target location as farther than those not wearing a heavy backpack. In other words, there can be an interaction between the perceptual result and the perceived effort required to reach the object perceived (Wilt, Proffitt, & Epstein, 2004).

The more effort one requires to reach something, the farther away it is perceived to be. Depth perception is a good example of how cues facilitate our perception. When we see an object that appears small, there is no automatic reason to believe it is farther away. Rather, the brain uses this contextual information to conclude that the smaller object is farther away.

9. Anomalies in Color Perception

Color perception deficits are much more common in men than in women, and they are genetically linked. However, they can also result from lesions to the ventromedial occipital and temporal lobes. There are several kinds of color deficiency, which are sometimes referred to as kinds of "color blindness." Least common is rod monochromacy, also called achromacy. People with this condition have no color vision at all. It is thus the only true form of pure color blindness. People with this condition have cones that are nonfunctional. They see only shades of gray, as a function of their vision through the rods of the eye.

Most people who suffer from deficits in color perception can still see some color, despite the name "color blindness." In dichromacy, only two of the mechanisms for color perception work, and one is malfunctioning. The result of this malfunction is one of three types of color blindness (color-perception deficits). The most common is red-green color blindness. People with this form of color-blindness have difficulty in distinguishing red from green, although they may be able to distinguish, for example, dark red from light green (Visual disabilities: Color-blindness, 2004).

The extreme form of red-green color blindness is called protanopia. The other types of color blindness are: deuteranopia (trouble seeing greens), and tritanopia (blues and greens can be confused, but yellows also can seem to disappear or to appear as light shades of reds). See the companion website for a picture showing a rainbow as seen by a person with normal color vision and by persons suffering from the three kinds of dichromacy.

CHECK YOUR KNOWLEDGE

1. What is shape constancy?

2. What are the main cues for depth perception?

3. What is visual agnosia?

4. To what does "modularity" refer?

5. What is the difference between monochromacy and dichromacy?

10. Why Does It Matter? Perception in Practice

Perceptual processes and change blindness play a significant role in accidents and efforts at accident prevention. About 50% of all collision accidents are a result of missing or delayed perception (Nakayama, 1978). Especially two-wheeled vehicles are often involved in "looked-but-failed-to-see" accidents, where the driver of the involved car states that he did indeed look in the direction of the cyclist, but failed to see the approaching motorcycle. It is possible that drivers develop a certain "scanning" strategy that they use in complex situations, such as at crossroads. The scanning strategy concentrates on the most common and dangerous threats but fails to recognize small deviations, or more uncommon objects like two-wheeled vehicles.

In addition, people tend to fail to recognize new objects after blinking and saccades (fast movements of both eyes in one direction). Generally, people are not aware of the danger of change blindness and believe that they will be able to see all obstacles when looking in a particular direction ("change blindness blindness", Simons & Rensink, 2005; Davis et al., 2008). This tendency has implications for the education of drivers with regard to their perceptual abilities. It also has implications for the design of traffic environments, which should be laid out in a way that facilitates complex traffic flow and makes drivers aware of unexpected obstacles, like bicycles (Galpin et al., 2009; Koustanai, Boloix, Van Elslande, & Bastien, 2008).

11. Key Themes

Several key themes, as outlined in Chapter 1, emerge in our study of perception.

Rationalism versus empiricism. How much of the way we perceive can be understood

as due to some kind of order in the environment that is relatively independent of our perceptual mechanisms? In the Gibsonian view, much of what we perceive derives from the structure of the stimulus, independent of our experience with it. In contrast, in the view of constructive perception, we construct what we perceive. We build up mechanisms for perceiving based on our experience with the environment. As a result, our perception is influenced at least as much by our intelligence (rationalism) as it is by the structure of the stimuli we perceive (empiricism).

Basic versus applied research. Research on perception has many applications, such as in understanding how we can construct machines that perceive. The U.S. Postal Service relies heavily on machines that read zip codes. To the extent that the machines are inaccurate, mail risks going astray. These machines cannot rely on strict template matching because people write numbers in different ways. So the machines must do at least some feature analysis.
Another application of perception research is in human factors. Human-factors researchers design machines and user interfaces to be user-friendly. An automobile driver or airplane pilot sometimes needs to make split-second decisions. The cockpits thus must have instrument panels that are well-lit, easy to read, and accessible for quick action. Basic research on human perception can inform developers what user-friendly means.

Domain generality versus domain specificity. Perhaps nowhere is this theme better illustrated than in research on face recognition. Is there something special about face recognition? It appears so. Yet many of the mechanisms that are used for face recognition are used for other kinds of perception as well. Thus, it appears that perceptual mechanisms may be mixed—some general across domains, others specific to domains such as face recognition.