Chapter 7
Depth Perception

The possibility to see the world in three dimensions remains quite fascinating considering the following peculiar fact: images arrive on the retina in two dimensions. Spontaneously, no one would doubt that there is a physical world that includes a third dimension. Nevertheless, the brain has to build this third dimension. The present chapter addresses this phenomenon, that of the construction of space or, more specifically, of depth.

In the preceding chapters, different aspects of visual perception sent the reader back to fundamental concepts of physics for understanding the nature of stimuli or to fundamental concepts of physiology or brain sciences for understanding biological bases of perception. In the present chapter, special attention is paid to psychological phenomena allowing to perceiving this third dimension. After a review of the main cues for perceiving depth, two phenomena are examined closely. The first one, perceptual constancy, is fundamental and is indeed not involved only in depth perception. The second one reveals cases, most often fascinating and fun, where these cues are misleading and induce illusions.

7.1 Cues for Perceiving a Third Dimension

The capacity of perceiving depth is based on the availability and contribution of several cues. These cues are kind of tricks allowing the brain to generate this impression of depth. However, using these tricks requires no voluntary effort; they are spontaneously activated by the data provided by the visual scene.

For understanding how these cues are studied in the field of perception, it is appropriate to make some distinctions between terms or concepts. First, for studying

S. Grondin, *Psychology of Perception*, DOI 10.1007/978-3-319-31791-5_7

the estimated distance, for example, researchers will distinguish absolute distance and relative distance. The idea of absolute distance, sometimes also called self-centered, refers to the estimation of the distance between an object and the one observing this object. In contrast, reference is made to relative distance, or exocentric distance, to designate the distance between the objects or between parts of these objects. If we generally find it hard to precisely estimate the absolute distance, we are rather good to decide if an object is closer to us than another object.

In addition to this absolute vs. relative distance distinction, it is important to keep in mind that the various depth cues that can be used can be classified in the following three dichotomous categories. Cues may be extraocular (nonvisual) rather than visual. Thus, information on depth might not be extracted specifically from the visual system per se, but from a source belonging to another sensory modality, namely, the kinesthetic system. Another useful distinction is the one between pictorial cues, which are static, and kinematic cues, which are dynamic, i.e., related to movement. Finally, there are binocular cues, as opposed to monocular cues. The presentation of the various cues below is based on this distinction.

7.1.1 Binocular Cues

The simple fact of having two eyes, and to have some distance between them, provides a better perspective on what is happening in our environment. It is possible to perceive depth with only one eye, but some cues require the joint operation of both eyes. These *binocular* cues are very powerful because they add precision to our appreciation of the third dimension.

7.1.1.1 Binocular Convergence

A first cue involving the contribution of both eyes is called *binocular convergence*. Looking at an object usually means that both eyes converge on it. If an object is far away, the angle between the focal point (the object) and the eye is small. If the object is close, the convergence angle is larger. According to the angle of convergence, the eyeballs are more or less displaced. These movements generate nonvisual cues. These cues are based on the kinesthetic information provided by receptors located within the extraocular muscles, i.e., the muscles that allow the movement of the eyes.

It is the convergence movement, rather than the state of convergence, which would provide the kinesthetic cues. This binocular convergence provides information on the absolute distance. Also, this source of information would be more effective when objects are close to the observer (say less than 6 m). Indeed, you can feel that there is some work being done in the eyeballs when you approach a finger to your nose and try to follow it with your eyes.

7.1.1.2 Retinal Disparity

A second binocular cue is called *retinal disparity*. The fact that there is some distance between the eyes is not trivial at all. This means that when looking at an object, an observer gets two points of view on it; in other words, for a given object, two images are formed. Disparity is the term designating the fact of receiving two images of the same thing, and binocular disparity refers to the disparity caused by the fact of having two eyes. Perceiving depth on the basis of binocular cues is also called stereoscopy (and sometimes binocular parallax), and the device designed to simulate a sense of depth with two different images of the same object, one for each eye, is called stereoscope.

It is easy to realize that each eye offers a unique point of view with the following exercise. First, place a finger 15 cm in front of you at eye level. By closing one eye and then the other alternately, you realize that different views of the finger appear. Equally important is the following fact. Place a second finger 15 cm behind the first, in the same axis. Now focus on the nearest finger, but pay attention to the farthest finger: it is seen in double. If you now focus on the farthest finger, but pay attention to the other one, it is now the closest one that is seen in double.

This difference in the clarity of the image as a function of the focus point is very important. It teaches us that not everything that is in our visual field is seen clearly. Horopter is the name given to the horizon line in front of us where vision is simple (no image seen in double); this line is actually an area called the Panum area. Depending on the distance of the fixation point, the size and shape of the area change slightly. Outside this area, the vision is in double. If simple vision is possible even if we have two eyes, therefore two images for one given fixation point, it is because for each given location on the retina of an eye, there is a precise correspondence point on the retina of the other eye.

Furthermore, it is important to note that the double vision of the close finger and that of the distant finger differ fundamentally. By placing two fingers as was done earlier and on focusing on the close finger, the farther finger is seen in double. By closing the right eye while continuing to focus on the close finger, the farther finger is seen with the left eye to the left of the close finger; now, by closing the left eye instead of the right one, the farther finger is seen with the right eye to the right of the near finger. Thus, when both eyes are open and we are fixating on a given point, what is located behind this point is seen in *uncrossed disparity*.

The demonstration works in the opposite direction. Now, the fixation point is the farther finger, but attention has to be allocated to the closer finger. By closing one eye and then the other alternately while keeping the fixation on the farther finger, one realizes that the closer finger is seen with the left eye to the right of the farther finger and with the right eye to the left of the farther finger. This is a case of *cross disparity*. The brain therefore benefits from a depth perception cue allowing to deciding whether object is in front a focal point or behind.

Finally, it should be noted that the visual field seen ahead binocularly covers a range of 120°. Added to this are approximately 40° of monocular vision to the left with the left eye and 40° of monocular vision to the right with the right eye.

7.1.2 Monocular Cues

Monocular cues of depth perception are sources of information about distance that remain available even when an observer uses only one eye. There are several monocular cues. Most of them are related to vision but one, *accommodation*, has a kinesthetic origin. In this case, the fact that an object is more or less distant causes a change in the shape of the lens. For instance, if an object is far away, the lens is less curved. Changing the shape of the lens requires changes in the contraction of the muscles involved in the accommodation process, and these muscular changes produce kinesthetic cues that the brain can interpret for assessing distance.

Because many objects have a typical size, it may happen that distance be estimated on the basis of this knowledge. For example, we know quite well the normal size of a card. If we are not under specific conditions such as those that cause optical-geometric illusions (see below), we can rely on the combination of this knowledge and retinal size for estimating distance. We call this index the *familiar size*. Thus, if we look at a coin that looks like a 25-cent coin, and if we do it in an environment where other depth perception cues are not available, we will assume that the size of this coin is normal for estimating how far away it is. If it should happen that this coin is actually smaller (because a friend is playing a trick or a researcher in psychology of perception studies the mechanisms of depth perception), this would be misleading, and the estimated distance would likely be incorrect.

Linear perspective is one of the most powerful depth perception cues, a cue that is most useful in the field of visual arts. When two lines like those shown in Fig. 7.1 converge to a vanishing point, they give a sense of depth. The points that are closer to each other appear to be farther away from the observer. The farther away is a part of the image, the smaller is the distance between each line on the retina. What we see in the real world in three dimensions can therefore be implemented on a two-dimensional image by adjusting the distance between the drawn objects and their size.

In fact, this linear perspective effect caused by the convergence could be considered a special case of a more general cue that James Jerome Gibson called *texture*. This is a mixture of both the linear perspective and the relative size of objects composing a visual scene. Most often we talk about texture gradients to denote the fact that the density and the size of the elements of a visual scene vary with their distance. Thus, as can be seen in Fig. 7.2, when dots are smaller and close to each other, they appear to be farther away. If what is viewed is farther away, the elements of the scene will be more compact.

Another visual cue for perceiving depth is called *occlusion*. This cue, also called *interposition*, refers to the fact that objects or parts of a visual scene are often hidden by other objects. What is covered necessarily appears to be farther away than what is causing occlusion. This powerful cue says nothing about the distance between the observer and the object, but gives an idea of the relative distance between objects. Figure 7.3 illustrates how powerful this cue is.

Fig. 7.1 Although the rails are parallel, the distance between them appears to decrease from the *bottom* of the image to the *middle*, which induces an impression of depth. This is an example of linear perspective

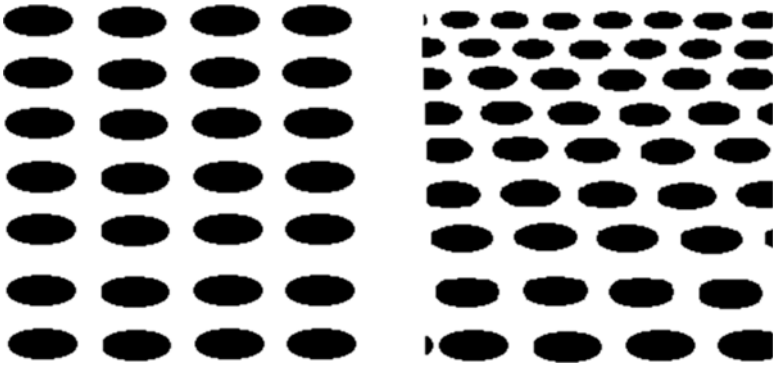

Fig. 7.2 When texture gradients are uniform, as on the *left*, no impression of depth is created; however, compressing points and using heterogeneous point size, on the *right*, give an impression of depth

Another cue based on the relationship between two objects allows to draw conclusions about their relative distance. This cue is called the *relative height*. The nearer an object is from the horizon point, the farther away it seems. Consequently, for the ground-related objects, i.e., below the horizon line (which is usually where our gaze is directed), the higher in the visual field an object is, the farther away it seems (Fig. 7.4). Conversely, for objects located above the horizon line, the lowest objects appear more distant.

Different arrangements of light and brightness can contribute to give an impression that objects or parts of the visual field are more or less close to an observer. For example, in the dark, the brighter of two objects is perceived as being closer.

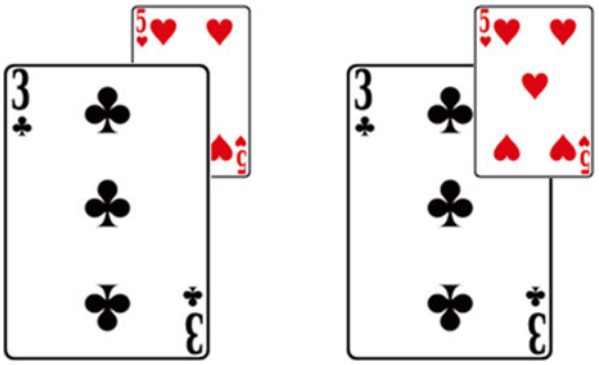

Fig. 7.3 If we are presented only the two cards on the *left*, it is easy to figure out that the cards have the same size, with the five of hearts being farther away than the three of clubs. It is not possible to reach the same conclusion for the cards on the *right* due to the interference caused by the occlusion factor. Because the five of hearts covers part of the three of clubs, the five of hearts must necessarily be in front of the other card (i.e., closer); as a result, it is not possible to believe that these cards are the same size

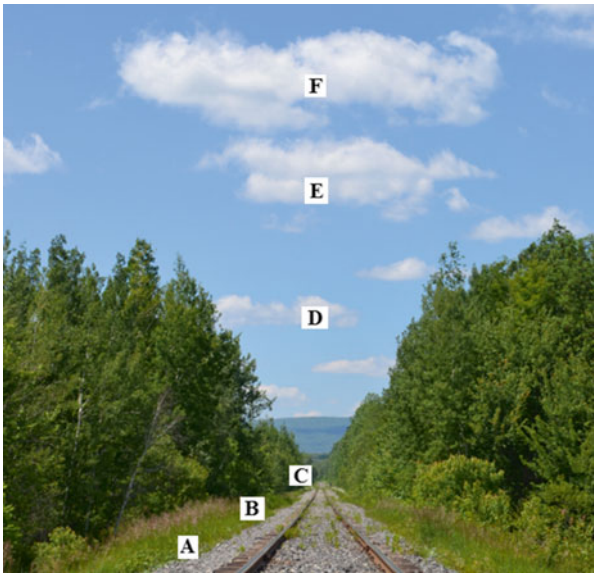

Fig. 7.4 Relative height in the visual field is a very strong indicator for distance perception. When looking at objects on the ground, those that are high in the field are farther away. It is therefore easy to understand that *C* is farther away than *B* and that *B* is farther away than *A*. When looking at the sky, the relative height is still a cue, but now, objects that are low in the visual field are farther away. Thus, cloud *D* is more distant than cloud *E*, which is itself farther away than cloud *F*

This cue is called *relative brightness*. Somewhat in the same vein, the use of shading allows to create an impression that something is more or less distant. In the case of Fig. 7.5, it seems that some circles are concave and others convex. Interestingly, this impression can be reversed by rotating the book 180°. In fact, the angle of illumination from a light source might well change the perception of an object, of an image, or of a face.

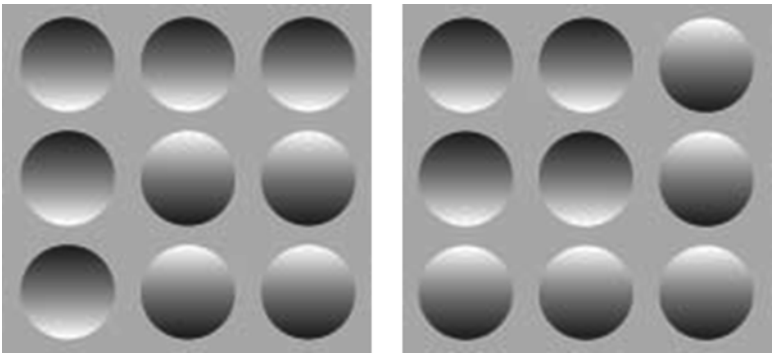

Fig. 7.5 Impressions of depth (concave versus convex) caused by shading. Images on the *left* and on the *right* are identical, but have been rotated 180°

It is also known that distance creates an attenuation of contours. This cue is called *aerial perspective* (or relative sharpness) and is more useful for estimating the distance of objects that are far away. The clearer the contour of an object, the closer the object appears because brightness contrasts are attenuated by distance. Thus, a car or even a mountain will appear much closer in clear weather than in fog conditions.

Another monocular cue, but linked to movement, is called *motion parallax*. The term parallax means shifting from one position to another. Here, we are talking about cases where a change is caused by the movement of an observer. To understand the relationship between the movement and the sense of distance, simply place a finger in front of you, and have your gaze fixated on it while moving your head to the left and to the right. When the head goes left, it seems that the finger goes right; when the head goes right, it seems that the finger goes to left. The wall behind the finger seems to go in the same direction as the head, but what lies between you and the finger (the fixation point) goes in the opposite direction. This provides an indication as to whether an object is in front or behind a fixation point. Even more important is the following point: the closer an object is, the greater the distance covered on the retina. Thus, the objects in the visual field give the impression of not moving at the same speed. The greater the speed, the closer the objects are. One can verify this statement by traveling on a road in the passenger seat: the gravel on the side of the road seems to fall quickly backward, whereas the mountain at some distance away, or a cloud, seems to follow you slowly. Now you know why it looks as if the moon follows us when traveling by car at night!

This section can be summarized by reference to Table 7.1 in which the cues for depth perception are classified as binocular or monocular, visual or nonvisual, and static or dynamic and on whether they are used to assess a relative or absolute distance.

In closing this portion of the chapter, it should be noted that the presence of certain cues in a visual scene determines the way we see. However, it may happen that

Table 7.1 Depth perception: summary and classification of cues

Cues	Type	Distance
Binocular		
Convergence[a]	K	A
Disparity	V	R
Monocular		
Accommodation	K	A
Familiar size	V	A
Relative height	V	R
Shading	V	R
Occlusion	V	R
Motion parallax[a]	V	R
Linear perspective	V	R
Texture	V	R

[a]Dynamic cues (other cues are static)
K kinesthetic, *V* visual, *A* absolute, *R* relative

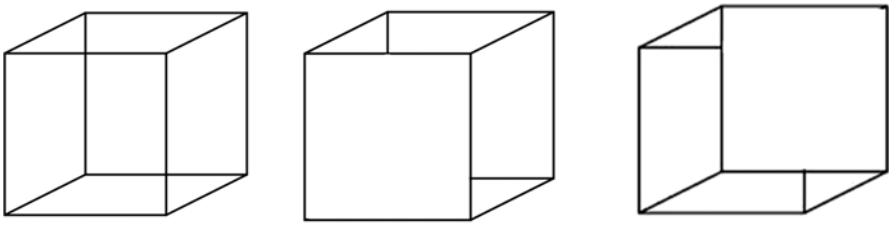

Fig. 7.6 Necker cube, on the *left*, could be perceived like the one at *center* or like the one on the *right*

the scene can be interpreted in different ways; in other words, cues may lead to some reversible images as is the case of the Necker cube (Fig. 7.6). On this figure, depending on what surface is perceived as occluding the other, a surface will be considered as being in the foreground or in the background. Similarly, the use in painting of some occluding effects can lead to the construction of pretty scenes that seem real, but could in no way be observed in nature. The Dutch artist Maurits Cornelis Escher has mastered the development of scenes involving this kind of deception. You may find some of the works of the artist if you just type his name on an Internet search engine. Similarly, typing "trompe-l'œil" on the Internet provides access to many other illustrations allowing to see how fine use of depth cues by painters can create powerful impressions, sometimes vertiginous, of a third dimension. We will return later to some particular impressions, namely, illusions, caused by the particular use of depth cues.

7.2 Perceptual Constancy

This section, which deals with the notion of perceptual constancy, could have been introduced at various places in this book, because it applies not only to size constancy but also to other dimensions, which will be detailed in the following paragraphs. The perceptual constancy is a basic mechanism of the perceptual system by which almost everything seems in order. Without that mechanism, we could not recognize anything. All physical stimuli would be chaos, and there would be no basis for perceiving.

What reaches the retina is continuously changing. If a chair is presented to an observer from an angle under which it has never been seen, the observer still manages to identify it as a chair. We can turn the chair in every sense and vary constantly the patterns of light energy it sends to the retina of the observer. This chair always maintains its objective characteristics, and the observer is able to know, without a shadow of a doubt, that two completely different energy patterns reaching his retina originate from the same object. In other words, seeing is not just a simple stimulation of retinal cells.

7.2.1 Types of Constancy

Among the different types of perceptual constancy, there is *shape constancy*. This constancy explains why an object maintains its shape even if different inclinations in different spatial planes cause as many variations of the projective image (see the previous chapter on form recognition). Similarly, under the *color constancy*, it is possible for an observer to recognize the hue of an object in spite of the fact that the light projected on this object changes its spectral composition, provided however that this change is not exaggerated. Also, the brightness of the object does not vary in spite of the differences in light intensity, and that is due to *brightness constancy*. Similarly, despite the differences in speed of the retinal image that can be caused by the distance in depth, it is possible to properly assess the speed of a moving object through a phenomenon referred to as *speed constancy*.

In the context of space perception, i.e., 3-D, the issue of perceptual constancy is closely linked to *size constancy*. The constancy refers to the capacity of maintaining the apparent size of objects or of people although the image size on the retina decreases with an increase of the distance between these objects or people and the observer. In other words, it is not because the retinal image of a person going away from the observer shrinks that this person appears to shrink. Unless there is in the environment a set of cues that may induce the observer in error, this observer continues to believe that this person is the same size.

7.2.2 *Interpretations and Investigations*

A classic question arises about the nature of size constancy: should distance be taken into account? In general, this question refers to the *size-distance invariance principle* (Kilpatrick & Ittelson, 1953). This hypothesis of invariance between size and distance basically states that an observer determines the apparent size on the basis of two combined elements of information, the perceived distance and the size of the retinal image. This idea is expressed by several authors in different forms. Thus, Helmholtz had already invoked the participation of a mechanism, the *unconscious inference*, to refer to the fact that the distance is taken into account in estimating the size of an object, this way of taking into account being settled without the help of conscious mechanisms. This theoretical perspective is also sometimes referred to as the algorithm theory (Epstein, 1977), as opposed to a relational theory. In the latter, the estimation of the size of an object or of person does not depend on some calculation of the distance between the observer and the object or person but rather on the relationship between the information available around the object or person. It is actually more a size-size type of relationship than a size-distance type. We will briefly return to this point of view in the next subsection.

The hypothesis about the need of taking the distance or not into account when size is evaluated has been tested in several empirical investigations. One way to illustrate the potential importance of distance in the evaluation of size is to use an afterimage. As we saw in Chap. 5 on color perception, a consecutive image is an image that remains somehow imprinted on the retina for a few seconds after prolonged stimulation. Remaining fixed on the retina, the image always maintains the same retinal size. The apparent size of this image depends on the distance of surface on which the image is projected. The farther away from the observer the projection surface is, the larger the image appears. This relationship between the apparent size of an afterimage and the distance from the observer to the projection surface is known as *Emmert's law*. This law illustrates the fact that the apparent size of an object depends not only on the size of the retinal image but also on the distance from which the object is perceived; therefore, apparent size likely depends on the fact of taking distance into account.

Among the various studies designed to test the size-distance invariance hypothesis, or the algorithm theory, the most classic is probably that of Holway and Boring (1941). In this experiment, some observers, including the authors, indicated the experimenter to adjust a comparison stimulus located about 10 ft (about 3 m) away. This adjustment was made for matching the size of a standard stimulus located in a long corridor at different distances, 10–120 ft (3–36 m), from the observer (Fig. 7.7). The stimuli, standard and comparison, were projected on screens. The images were uniform circular illuminations. In each experimental distance, the standard stimulus was adjusted so that the retinal image was kept constant, i.e., constantly subtended a visual angle of 1°.

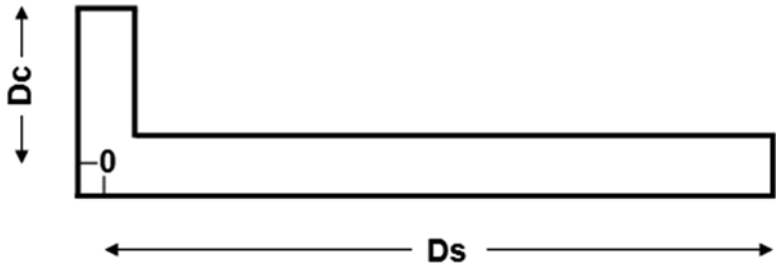

Fig. 7.7 Illustration of the experimental setting designed by Holway and Boring (1941). *Dc* distance of the comparison stimulus (10 ft ~ 3 m), *Ds* distance of standard stimulus (from 10 to 120 ft ~ de 3–36 m), *0* observer

The idea behind the experiment was to see if the adjustment of the comparison stimulus would be consistent with the actual size of the standard stimulus, as predicted by the size-distance invariance hypothesis. Thus, if the distance is not taken into account in the adjustment, the adjustment will always remain the same; however, if distance is taken into consideration, the adjustment will change as a function of the actual size of the stimulus or will get close to real size.

Holway and Boring pushed their reasoning a little further. If the distance is really considered, then different conditions for estimating distance should have an effect on the precision of the adjustment of the comparison stimulus. Thus, four experimental conditions were designed:

1. A binocular vision condition where the adjustment was expected to be the best
2. A condition where the only restriction was to use monocular vision
3. A monocular condition with vision through an artificial pupil, which was expected to reduce the cues provided by the motion parallax
4. A monocular vision condition through an artificial pupil and with low light conditions in order to reduce as much as possible potential sources of information on distance

Holway and Boring (1941) found that in conditions where cues were available for assessing distance, the adjustment of the comparison stimulus approximates the actual size of the object. In other words, even if the retinal size of the standard stimulus remains the same, the perceived size of the circular illumination changes according to the distance: the greater the distance, the greater the luminous circle, and the adjustment of the comparison stimulus is made accordingly. Figure 7.8 illustrates the results obtained in each condition. It should be noted that the loss of cues leads directly, as suggested by the slope of each function, to a decrease in the estimated size of the standard stimulus. All these results can be interpreted as supporting the size-distance invariance hypothesis, an idea often reported for explaining size constancy.

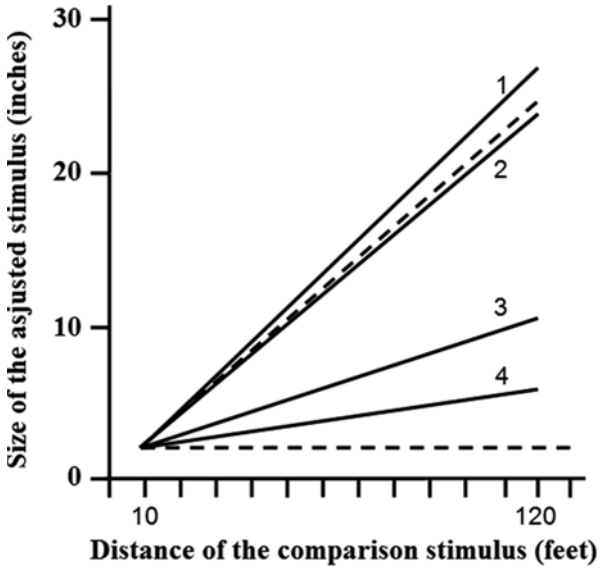

Fig. 7.8 Results of the experiment by Holway and Boring (1941—see their figure 22) where are grouped four experimental conditions: (*1*) binocular vision, (*2*) monocular vision, (*3*) monocular vision with artificial pupil, and (*4*) monocular vision with artificial pupil and reduced cues. The *broken lines* show the expected results would the perceptual constancy been perfect (*diagonal line*) and null (*horizontal line*) (1 in. ~ 2,54 cm; 1 ft ~ 30 cm)

7.2.3 Gibson's Perspective

Despite the elegance of this demonstration by Holway and Boring, other authors argue that this explanation based on the size-distance invariance can be faulted (Kilpatrick & Ittelson, 1953); it applies to the results in certain circumstances but cannot be a generalized. In fact, by removing cues of depth perception, the quality of the relational information is also reduced.

There is a radical position in the field of visual perception stating that there is no need for cognitive processing or inference mechanisms for estimating, for instance, depth. According to Gibson (1966, 1979), all the perceptual system needs is already available in the environment. In this Gibsonian perspective, everything that is in the environment (surfaces or objects) reaches the observer with specific physical characteristics. The movements of the observer determine what reaches the eye, and the material getting to this point is already organized. In the experiment by Holway and Boring, it was not possible for the observer to benefit from the cues normally provided by movements, especially in the condition involving to looking through an artificial pupil. Such an experimental design hinders the proper functioning, if not to say connivance, between the viewer and the environment.

Gibson therefore adopts what is referred to as an *ecological position* in which only natural situations can really contribute to our understanding of the visual

system. In this Gibsonian psychology, the environment provides us spontaneously not only precise physical stimuli but also information relative to the function of what is observed (e.g., when it is an object). In other words, seeing a chair also activates in the brain of the observer what a chair is for, i.e., sitting down. In the Gibsonian terminology, the idea that perceiving is inseparable from the function is called *affordance* (that is to say, what is made possible by what is observed).

7.3 Illusions

The perceptual systems are generally very reliable and allow to be adapted to the requirements of the environment and its characteristics. Despite the effectiveness of these systems, it happens that an observer is misled when these characteristics are somewhat special. In the field of visual perception, such misinterpretations have quite amazing, and sometimes even spectacular, consequences. These misinterpretations are caused not by a system failure as the inability to maintain perceptual constancy but by the objective characteristics of the environment.

These errors are called optical illusions or optical-geometric illusions. As they depend on the normal functioning of the visual system, these illusions provide an opportunity to inform us about the nature of perceptual processes. They should not be confused with *hallucinations*, which are a phenomenon where there is an impression of perception even though there is no perceptual object around (no physical stimuli) or *mirages*, which are a physical phenomenon caused by reflections of light rays in particular conditions.

7.3.1 Variety of Illusions

There are of course very strong visual effects like those caused by the subjective contours described in the previous chapter. In addition to these effects, there are hundreds of illusions that an interested reader can discover by consulting older books (see Coren & Girgus, 1978 or Shepard, 1990) or some specialized websites on the Internet. We present here only some of the most classic or of the most spectacular illusions. Many of these illusions were discovered in the nineteenth century, and in most cases, a given illusion was named after the person who reported it.

The classification of these illusions into a limited number of categories remains a difficult exercise (Coren, Girgus, Ehrlichman, & Hakstian, 1976). Some classifications like that reported by Gregory (1997) require many distinctions; that of Piaget is simpler. Piaget is famous for his work on the development of intelligence, but nevertheless studied in depth the impact of perception on knowledge. Some of his works, including those grouped in a book entitled *Les mécanismes perceptifs* (Piaget, 1961), concern the illusions in particular and their changes in magnitude with age. Inspired by Alfred Binet, who is distinguishing innate vs. acquired

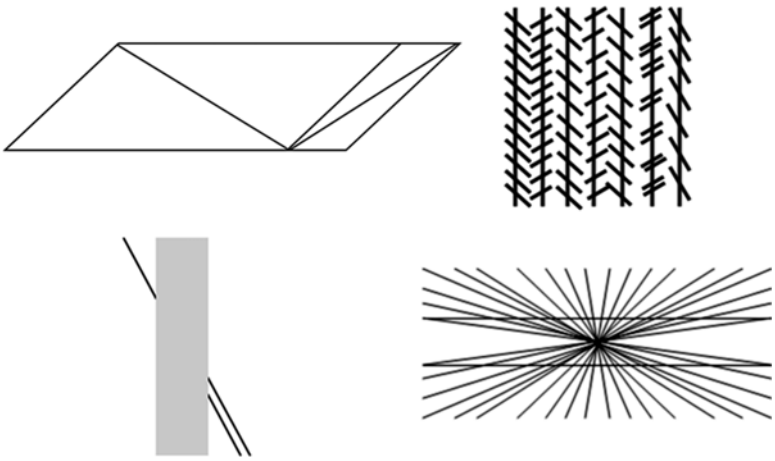

Fig. 7.9 In Sander's illusion (*top left*), the *diagonal lines* crossing the parallelograms are of the same length; in Poggendorff's illusion (*bottom left*), we are under the impression that, of the two segments on the *right of the gray rectangle*, it is the one on top that is in continuity with the one on the *left of the rectangle* (the reader should check); in Zollner's illusion (*top right*), the *vertical lines* are really parallel; and, similarly, in Hering's illusion (*bottom right*), the *horizontal lines* are really parallel

optical-geometric illusions, Piaget rather speaks in terms of primary illusions vs. secondary illusions. The fundamental property of a primary illusion, also called field effect, is that it does not vary qualitatively with age. However, their quantitative aspect, that is to say the strength of such an illusion, does vary with age. Also, Piaget does not say like Binet that the effect is innate. Secondary illusions are rather those arising from perceptual activities. These activities cause a decrease in some primary illusions and the emergence of new illusions.

Figure 7.9 reports a series of illusions based on angle effects. This category of illusions is very powerful. Among them, you will discover the spectacular Sander's illusion where the diagonal lines passing across parallelograms are surprisingly of the same length. Zollner's, Hering's, and Poggendorff's illusions are also based on angle effects.

Another example of angle effect, perhaps the best known, is the Müller-Lyer illusion (Fig. 7.10). This illusion could be explained by an assimilation effect or central tendency effect. According to this view, the EF and GH segments are taken into account in the estimation of segment AB (Fig. 7.10, right). Segments EF and GH being on average shorter than segments IJ and KL, it follows that the segment AB is perceived as being shorter than the segment CD.

For explaining the tendency to consider the segment AB to be shorter than segment CD (Fig. 7.10, right), some authors argue that these segments automatically generate the depth cues frequently observed on a daily basis (Fig. 7.11). Indeed, this illusion would be less pronounced with non-occidental populations less accustomed to the architecture made of many angles and squares as is often the cases in Western countries.

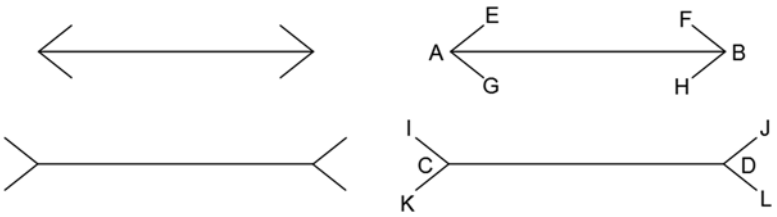

Fig. 7.10 The Müller-Lyer illusion (on the *left*), where the *horizontal line*, *bottom part*, seems longer than the *top left horizontal line*; letters on the illustration on the *right* serve an explanation reported in the text

Fig. 7.11 These cabinets contain clues reminding the Müller-Lyer illusion. The two *long vertical black lines* appear to be of equal length. However, the one on the *right* is shorter by about 15 %. Indeed, the two *vertical lines* are not placed in the same context. Even by adding the length of the wooden parts just below and just above the line on the *right*, this line remains shorter than the *black line* on the *left*

Other illusions are essentially based on perspective effects. A simple case is that of the Ponzo illusion (Fig. 7.12). Also, it is possible to create a variant of this illusion with the railroad track illustration used earlier and the addition on the picture of same size segments at different locations on the track. If the size of the segments is not adjusted for perspective, the highest segment appears longer, and the lowest segment appears shorter.

A spectacular case involving a perspective effect is that of the Ames room. This room is not square as would suggest our knowledge of what a normal room is. It rather has a side (the photo in Fig. 7.13) deeper and higher than the other. A major visual distortion can happen when looking at persons in such a room. If we pay attention at the relative size of the two persons, one on the left and the other on the right, and if we assume they are in a normal environment (in which one would think they are at the same distance from us as would spontaneously suggest normal size

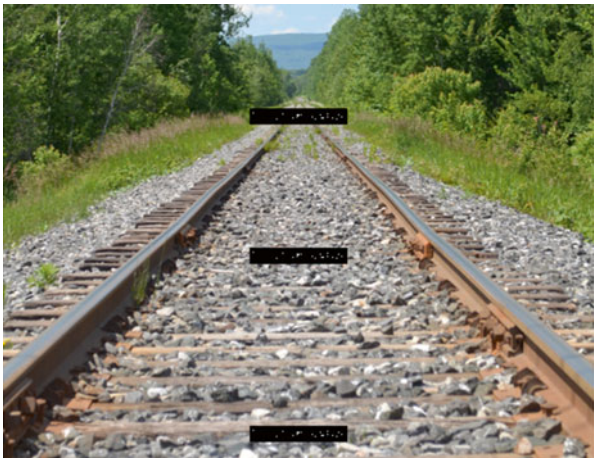

Fig. 7.12 Illustration of the Ponzo illusion

constancy mechanisms), the person on the right looks oversized compared to the person on the left. Also if one person was to move along the back wall, the size of that person would change: the person on the right would shrink if going left, and the one on the left would grow if going right.

There are other ways of generating strong illusions. One of them is to take images of different sizes close to each other and to compare them. Among the illusions of this kind, those of Delboeuf and of Titchener (Fig. 7.14) are noteworthy. Another classic illusion is the Oppel-Kundt: a segment divided into several parts is perceived to be longer than a segment of equal length but undivided (Fig. 7.15).

The length of a segment also depends on its orientation. Thus, a segment of a given length appears longer vertically than horizontally (Fig. 7.16). According to Künnapas, who wrote a series of articles about this illusion in the 1950s (see Prinzmetal & Gettleman, 1993), a frame effect is causing the illusion. Because the visual field is elliptical, a vertical segment is closer than a horizontal segment of the same length to the frame (i.e., closer to the ends of the visual field). Coren and Girgus (1978) rather hypothesized that the vertical appears longer because it involves a depth cue and the horizontal does not. If you are asked to indicate the midpoint of a vertical line, you will probably not indicate a location dividing the line in two equal parts, but a point located a bit higher than midpoint because higher means farther away (more distance).

7.3.2 The Moon Illusion

Because of its ubiquity and also because it has intrigued philosophers and scientists for a long time, the moon illusion deserves we spend some time on it. This illusion is so strong that we forget or even doubt that it is an illusion. This illusion is even more interesting that a plausible explanation requires the perfect integration of the

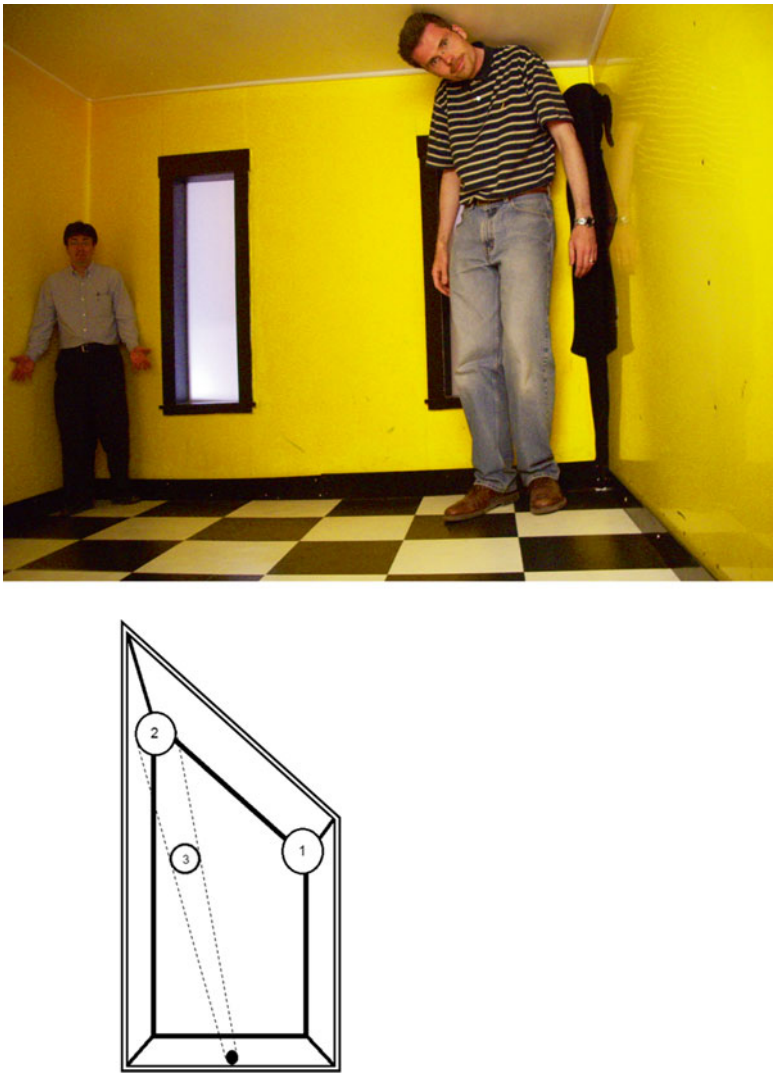

Fig. 7.13 The effect generated when two people are in the Ames room (photo on *top*). Below, a bird's-eye view of the room. If an observer (*black dot*) believes that *person 2* is at *position 3*, i.e., at the same distance as *person 1*, as suggested by the depth cues in the room, then the observer will perceive *person 2* as much smaller because the retinal size of the latter is much smaller than that of *person 1*

fundamental notion of perceptual constancy. But what is the moon illusion? This illusion refers to the fact that the moon appears larger when it is on the horizon than when it is at its highest point in the sky (the zenith). This difference is estimated at approximately 30 %, but may be sometimes smaller, sometimes much greater, depending on the exact conditions of testing and on the observers.

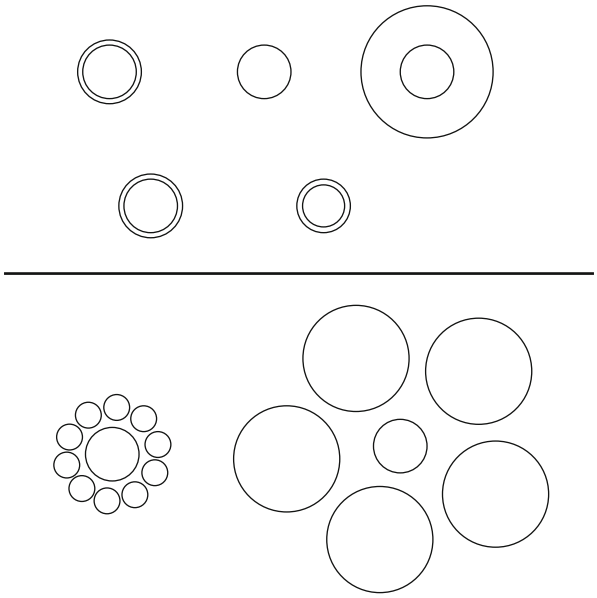

Fig. 7.14 Illustration of the Delboeuf (*top*) and Titchener (*below*) illusions

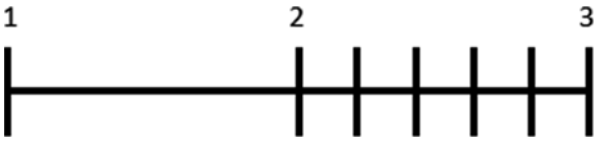

Fig. 7.15 Illustration of the Oppel-Kundt illusion. The distances between *1* and *2* and between *2* and *3* are the same although the distance between *2* and *3* seems larger

Fig. 7.16 Illustrations of the *horizontal-vertical* illusion. Is the wizard hat (*left*) wider than is tall or taller than is wide? Or does height and width seem almost equal? Just measure it! On the *right*, do *horizontal* and *vertical lines* have equal length?

Since the distance that separates us from the moon remains pretty much equivalent regardless of where it is located, its projective size remains the same, whether at the zenith or horizon. According to Irvin Rock and Lloyd Kaufman (Kaufman & Rock, 1962; Rock & Kaufman, 1962), the illusion is not caused by the different angles of the observer's gaze, as some researchers believed until then, but to the

presence or absence of objects (the ground) between the observer and the moon. Rock and Kaufman rather emphasized the importance of apparent distance, also referred to as the size-distance invariance hypothesis.

In order to understand this explanation, it is crucial to remember the idea of perceptual constancy: if two objects have the same retinal size, the one that appears farther away is perceived as larger. What would happen would the brain believe that the moon is farther when on the horizon than at its zenith? Because we know that the retinal image is the same in both cases, we must conclude that the brain would interpret that the moon is larger on the horizon. In other words, we believe that the moon is very large on the horizon because our brain believes it is far away. This may seem counterintuitive for someone concluding that the moon appears to be so close because it looks so big. A full understanding requires that you keep in mind the fact we are dealing with perceptual mechanisms engaged automatically or unconsciously by the brain.

The critical question at this point becomes the following one: are there at least reasons to believe that the moon seems farther away on the horizon than at the zenith? The answer is yes, according to Kaufman and Rock. Consider the following experiment where observers were asked to point out the midpoint between the zenith, 90°, and the horizon, 0°. Rather than pointing the midpoint, which is 45°, these observers rather tended to point a direction a little closer to the horizon than to the zenith. As shown in Fig. 7.17, observers do not point to the midpoint of a sky that would be perceived as semicircular; they point rather to what is midpoint of a sky perceived as being flattened. If the sky is perceived as being flat, the moon is necessarily perceived as more distant when on the horizon than at the zenith.

There is a second reason to believe that the moon seems farther away on the horizon than at the zenith. It is recognized that the perceptual system is sensitive to

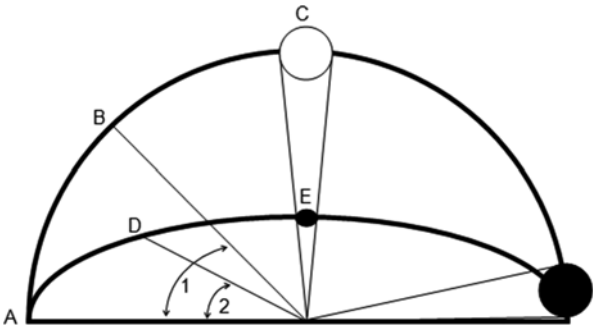

Fig. 7.17 When asking observers to indicate midpoint between the horizon and zenith, they do not indicate Point *B* (a 45° angle; see digit *1* in the figure); they point rather in the direction of Point *D* (angle *2*). Point *B* is midpoint between *A* and *C*, *C* (*large white disk*) indicating the real location of the moon, *D* being midpoint between *A* and *E*, and *E* (the *small black dot*) being where the observer believes the moon is. To an observer, indicating midpoint corresponds to pointing *D*, assuming that the sky is perceived as being flat rather than *semicircular*. The moon is therefore judged as being closer (Point *E*) when standing at the zenith than when located at the horizon (the *large black disk*) (Kaufman & Rock, 1962)

the presence of benchmarks in the environment. With more landmarks ahead of us, we tend to perceive distances as larger. When looking at the moon at its zenith, there are no landmarks to guide the estimation of distance; however, most often the land offers several landmarks such as trees, cars, or houses. These landmarks help give the brain the impression that the moon on the horizon is far away from us.

In short, the moon would be perceived as being larger on the horizon than at the zenith because the brain would believe it is farther on the horizon. This explanation makes sense only if one understands the idea of perceptual constancy, that is, the principle stating that perceived distance and projective size are closely connected when estimating the size of objects. Many explanations and descriptions about the moon illusion can be found in Hershenson (1989) or Ross and Plug (2002):

> The fact that landmarks contribute to perceiving depth led to a basic rule of water safety. If you capsize a boat after moving far from the shore of a lake, be careful before deciding to swim back rather than trying to grab the boat. Because there are usually no benchmarks in the water (sometimes an island, sometimes other boats), you might get the incorrect impression of still being close to shore. Inadequate assessment of distance may cause exhaustion before reaching the shore.