Perceptual Organization: Auditory

See Auditory Scene Analysis

Perceptual Organization: Vision

Perceptual organization is part of the process by which sensory data collected by our receptors is combined and integrated into structured, organized percepts. The eyes signal information about the light around us, but humans do not consciously experience these separate pieces of information—we don't see countless patches of color from the wavelengths registered by our cones. Instead, we see whole objects and surfaces smoothly integrated into scenes. These wholes are the products of perceptual organization. Sometimes these integrations contain surprises, where the whole pattern may be strikingly different from the sum of its parts: an artist draws a few curves on the canvas and a face emerges; or a set of black dots comprising a newspaper photograph combine into a rich scene; or three black disks with wedged-shaped notches yield the perception of a white triangle (see *Perceptual Segregation,* Figure 2). Such novel wholes are called *gestalts*. This entry reviews what perceptual organization does, how it does it, and its practical implications.

A common view sees visual perception beginning with the registration and transduction of light within the retina into neural signals, a process that takes place without our conscious awareness. This is followed by the organization and interpretation of those neural signals into structured, conscious percepts. These processes are sometimes called *sensation* and *perception,* respectively. There seems to be no clear border between them, however; rather, sensation flows smoothly into perception. The latter explains why we usually see whole scenes rather than swirls of megapixels or innumerable separate pieces of a jigsaw puzzle. Determining how we perceive wholes rather than just local parts has proven to be a challenge.

Perceptual organization is an important part of that second process in vision because it deals with how those elementary pieces, sometimes called basic features, are assembled into organized wholes. A commonly held view is that while the early phase of vision proceeds automatically and in parallel (simultaneously) across the visual field, the integration of the component features requires attention and takes place only within a limited region of space at any one time, within a spotlight of attention. If a red ball is tossed our way, one part of our visual system processes the wavelengths leading to red, another detects the curved edges defining a ball, another processes the motion of the ball toward us, and yet others detect the ball's size, texture, and spin. These components are subsequently combined into final percepts through attention, a process called feature binding.

Gestalts

Perceptual organization specifies a very different way parts combine into integrated configurations and how those parts and wholes can be identified. The structuralists, such as Edward Titchener (1867–1927), held that the combination was additive: When the parts are perceived, their sum defines the whole. In this manner, the perception of a dog equals the sum of the percepts of its parts: four legs, a torso, one tail, a head with eyes, ears, nose, and mouth, and fur of a certain length and color. Gestalt psychologists rejected this notion of additivity, claiming that "the whole is different from the sum of its parts" (often misquoted as "more than the sum"). They noted that wholes often possess configural qualities or patterns—gestalts—and that the relationship among the parts was often more important than the parts themselves (e.g., a melody is defined by the relationship among its notes and can thus be played in any key).

The Gestalt movement was launched in the early 1900s with the study of apparent motion, an illusion arising when two stationary lights are flickered alternately, as in movie marquees or in roadside warning signals. Gestaltist researchers produced demonstrations showing that when parts or elements are combined in certain ways, something novel emerges from their configuration, often something surprising. Similarly, they demonstrated that a given part may assume one appearance in one configural context but another in a different context.

Consider the Gestaltists' apparent motion stimulus. We start with a flashing light. Then we introduce a second light flashing in the opposite rhythm, such that whenever one light is illuminated the other is extinguished. When looking at both lights, we see something different from the sum of what we saw from the two lights individually: We now perceive a single, nonflashing light moving back and forth between the two locations. (Seeing just one light with no flashing illustrates how the whole can be less than the sum of its parts.) Here, motion is an emergent feature, something categorically different and surprising because it is a property of neither light individually. The whole we perceive—the configuration or gestalt—differs from the sum of its parts.

Many essential issues of perceptual organization are apparent in Figure 1(a), which shows an old, weathered photograph that is hard to interpret. We can sense the image well enough, in that we can report the gray level of any spot in the picture; we could even duplicate it with a paintbrush, but we might still fail to recognize the familiar object it depicts.

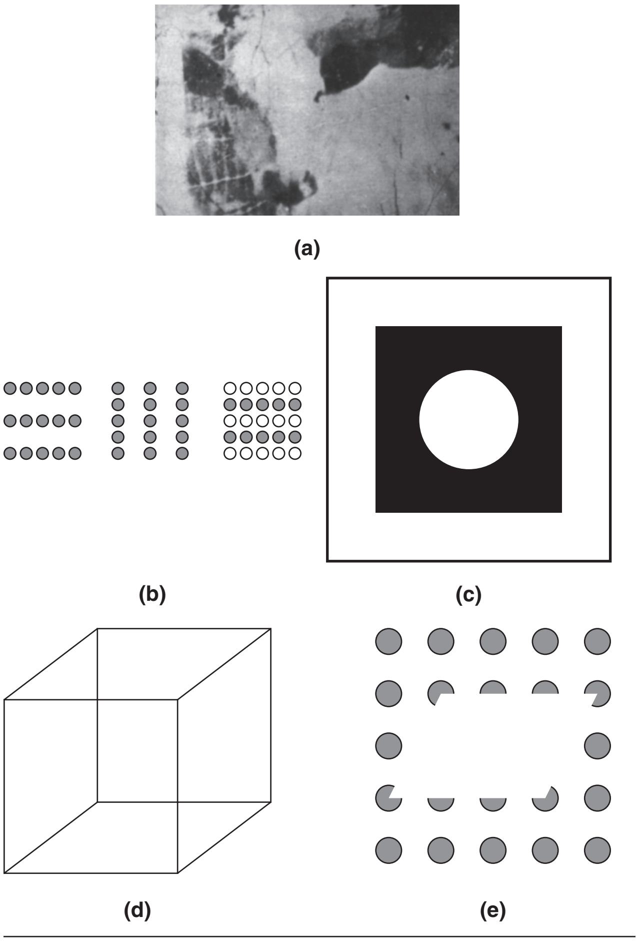

Figure 1

Notes: (a) What's in this image? (b) Grouping by proximity and similarity (c) A circle on top of a square, or a square with a circular hole? (d) Necker Cube (e) Do we see the illusory parallelogram because it is simpler or more probable?

Four Key Phenomena

Perceptual organization focuses on four key phenomena:

1. *Grouping and part-whole relationships:* determining which regions of an image go with which others to form unitary objects.
2. *Figure-ground segregation:* determining which regions represent opaque objects blocking our view of ("occluding") other, more distant objects; and which side of an edge is the figure side and which side belongs to the ground continuing behind.
3. *Perceptual coupling:* determining the appropriate relationship between two linked dimensions in the image. As an object moves away from us, the image it projects to our eye shrinks until it has vanished. If a medium-sized image strikes our retinas, did it come from a large object at a great distance, a small object at a short distance, or an intermediate-sized object at a moderate distance?
4. *Multistability (bistable perception):* some stimuli may be perceived equally correctly in two different ways. Interestingly, our visual system often alternates spontaneously between possible interpretations, abruptly and unrelentingly flipping as though the stimulus were changing.

Let us examine these four phenomena in greater detail, with examples.

Grouping and Part-Whole Relationships

Today's best high-definition televisions have a resolution of about 2,073,600 pixels. The picture comprises a 1920 × 1080 grid, with each cell containing a number indicating how bright the picture is at that spot. (Even better, think of the picture as a stack of three grids or spreadsheets, representing the three color channels). The same is true for a newspaper picture: a simple grid of closely spaced dots varying in size. You may think of the image leaving the eye for the brain similarly—a grid of neurally coded intensity values, ready to be recognized as objects in the cortex. The challenge of grouping is to specify which pixels go with which others to form those objects. Neither your television set nor your digital camera solves this problem; they have no idea what is in the image they are transmitting! Humans, however, perform grouping continuously, accurately, and effortlessly whenever our eyes are open.

To see how to solve this problem, think about solving jigsaw puzzles. Faced with a jumble of pieces, you search for ones with similar colors and for edges continuing smoothly from one piece into another. Research indicates our visual system works similarly, using principles to determine grouping. Some common grouping principles are:

- *Proximity:* the closer together any two elements are in an image, the more likely they belong to the same object (see Figure 1b).
- *Similarity:* the more alike any two elements are (more similar in color, size, orientation, distance, etc.), the more likely they belong to the same object (Figure 1b).
- *Common fate:* the more similarly any two elements change over time (e.g., in their pattern of motion) the more likely they belong together.
- *Good continuation:* the more smoothly one edge or contour blends into another one, the more likely they are parts of a single contour.
- *Closure/convexity:* when connecting contours into objects, curves that can be assembled into closed or convex objects are more likely to belong together than ones that cannot.
- *Common region:* any two elements that are contained within a common region (e.g., encircled by a single contour) are more likely to belong to the same object.
- *Connectedness:* any two elements that are physically connected to one another are more likely to be parts of the same object than two that are not.

These principles operate like rules of thumb or guidelines rather than laws, allowing us to connect picture elements appropriately and to segment (parse) an image correctly. Consider looking out at a sea of faces in a crowd. If we could not group correctly the various facial features that belong to a single person—the eyes, ears, nose, and mouth—we might end up trying to recognize an individual by combining the right eye of one person, the left eye of another, the nose of a third, and the mouth of a fourth. Instead, our visual system may use principles like proximity, similarity, common area, and connectedness to group the features correctly, all without our conscious awareness of these principles. When a long snake slithers through the grass and under a log, we can recognize that snake as a single object because its components are close together, because the coloration across its body is relatively uniform and different from the background, because the contours of its body continue smoothly without sharp angles, and because the whole animal moves together.

Figure-Ground Segregation

Much of the information in an image is conveyed by the edges it contains. Line drawings and cartoons depict scenes using only those edges. Looking around, we see edges everywhere—the edges of our desk and chair, of our arms, hands, and fingers, edges of the room we are in and of the trees outside the window. When our visual system encounters an edge, it has some figuring to do, including whether the edge comes from one object occluding our view of another—as when you put your hand across the sleeve of your shirt. (Other edges arise when two objects merely abut, as with tiles; or when two surfaces come together, as with a corner in a room, or from shading.) With occlusion edges, the visual system tries to determine which side of the edge belongs to the nearer, occluding object (i.e., the figure) and which does not. When your hand covers your sleeve, it is important that the brain correctly assign the edge where they meet to your hand and not to your sleeve, which continues unseen behind your hand and whose own edges have nothing to do with the edges your hand leaves on it.

Imagine you see a circular black patch on an otherwise white wall. Is the round region a black disk that has been taped or painted onto the wall, or is it a hole passing through the wall into the dark? See Figure 1(c) for a similar illustration of the puzzle of figure-ground segregation. If we make the assignment incorrectly, we may fail to recognize objects, as when one looks at a map of Europe and mistakes seas and oceans for land. Escher produced many drawings playing with these phenomena, and the Federal Express® logo famously contains a figure that is hidden until you reverse your figure-ground assignments.

As with grouping, several principles help us make the correct assignments:

- *Convexity:* the convex side of a region is most likely the figure. This helps explain why a dark circle on a wall is more often seen as a painted circle than as a dark hole.
- *Size:* the smaller region is more likely to be the figure. This too helps with the circle.
- *Motion:* the moving region is more likely to be the figure.
- *Symmetry:* the symmetric or "good" region becomes the figure.
- *Distance:* the nearer region becomes the figure.

Perceptual Coupling

The separate components or features of a stimulus are often coupled physically, such as the size of the image an object projects to the retina and its distance from the perceiver. Our thumbnail at arm's length may easily cover the moon, but we are not fooled into thinking the two are the same actual size. These two dimensions are coupled in our perception too, as is well demonstrated in Emmert's law. Suppose we create an afterimage by staring at a bright square and then look at a nearby brick wall. We see the afterimage of that square on the wall, as though it were really there. If we now look at a more distant wall, we again see the square, but now it appears larger, in that it covers more bricks. If we look at our hand, the afterimage will appear to shrink to fit in our hand. Similarly, if we look at a wall that is oblique to our line of sight, the square afterimage becomes trapezoidal, demonstrating the perceptual coupling now between perceived shape and perceived orientation. These demonstrations reveal how the various components and dimensions of a whole stimulus are organized and linked, rather than perceived independently.

Multistability

Many stimuli entering the eye may be interpreted in multiple ways: an ellipse can also be perceived as a circle viewed from an angle, a diamond may be seen as a tilted square, and a triangle may be seen pointing in any of three directions. For some stimuli, rather than sticking with one interpretation, the visual system cycles though different ones. Despite our conscious knowledge that the stimulus is unchanging, our perception of it flip-flops, creating a rare, peculiar, and often amusing experience. The most famous demonstration of multistability comes from the Necker cube (Figure 1d), a simple line drawing of a wire cube whose perceived orientation reverses nearly uncontrollably every few seconds as it is observed. Multistability arises too in motion perception: The red and blue stripes of a barber pole seem to be moving upward as the pole rotates, despite the fact that they are moving horizontally. If the pole is made short and wide, however, the stripes now appear as they should, moving horizontally. If the pole is made just as wide as it is tall, however, perceivers can now see motion in a variety of directions, with our perception shifting incessantly.

All four of these phenomena have implications for how we perceive our visual world. Some lead to illusions; for example, two identical grey surfaces seem to have different shades if we alter our perception of the surfaces' orientation toward a light source. Similarly, when two regions of a stimulus are perceptually grouped, it is difficult for us to attend selectively to just one; and it is easier for us to make judgments of two parts of a single object than of one part each of two objects.

Theoretical Perspectives

Some theoretical notions have been advanced to account for the four phenomena of perceptual organization. One is the claim that gestalts or wholes are primary in perception, with information about parts being secondary. The claim is that we see the forest before the trees, or that we perceive and remember a melody without noticing the particular notes or the key in which it is played. There is good evidence that with some wholes this is what happens (e.g., people are better at discriminating arrows from triangles than they are at discriminating the difference in orientation of a single line segment that changes an arrow into a triangle). People—and animals—are often better processing holistic properties of stimuli (such as closure) than more primitive, localized properties.

Historically, there have been two overarching hypotheses about why our percepts are organized as they are. The first, generated by the Gestaltists, is the Prägnanz principle, which holds that we organize our percepts in the simplest way possible, much as a soap bubble assumes the simplest possible shape—a sphere. Thus, we perceive an ellipse as a circle at an angle because a circle is simpler than an ellipse (having only one parameter, its diameter, compared with an ellipse's two). The second hypothesis is the likelihood principle, espoused by von Helmholtz, which holds that we organize our percepts in the way that most likely reflects the objects in our world. Thus, we perceive trapezoids as rectangles viewed from an angle because rectangles are so much more common in our environment than trapezoids (think about rectangular doors, buildings, signs, containers, etc.). See Figure 1(e) for another example. A possible reconciliation between these hypotheses notes that simple structures are also more common in nature than complex ones.

Practical Implications

Perceptual organization has ramifications for life outside the lab. An example is camouflage, as found in the animal kingdom with protective coloring and in humans within the military. To disguise ourselves to prevent detection by others, we try to group with our background so we don't stand out as figure on ground. We do this by matching our color and texture characteristics to our surroundings and by freezing our motion, lest we group as a unit by common fate. Perceptual organization also figures prominently into the design of instrument panels and computer interfaces. Here, designers try to group controls for similar functions by placing them close together (proximity), by coloring them the same (similarly), and by enclosing them within the same boundary (common region)—as with windows on computer screens. Similarly, we try to draw attention to controls by making their regions stand out as figures, by giving them unique colors, by making them bright and flashing, and by making them move. Perhaps most importantly, perceptual organization helps us perceive the world both accurately and quickly—two virtues that bring great rewards in competitive or dangerous environments.

James R. Pomerantz
and Mary C. Portillo

See also Bistable Perception; Constancy; Context Effects in Perception; Feature Integration Theory; Gestalt Approach; Object Perception; Perceptual Segregation

Further Readings

Garner, W. R. (1974). *The processing of information and structure*. Mahwah, NJ: Lawrence Erlbaum.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kimchi, R., Behrmann, M., & Olson, C. (Eds.). (2003). *Perceptual organization in vision: Behavioral and neural perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Köhler, W. (1947). *Gestalt psychology*. New York: Liveright. (Original work published 1929)

Kubovy, M., & Pomerantz, J. R. (Eds.). (1981). *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge: Bradford Books/MIT Press.

Perceptual Representation (Philosophy)

Philosophers and psychologists talk about perceptual states as *representations*. One of the most important questions in the philosophy of perception is whether and in what sense could perceptual states be considered representations. A further crucial question is in what way perceptual representations are different from nonperceptual mental representations, such as beliefs or thoughts. These questions about representations will be discussed in this entry.

Representations

There are various kinds of representation. Leonardo da Vinci's *Mona Lisa* represents a woman: It is a pictorial representation. If you are thinking about Paris, your thought represents Paris: It is a mental representation. Representations refer to things that may be far away (like Paris) or that may not even exist (or, does not exist any more, like the sitter of *Mona Lisa*). They refer to something, they are about something, and what they are about is the object of the representation. The same object could be represented in different ways. The represented