

Object Perception

Scott P. Johnson

Subject: Cognitive Psychology/Neuroscience, Developmental Psychology

Online Publication Date: Jul 2018 DOI: 10.1093/acrefore/9780190236557.013.62

Summary and Keywords

Visual scenes tend to be very complex: a multitude of overlapping surfaces varying in shape, color, texture, and depth relative to the observer. Yet most observers effortlessly perceive that the visual environment is composed of distinct objects, laid out across space, each with a particular shape that can be inferred from partial views and incomplete information. Moreover, observers generally expect objects to be continuous across space and time, to have a certain shape, and to be solid in three-dimensional (3D) space. The cortical visual system processes information for objects first by coding visual features, then by linking features into units, and last by interpretation of units as objects that may be recognizable or otherwise relevant to the observer. This way of conceptualizing object perception maps roughly onto processes of lower-, middle-, and higher-level visual processing that have long formed the basis for investigations of visual perception in adults, as well as theories of object perception, the ways visual deprivation reduces object perception skills, and the developmental time course of object perception in infancy.

Keywords: object perception, visual perception, theories of object perception, cortical visual system, critical periods, visual development

Introduction

When observers encounter a visual scene, they quickly form an impression of its contents and they make moment-to-moment, context-appropriate decisions about appropriate actions. Vision works in concert with other sensory systems (audition, proprioception, taste, and smell) to impart coherent interpretations of the identities, locations, and movements of objects and people in our surroundings. Visual scenes tend to be very complex: a multitude of overlapping and adjacent surfaces with distinct shapes, colors, textures, and depths relative to the observer. The input to the visual system and the subjective experience of the visual world, however, are quite different. Visual input under most circumstances is continuous and unbroken across the retina; there are no gaps in the input, and few photoreceptors go unstimulated as the eyes take in visual information. The subjective experience of the visual environment, in contrast, is one of largely empty space interspersed with objects at various distances. Subjective experience is at odds with visual input in a second way. Although objects are generally experienced as having a regular, solid

Object Perception

shape, most objects often cannot be seen in their entirety because of occlusion—occlusion of far objects by nearer ones, and self-occlusion of the far sides of individual objects due to opacity.

Yet observers' typical visual experience is not one of incomplete fragments of surfaces, but instead one of objects, most of which have a shape that can be inferred from partial views and incomplete information. In everyday settings, observers hold certain common-sense expectations about the objects they see. Most objects, for example, can be expected to be continuous across space and time despite gaps in perception due to occlusion, and are perceived as separate from neighboring objects. Observers expect objects to have a certain shape and to have a coherent structure, solid in three-dimensional (3D) space. Observers classify objects into categories according to appearance or function, and they recognize familiar objects and distinguish them from new ones that they encounter. Motion of observers and of objects provides additional information to determine the contents of the surroundings. Humans can move through the environment, obtain new perspectives, and see parts of objects invisible from previous vantage points. As objects move, observers can track them across periods of temporary invisibility, often predicting their reappearance. Finally, actions are planned around these expectations: Objects might be either avoided or approached as they and the observer move about, depending on the observer's goals.

The visual system generally provides fast and highly accurate information about near and distant objects in our surroundings, because object perception is the *raison d'être* of visual perception. Object perception may be taken for granted because it seems so effortless and seamless, yet it is remarkable to consider the mechanisms underlying it: intricate cortical machinery comprising several dozen areas of the brain, each responsible for processing a distinct aspect of the visual scene or coordinating the outputs of other areas (Zeki, 1993), and elaborate action systems comprising the eyes, head, and body, each with independent control systems, working in tandem to explore the visual environment (Gibson, 1950).

Several steps are involved in the processing of visual input that leads to the subjective experience of objects. First, input from a visual scene reaching different parts of the retina (the first stage of visual processing) must be coded according to variations in color, luminance, motion, texture, pattern, shape, orientation, and distance. Next, the outputs of these processes must be recombined into structured units—the building blocks of objects—and, as appropriate, units must be perceived as complete across space and time despite gaps in perception. The gaps may be due to occlusion, to movement of the observer, or to movement in the environment. The process of filling in the gaps is known as perceptual completion, and it includes deduction of 3D shape from limited views due to self-occlusion. Next, higher-order visual processing is performed as necessary, such as recognition of objects, categorization, tracking identity of objects over time, and planning relevant actions based on perceived affordances of objects and the needs of the observer.

Object Perception

Object perception, therefore, rests on a foundation of initial coding of visual features, followed in succession by linking of features into units, and finally by interpretation of units as objects that may be recognizable or otherwise relevant to the observer. This way of conceptualizing object perception maps roughly onto processes of lower-, middle-, and higher-level visual processing that have long formed the basis for investigations of visual perception in adults (e.g., Marr, 1982; Palmer, 1999), as well as theories of object perception. The remainder of this article covers theories of object perception, the neural bases of object perception, and finally visual development, including critical periods for visual functions and the developmental time course of different aspects of object perception.

Theories of Object Perception

Koffka (1935) asked, “Why do things look as they do?” This question has motivated a number of theories of visual perception and the theories provide distinct answers, depending on their underlying assumptions and the evidence they bring to bear. One of the earliest attempts to understand object perception had its origins in the 19th-century theory of structuralism, espoused by Wilhelm Wundt in Germany and Edward Titchener in the United States, which was largely consistent with the views of the British empiricists, such as George Berkeley, David Hume, and John Locke (see Palmer, 1999). The theory of structuralism held that perception arises from assembly of sensory primitives in a given sense modality, through a process of repeated associations of the primitives in time and space. The associations are presumably formed early in life from exposure to structured objects and events (a concept that is discussed further in this article).

In opposition to structuralist theory, theorists in the Gestalt tradition operating in the early to mid-20th century, such as Wolfgang Köhler and Max Wertheimer (as well as Koffka) from Germany, argued that structure cannot be reduced to the sum of the parts (Koffka, 1935; Palmer, 1999). Rather, many configurations, such as illusory figures, have emergent properties that are inherently holistic. Perceptual experience was proposed to correspond to the simplest and most regular interpretation of a particular visual array, consonant with a general “minimum principle,” or *Prägnanz* (Koffka, 1935). When confronted with a scene in which a palm tree is seen on a beach with the shoreline behind, for example, an adult observer will usually report perception of a continuous shoreline, despite the shoreline’s partial occlusion by the tree. The determination of continuity can be made on the basis of the alignment of the shore’s edges to the left and right (the Gestalt principle of good continuation), the resemblance of the visible portions of the shore’s surface (symmetry and similarity), the regularity and simplicity of the shoreline in general (good form), and the common motion of waves visible on either side of the tree (common fate). (And, of course, it is highly unlikely that two different shorelines would line up precisely.) Shapes that are defined by such principles are more coherent, regular, and simple than disconnected and disorganized forms. The minimum principle and *Prägnanz* were thought to arise from a tendency of neural activity toward minimum work and minimum energy

Object Perception

(analogous to other physical systems), which drive the visual system toward simplicity (Koffka, 1935).

Because this predisposition is inherent in the visual system, according to the Gestalt view, it follows that infants and children should experience the visual array in ways similar to adults. Some researchers suggested that perceptual experience is never disorganized: An organized world could not arise solely from experience because experience cannot operate over inherently disorganized inputs (Zuckerman & Rock, 1957). Necessarily, therefore, the starting point of visual organization is inherently organized, and in this respect Gestalt theory was consistent with the views of rationalist and nativist schools of thought, exemplified by such philosophers as René Descartes and Immanuel Kant (in his early writings). Holistic perception necessarily arises from underlying holistic processes, so goes the argument, and must originate in the intrinsic structure of sensory systems and neural circuits in the brain. From this standpoint, therefore, intrinsic sensory and cortical structures in the visual system are responsible for observers' typical perception of coherent objects.

These views were challenged by additional advances in the 20th century. Four advances are particularly important for understanding theories of object perception. The first advance was information-processing theory, rooted in advances in computer technology, in particular the invention of devices that could be programmed with algorithms to carry out a variety of procedures based on the inputs that were provided. The second advance was the theory of "ecological optics" espoused by Gibson (1979). Gibson suggested that perception is best understood by examining the structure of the perceiver's environment—for example, the information in light reflected from objects as it is received by the organism. A central idea in this account is that mobile organisms are able to exploit visual information to maximum effect because motion and change provide important information for perception: The eyes rotate within the head, which moves relative to a body, which perambulates and explores the world. Moreover, motion of objects and events in the environment provide vital information about object properties, segregation, distance, and coherence. The third advance was constructivism—advocated by the psychologists Richard Gregory in the United Kingdom and Julian Hochberg and Irvin Rock in the United States—a theory about the mechanisms of perception that extract information from the environment and, importantly, fill in the missing pieces via processes of inference. In the constructivist view, perception of holistic structure from relatively underspecified input (as in the case of occlusion) implies a set of heuristics by which optical information leads to subjective experience, such as the "likelihood principle," a probabilistic computation concerning which interpretation of a given scene is most likely given current retinal input and past experience (nowadays known as the Bayesian approach). The fourth advance was Piaget's (1954) theory of cognitive development (itself a constructivist theory), which emphasized the contributions of action systems to cognition, especially the child's recognition of her own body as an independent object and her own movements as movements of objects through space, akin to movements of other objects she sees. Piaget proposed that, prior to the advent of coordinated visual and manual action skills in infancy, the visual environment is essentially a "sensory tableau" in which images without permanence

Object Perception

or substance shift erratically and capriciously; objects, as adults understand them, do not yet exist. Veridical object perception, therefore, was thought to be an outcome of coordination of perception and action systems.

The focus on (a) the information in the external environment, (b) the uptake of available information, (c) the mechanisms by which the observer receives and interprets the information, and (d) the developmental origins of these mechanisms provides clarity with respect to understanding object perception. The task of the observer is to use his or her perceptual systems—vision, hearing, touch, and so forth—to explore the world and to obtain information about its properties. The information must be attended to, encoded, stored, retrieved, and acted upon.

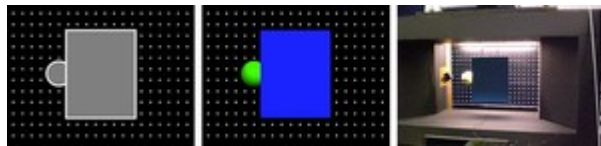


Figure 1. Levels of representation in Marr's (1982) theory of object perception. Left: A primal sketch representing edges and contrast. Center: A 2.5-D sketch representing distinct surfaces and relative depth relations. Right: Objects in 3D space.

The simple accrual of associations within a single sensory modality, as proposed by the structuralists, seems inadequate to account for the complex nature of object perception. However, our current state of knowledge benefits from more recent theories of object perception that revisit some of the other themes pointed out by Gestalt and ecological psychologists. For example, the theory of “visual interpolation” (Kellman & Shipley, 1991) builds on Gestalt theory by formalizing the “unit formation” process underlying perceptual completion, with the goal of understanding conditions that lead observers to perceive edge connectedness despite incomplete sensory information (gaps in space and time). Marr's (1982) theory of vision as an information-processing system argued that cognition could be studied at three independent levels: the computational level, specifying computations needed to solve a particular problem (e.g., object perception presupposes figure/ground segmentation); the algorithmic level, specifying mechanisms by which computations are carried out; and the implementational level, specifying how the algorithm is accomplished in a physical structure (a neural network or brain). Marr also proposed a theory of object perception in which visual function proceeds in three stages: an initial “primal sketch” capturing fundamental visual properties of a scene, such as edges and contrast; a second “2.5-D sketch” delineating distinct surfaces; and a final representation of 3D objects laid out in space (see Figure 1). The theory of “embodied cognition” holds that object perception and other cognitive processes are rooted in the body's interactions with the environment (Wilson, 2002), a position that echoes the theory of ecological optics (Gibson, 1979) as well as Piaget's (1954) view that object perception is built from reciprocal processes of perception and action. Finally, Bayesian theories of object perception propose that the visual system integrates information from prior knowledge and current

Object Perception

inputs in a probabilistic fashion to achieve inferences that guide visual attention and bind features into coherent structures (Kersten, Mamassian, & Yuille, 2004). The possibility that object perception develops in part from processes reliant on experience, such as statistical learning, provides a mechanistic account consistent with the constructivist theories of cognition and perception mentioned previously.

Neural Bases of Object Perception

The importance of accurate visual perception of our surroundings is manifested by the allotment of cortical tissue devoted to vision: By some estimates, over 50% of cortex in the macaque monkey (closely phylogenetically related to humans) is involved in visual perception, and there are over 30 anatomically and physiologically distinct cortical areas that participate in visual or visuomotor processing (Felleman & Van Essen, 1991; Van Essen, Glasser, Dierker, & Harrell, 2011). The visual system, like the rest of the brain, is organized hierarchically. Its purpose is to transduce light reflected from surfaces in the environment into neural signals that are relayed to the brain for processing and decision-making. Light is first transmitted through the cornea, the outer protective covering of the eye, and then the lens, which helps to focus reflected light onto the retina, the thin film of tissue covering the back of the eyeball. The retina is composed of layers of photoreceptors as well as a rich network of connections, nonsensory neurons, and supporting tissues that provide initial processing of visual information. Different kinds of photoreceptor accomplish different tasks: There are specialized cells and circuits in the retina for color and contrast, for example, and they help determine how information is subsequently routed to appropriate channels up the visual hierarchy in the brain.

Neural signals from the retina are routed to a midbrain structure called the lateral geniculate nucleus and then to the primary cortical visual area. Successively higher visual areas are specialized for visual attributes in larger portions of the visual field and participate in more complex visual functions (Mishkin, Ungerleider, & Macko, 1983). Reciprocal connections carry information to secondary visual areas (e.g., V2, V3, V4, and the medial temporal area, or MT), which participate in processing of color, contrast, motion, and other low-level visual attributes. Primary visual cortex is the origin of two multisynaptic corticocortical pathways. These pathways diverge into two partly segregated, yet interconnected streams (Goodale & Milner, 1992; Milner & Goodale, 2008; Mishkin et al., 1983). The first, known as the ventral stream, connects to temporal cortex. This pathway is specialized for object recognition, which is localized to an area known as the inferotemporal cortex (IT; Tanaka, 1997). The IT projects to the perirhinal cortex and other areas involved in categorization of visual stimuli and formation of visual memories (e.g., entorhinal cortex and hippocampus) as well as a part of the frontal lobe, the lateral prefrontal cortex, which is involved in learning contingent relations among stimuli and in action planning (Buschman & Miller, 2014; Miyashita & Hayashi, 2000). The second visual stream, the dorsal pathway, connects primary and secondary visual areas to parietal areas and codes information about object location and object-oriented action. The posterior parietal region is particularly important for voluntary action planning and the coordina-

Object Perception

tion of somatosensory, proprioceptive, and visual inputs. Parietal cortex also has reciprocal connections to and from the IT and prefrontal cortex. The IT is thus richly interconnected both with lower-level areas responsible for feature analysis and with higher-level areas responsible for object memory and behavior, and it is sometimes referred to as an association cortex. IT, therefore, is a central locus of object-oriented cortical activity.

Perceptual completion may be accomplished in part with relatively low-level mechanisms in cortical areas V1 and V2, which code for edge connectedness and send signals to ventral locations, including the IT. Connections between individual visual neurons allow information about edge orientation and motion to be passed to neighboring neurons, and these cell-to-cell activations are strongest within and across cell groups that code similar orientations (Roelfsema & Singer, 1998). The networks of neurons that respond preferentially to each orientation are characterized by long-range connections, the extended growth of axons and dendrites of individual cells. Long-range interactions may extend across several millimeters of cortex and can provide information about edge connectedness across a span of at least several degrees of visual field, even across a spatial gap (Heydt, Peterhans, & Baumgartner, 1984; Kellman & Shipley, 1991). The spreading of activation across networks occurs in part via cooperative responses across neurons; they fire in bursts of synchronized oscillatory activity (Singer & Gray, 1995). Neighboring cells that code for similar orientations, then, are connected both by virtue of their intrinsic wiring patterns and by their firing in a coordinated, organized fashion. This scheme is very effective at detecting connectedness, so much so that an area known as the lateral occipital complex, which straddles secondary visual areas and the IT in humans and contributes to segmentation regions of visual scenes (Stanley & Rubin, 2003) and object recognition (Grill-Spector, Kourtzi, & Kanwisher, 2001), responds as well to partly occluded familiar objects as it does to fully unobstructed views of the same objects (Lerner, Hendler, & Malach, 2002).

IT and adjacent areas in the temporal lobe also play a central role in maintaining short-term representations of objects and in analyzing global shape (Kanwisher, Woods, Iacononi, & Mazziotta, 1997), as well as in integrating visual features (Grill-Spector, 2003). In addition, there are specialized regions (in and around the area known as fusiform gyrus) that are most highly active when observers view objects from specific categories of stimuli—faces (Kanwisher & Yovel, 2006; Perrett, Rolls, & Caan, 1982), bodies (Downing, Jiang, Shaman, & Kanwisher (2001), hands (Desimone, Albright, Gross, & Bruce, 1984), artifacts (Chao, Haxby, & Martin, 1999; Martin, Wiggs, Ungerleider, & Haxby, 1996), and locations (Epstein & Kanwisher, 1998)—with each category associated with a specific cortical locus. Under some conditions, such areas can be activated by viewing even object parts (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001), but most neurons in the ventral stream are tuned broadly and respond to a variety of objects and features (Grill-Spector et al., 2001). An important question, therefore, is the extent to which brain areas specialized for specific object categories can be considered “modules” that are dedicated to those categories, or whether they are part of a more general object recognition system (Grill-Spector, 2003). One possibility is that loci of cortical object representations cluster according to level of processing rather than visual attributes. Processing of details that

Object Perception

lead to individuation of distinct items (as occurs in face recognition) as opposed to more generic categories may require dedicated computations that are localized to particular cortical regions. In this view, the so-called fusiform face area (Kanwisher & Yovel, 2006) is not a region dedicated to face recognition, but rather a region for subordinate identification of object category members that has become automated by expertise (Tarr & Gauthier, 2000). Evidence in favor of the expertise hypothesis comes from studies showing that brain areas involved in face recognition are also active in individuals who are experts in identifying different kinds of cars and birds, but are not active in non-experts (Gauthier, Skudlarski, Gore, & Anderson, 2000). Other evidence, however, suggests that the fusiform face area is more sensitive to faces than to other visual stimuli, even in experts, and also shows classic face-selective processing effects, such as holistic processing (for a review, see McKone, Kanwisher, & Duchaine, 2006).

Critical Periods in Object Perception

Learning and experience strongly affect the brain's responses to objects, as discussed in the previous section. Furthermore, research on visual development has shown that experience also alters brain function (specifically, the visual brain) during critical periods of development in infancy and childhood. A critical period is a time when some function or ability must be stimulated or it will be lost permanently (for reviews, see Daw, 1995, 2003).

Critical periods were first revealed by Hubel and Wiesel (1963, 1970), whose work initiated the formal study of the physiology of visual development. Using kittens, Hubel and Wiesel covered a single eye at each animal's birth and left the patch in place for a duration ranging from one to several months. They then investigated the effects of visual deprivation by patching the unaffected eye and documenting visual function of the formerly deprived eye alone, which was now uncovered and exposed to the visual environment. The deprived eye was effectively blind, as revealed by both behavioral and neural effects. Behavioral effects included the kittens' inability to navigate visually or to respond to objects introduced by the experimenters, although the animals behaved normally under the same circumstances when permitted to use the unaffected eye. Neural effects were examined by recording from single cells in visual cortex, and recordings showed that, in general, few cortical cells could be driven by the deprived eye in cortical regions normally responsive to input from both eyes, such as the postlateral gyrus. Wiesel and Hubel also reported the effects of eye closure in animals that were allowed some visual experience prior to deprivation. The unaffected eye dominated activity of cells in the visual cortex, but this effect depended on both the extent of visual experience prior to deprivation and the duration of deprivation. In humans, early deprivation of typical visual experience has been shown to affect acuity, peripheral vision, motion perception, and binocular (stereo) vision; interestingly, there are different critical periods for different visual functions, each with a characteristic time course and duration (for review, see Lewis & Maurer, 2005).

Object Perception

Evidence for a critical period for holistic face perception comes from a study of individuals born with cataracts who underwent surgery during infancy to correct the problem (Le Grand, Mondloch, Maurer, & Brent, 2001). Each individual had at least 9 years of visual experience after surgery. The patients were tested with face recognition tasks, including, importantly, tests of inversion effects, as faces become difficult to recognize when upside-down (Yin, 1969). The individuals demonstrated a specific deficit in recognition from holistic or “configural” information (the spacing of facial features, such as eyes, nose, and mouth) but not from “featural” information (differences among features), where performance was not reliably different from age-matched controls (see Figure 2). A particularly striking finding concerned the timing of cataract replacement, which for every patient occurred when they were less than 7 months old, and in a few cases occurred when they were as young as 2 to 3 months old. The critical period for development of holistic face processing, therefore, appears to be exceedingly brief. Interestingly, infants who are 2 to 3 months old show no signs of the inversion effect (Cashon & Cohen, 2003), and adult levels of sensitivity to some kinds of holistic information in faces is not evidenced until children are several years old (Mondloch, Geldart, Maurer, & Le Grand, 2003).

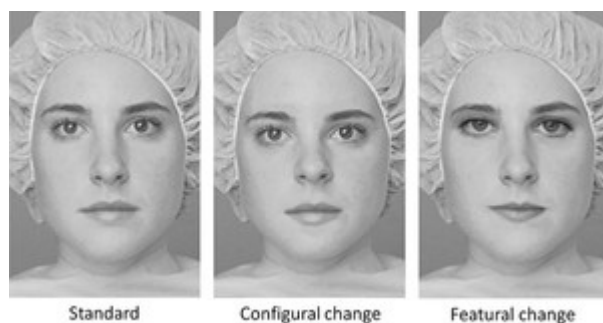


Figure 2. Configural versus featural information. The configural change face is identical to the standard face, except the spacing of the features is displaced (while the features are the same). The featural change face is identical to the standard face, except the features are different (while the spacing is the same). Adapted from Le Grand et al. (2001).

Some types of holistic object perception appear to be compromised by visual deprivation, but the evidence is complex. A study of illusory contour perception in patients who underwent congenital cataract correction provides additional evidence for reductions in configural processing (in this case, perceptual completion and feature binding) caused by early visual deprivation (Putzar, Hötting, Rösler, & Röder, 2007). Patients whose surgery took place after they were 6 months old showed higher reaction times and miss rates when searching for illusory shapes among distracters compared to their reaction times and miss rates when searching for real shapes; patients who had cataract surgery before they were 6 months old, as well as control participants, showed less of a difference on these measures. Interviews conducted after testing revealed that the post-6-month patient group did not perceive the illusory figures at all. In line with the results of the study on

Object Perception

face perception, these results indicate that the first several months after birth are a critical period for spatial integration of visual information.

A case study of MM, a man who lost his vision at age 3.5 years and had cataract replacement nearly 40 years later, revealed that MM had marked deficits in object perception (Fine et al., 2003; for a complete account, see Kurson, 2008). Five months after surgery, MM was unable to detect transparency in overlapping forms, to see depth from perspective in a Necker cube, or to identify a shape (a Kanizsa square) defined by illusory contours (see Figure 3). The latter finding was consistent with the results of the study by Putzar et al. (2007), even though MM had had typical visual experiences for the first 3.5 years of life. After his surgery, MM was also limited in object recognition and had difficulty in discriminating faces and in identifying emotional expression—he reportedly relied on individual features rather than holistic information, which is available to typically sighted perceivers. Cortical areas that give strong responses in typical observers when viewing faces and objects (lingual and fusiform gyri) were largely inactive in MM. Other visual functions, however, were preserved, such as color and contrast sensitivity and motion perception, implying that they are more robust to early deprivation. MM’s object perception skills remained largely unimproved after more than 10 years of postoperative visual experience (Huber et al., 2015). Interestingly, a case study of KP, a man who lost his vision from the age of 17 until cataract surgery at age 71 (53 years of visual deprivation), revealed fewer effects of deprivation on object perception, even with partly occluded objects (Šikl et al., 2013). However, KP’s performance likely was based on matching visual input with stored representations of typical object appearance; when objects were unfamiliar or had a modern design, he had difficulty identifying them. Moreover, KP’s acuity and contrast sensitivity were impaired, likely due to the extensive deprivation, and like some of the cataract patients mentioned previously, he struggled with tasks requiring perceptual organization and holistic processing.

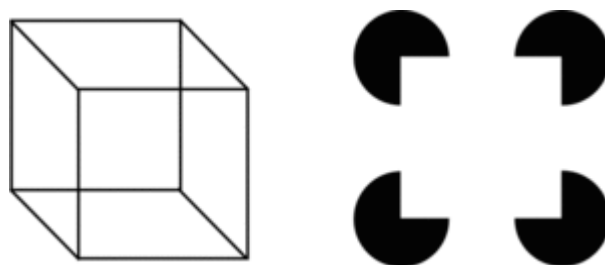


Figure 3. Left: Necker cube. Right: Illusory square, also known as a Kanizsa square.

A study of three younger cataract patients (who had surgery at the ages of 7, 13, and 29 years) who were tested within months after treatment reported that the individuals had difficulty using Gestalt cues (e.g., good continuation) for appropriate segmentation of visual images, in particular binding edges common to objects in multi-object visual stimuli, although the individuals could recognize basic shapes under conditions of partial occlusion (Ostrovsky, Meyers, Ganesh, Mathur, & Sinha, 2009). In contrast, motion cues facilitated perception of edge connectedness and perception of distinct objects in cluttered

Object Perception

scenes, and motion cues even helped support recognition of the same objects viewed in static images.

Taken together, these studies indicate that there are critical periods for normal visual function in several areas, such as acuity and contrast sensitivity, and more importantly for the present discussion, the development of holistic object perception.

Development of Object Perception

As noted already, early learning and experience are vital to object perception, both in the mechanisms that shape brain structure and in the ways that deprivation disrupts holistic object processing. A comprehensive account of object perception also requires an appreciation for the developmental changes that bring a child to an accurate understanding of the visual environment. Investigation of the development of object perception (specifically, object permanence) in infancy had its origins in Piaget's (1954) constructivist account, according to which infants progressively construct an objective knowledge of the world through their own experience with manual activity and coordination of manual and visual skills. In the 1980s, investigators developed ways of testing young infants' responses to hidden objects that relied on simple measures of looking time rather than on coordinated manual activity, and the experiments led to the view that knowledge of the physical world, including object permanence, is innate (Baillargeon, Spelke, & Wasserman, 1985; Spelke, Breinlinger, Macomber, & Jacobson, 1992). A central aspect of innate knowledge is the principle of object persistence (Baillargeon, 2008), which describes the tendency to perceive moving objects as continuing to exist as they become hidden. Experiments with infants from birth through the first 6 postnatal months provided evidence against this perspective, instead revealing developmental processes in infant object perception, in particular perception of objects as coherent and persisting under conditions of partial or complete occlusion (for a review, see Bremner, Slater, & Johnson, 2015).

Key developmental processes that lead to accurate object perception include a gradual increase in infants' ability to perceive connectedness of edges of partly occluded objects. Motion is an important cue for object segmentation in infancy (Kellman & Spelke, 1983): perceptual completion has been demonstrated in 2-month-old infants viewing displays in which moving rod parts are separated by a gap imposed by an occluder (Johnson & Aslin, 1995). Infants between 2 and 4 months old come to tolerate greater spatial gaps across which edges must be interpolated, and there are improvements in detecting alignment or misalignment of candidate edges for assignment to the same surface (Johnson, 2004). This may stem from improvements in information-processing skills, such as directed attention to relevant features of the display that specify occlusion (Johnson, Slemmer, & Amso, 2004). Perception of object persistence seems to develop in a similar fashion, as infants between 2 and 6 months old are able to maintain representations of moving, temporarily hidden objects across increasingly greater temporal and spatial gaps (Johnson et al., 2003).

Object Perception

For adults, deletion and accretion at virtual edges with no visible occluder can yield a “tunnel effect,” such that the object appears to disappear and reappear as if passing through a slit in the occluding surface (Burke, 1952; Michotte, Thines, & Crabbe, 1991). These are important cues to object persistence for infants as well: 5- to 9-month-old infants’ predictive tracking of a moving object (using anticipatory eye movements) passing behind an occluder was reduced if the object underwent instantaneous disappearance/reappearance or implosion/explosion at the occluder boundaries, rather than undergoing deletion and accretion. However, although deletion and accretion may be necessary cues for perception of object persistence by young infants, they do not appear to be sufficient in themselves. Infants 2 and 4 months old have been shown to perceive the visible segments of partly occluded objects and partly occluded object trajectories (i.e., in a temporarily hidden object event of the type described previously) as disjoint surfaces or trajectories, if the temporal or spatial gap is large (Johnson, Bremner, et al., 2003). Also, object persistence is less likely to be perceived by infants if the object’s trajectory is oblique relative to the occluder (Bremner et al., 2007).

Another important developmental mechanism is experience viewing moving objects. Repeated exposure to moving objects facilitates perception of object persistence; when an object previously seen traveling back and forth is then seen to move behind an occluder and subsequently emerge, 4-month-old children show a stronger tendency to anticipate the object’s reappearance, shifting their gaze to the far edge prior to the object’s re-emergence, than is shown by infants who did not first see the unoccluded movement (Johnson, Amso, & Slemmer, 2003).

In summary, older infants achieve perceptual completion and perceive persistence across longer spatial and temporal gaps than younger infants, and they require fewer cues to specify occlusion (and hence persistence). General principles of object perception, such as object persistence, thus seem to have their bases in development of early perceptual capacities. There is ample information in the visual world for segregated objects, laid out in depth (Gibson, 1979). Development is about improvement in detection and utilization of this information, and constructing representations of object properties (such as persistence) from repeated exposures.

Conclusions

Object perception consists of an interplay between bottom-up processing and top-down knowledge. The human visual system is organized in stages: the registration of low-level features, the segmentation of input of separate surfaces, the disambiguation of complex scenes, the identification of distinct objects, and the recognition of items from different object categories that can be highly specific. Interestingly, developmental processes of object perception in infancy also proceed from detection of simple visual attributes, to integrating features, and eventually to perceiving objects as permanent and cohesive across time and space. Deprivation of typical visual experience can severely disrupt some aspects of these developmental processes, but some of the more fundamental visual

Object Perception

skills, such as motion detection (also paramount for object segmentation, especially early in development), may be somewhat resistant to deprivation's deleterious effects.

Further Reading

Atkinson, J. (2000). *The developing visual brain*. New York, NY: Oxford University Press.

Cornsweet, T. N. (1970). *Visual perception*. New York, NY: Academic Press.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. New York, NY: Oxford University Press.

References

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, 3, 2-13.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20, 191-208.

Bertenthal, B. J., Longo, M. R., & Kenny, S. (2007). Phenomenal permanence and the development of predictive tracking in infancy. *Child Development*, 78, 350-363.

Bremner, J. G., Johnson, S. P., Slater, A., Mason, U., Cheshire, A., & Spring, J. (2007). Conditions for young infants' failure to perceive trajectory continuity. *Developmental Science*, 10, 613-624.

Bremner, J. G., Slater, A. M., & Johnson, S. P. (2015). Perception of object persistence: The origins of object permanence in infancy. *Child Development Perspectives*, 9, 7-13.

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4, 121-138.

Buschman, T. J., & Miller, E. K. (2014). Goal-direction and top-down control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20130471.

Cashon, C. H., & Cohen, L. B. (2003). Beyond U-shaped development in infants' processing of faces: An information-processing account. *Journal of Cognition and Development*, 5, 59-80.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, 2, 913-919.

Daw, N. W. (1995). *Visual development*. New York, NY: Plenum Press.

Daw, N. W. (2003). Critical periods in the visual system. In B. Hopkins & S. P. Johnson (Eds.), *Neurobiology of infant vision* (pp. 43-103). Westport, CT: Praeger.

Object Perception

- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *5539*, 2470–2473.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47.
- Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., . . . MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, *6*, 915–916.
- Heydt, R. von der, Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262.
- Hubel, D. H., & Wiesel, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, *26*, 994–1002.
- Hubel, D. H., & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, *206*, 419–436.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.
- Gibson, J. J. (1950). *The perception of the visual world*. Westport, CT: Greenwood Press.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 20–25.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*, 1–8.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- Huber, E., Webster, J. M., Brewer, A. A., MacLeod, D. I. A., Wandell, B. A., Boynton, G. M., . . . Fine, I. (2015). A lack of experience-dependent plasticity after more than a decade of restored sight. *Psychological Science*, *26*, 393–401.
- Johnson, S. P. (2004). Development of perceptual completion in infancy. *Psychological Science*, *15*, 769–775.

Object Perception

- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences (USA)*, *100*, 10568-10573.
- Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, *31*, 739-745.
- Johnson, S. P., Bremner, J. G., Slater, A., Mason, U., Foster, K., & Cheshire, A. (2003). Infants' perception of object trajectories. *Child Development*, *74*, 94-108.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, *6*, 185-201.
- Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*, *9*, 133-142.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *361*, 2109-2128.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141-221.
- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483-524.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271-304.
- Koffka, K. (1935). *Principles of Gestalt psychology*. London, UK: Routledge & Kegan Paul.
- Kurson, R. (2008). *Crashing through: The extraordinary true story of the man who dared to see*. New York, NY: Random House.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *410*, 890.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, *11*, 287-297.
- Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cerebral Cortex*, *12*, 163-177.
- Lewis, T. L., & Maurer, D. (2005). Multiple sensitive periods in human visual development: Evidence from visually deprived children. *Developmental Psychobiology*, *3*, 163-183.

Object Perception

- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Freeman.
- Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.
- McKone, E., Kanwisher, N., & Duchaine, B. C. (2006). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*, 8–15.
- Michotte, A., Thines, G., & Crabbe, G. (1991). *Les compléments amodaux des structures perceptives*. Louvain: Publications Universitaires. Excerpted in G. Thines, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 140–167). Hillsdale, NJ: Erlbaum. (Original work published 1964.)
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*, 774–785.
- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417.
- Miyashita, Y., & Hayashi, T. (2000). Neural representation of visual objects: Encoding and top-down activation. *Current Opinion in Neurobiology*, *10*, 187–194.
- Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology*, *86*, 67–84.
- Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U., & Sinha, P. (2009). Visual parsing after recovery from blindness. *Psychological Science*, *20*, 1484–1491.
- Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342.
- Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). New York, NY: Basic Books. (Original work published 1937.)
- Putzar, L., Hötting, K., Rösler, F., & Röder, B. (2007). The development of visual feature binding processes after visual deprivation in early infancy. *Vision Research*, *47*, 2616–2626.
- Roelfsema, P. R., & Singer, W. (1998). Detecting connectedness. *Cerebral Cortex*, *8*, 385–396.
- Šikl, R., Šimeček, M., Porubanová-Norquist, M., Bezdíček, O., Kremláček, J., Stodůlka, P., . . . Ostrovsky, Y. (2013). Vision after 53 years of blindness. *Iperception*, *4*, 498–507.

Object Perception

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18*, 555–586.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605–632.

Stanley, D. A., & Rubin, N. (2003). fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron, 37*, 323–331.

Tanaka, K. (1997). Mechanisms of visual object recognition: Monkey and human studies. *Current Opinion in Neurobiology, 7*, 523–529.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*, 764–769.

Van Essen, D. C., Glasser, M. F., Dierker, D. L., & Harrell, J. (2011). Cortical parcellations of the macaque monkey analyzed on surface-based atlases. *Cerebral Cortex, 22*, 2227–2240.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625–636.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Zeki, S. (1993). *A vision of the brain*. Cambridge, MA: Blackwell.

Zuckerman, C. B., & Rock, I. (1957). A re-appraisal of the roles of past experience and innate organizing processes in visual perception. *Psychological Bulletin, 54*, 269–296.

Scott P. Johnson

Department of Psychology, University of California Los Angeles