

Chapter 2: Basic processes in visual perception

There has been considerable progress in understanding visual perception in recent years. Much of this is due to the efforts of cognitive neuroscientists, thanks to whom we now have reasonable knowledge of the brain systems involved in visual perception.

- **Vision and the brain**

There are three major consequences when a visual stimulus reaches receptors in the retina: reception, transduction and coding:

- The amount of light entering the eye is determined by the pupil.
- The lens adjusts in shape to bring images into focus on the retina.

There are two types of visual receptor cells in the retina: cones and rods. There are 6 million cones, mostly in the fovea, which are specialised for colour vision and sharpness. There are 125 million rods, which are specialised for vision in dim light and for movement detection:

- Impulses leave the eye via the optic nerve.
- The main pathway between eye and cortex is the retina-geniculate-striate pathway.

Two stimuli adjacent to each other in the retinal image will also be adjacent to each other at higher levels within that system (*retinopy*). Signals proceed along two optic tracts within the brain. One tract contains information from the left half of each eye and the other tract from the right half. Nerves reach the primary visual cortex (V1) within the occipital lobe before spreading to secondary visual areas. There are two relatively independent channels within this system:

- The P (parvocellular) pathway, sensitive to colour and detail, has most input from cones.
- The M (magnocellular) pathway, sensitive to movement, has most input from rods.

WEBLINK: [Anatomy, physiology and pathology of the human eye](#)

The main route between the eye and the cortex is divided into P and M pathways. There are two main pathways in the visual cortex, one terminating in the parietal cortex and the other terminating in the inferotemporal cortex. According to Zeki's functional specialisation theory, different parts of the cortex are specialised for different visual functions. There is some support for this view, but there is much less specialisation than claimed by Zeki.

Neurons from P and M pathways mainly project to V1. The P pathway associates with the ventral or "what" pathway, concerned with form and colour processing, and proceeds to the inferotemporal cortex. The M pathway associates with the dorsal or "how" pathway, concerned with movement processing, and proceeds to the posterior parietal cortex. However, information processing in the two pathways is by no means totally or cleanly segregated (Leopold, 2012).

The receptive field for a given neuron is the region of retina in which light affects activity. Lateral inhibition is a reduction of activity in one neuron caused by activity in a neighbouring neuron. It is useful because it increases the contrast at the edges of objects. V1 and V2 occupy relatively large areas within the cortex. There is increasing evidence that early visual processing in V1 and V2 is very extensive, for example the macaque monkey study by Hegdé and Van Essen (2000).

In addition to the initial “feedforward sweep” of early visual processing, [Lamme \(2006\)](#) describes a second phase of recurrent processing in which feedback signals proceed in the opposite direction.

WEBLINK: [The visual cortex](#)

Zeki (1993, 2001) proposed that different parts of the cortex are specialised for different visual functions. The importance of V1 is shown by lesions at any point from retina to V1, which cause virtually total blindness in the affected part:

- V1 and V2 are involved at an early stage and respond to colour and form.
- V3 and V3a respond to form (especially in motion) but not to colour.
- V4 responds to colour and line orientation.
- V5 is specialised for visual motion.

Form processing

Several visual areas are involved in form processing. However, the cognitive neuroscience approach has focused mainly on the inferotemporal cortex (IT). Baldassi et al. (2013) measured neuronal activity within anterior inferotemporal cortex in two monkeys. Many neurons responded on the basis of aspects of form or shape (round, star-like, horizontal thin, pointy, vertical thin) rather than object category. Neurons in the anterior inferotemporal cortex may be very specific in their responsiveness (high object selectivity and low tolerance) or show high responsiveness (low object selectivity and tolerance). Zeki (1992) claimed that no one has ever reported a complete and specific loss of form vision.

Colour processing

Patients with achromatopsia show little or no colour perception but have near normal perception of form, motion and fine detail. [Bouvier and Engel \(2006\)](#) reported that nearly all cases of achromatopsia showed damage in or near to V4. However, these patients also showed deficits in spatial vision. Wade et al. (2002) had previously found area V4 was actively involved in colour processing but other areas (V1 and V2) were also activated. However, there is much evidence that other visual areas are also involved in colour processing. Area V4 may also be involved in other aspects of visual processing apart from colour processing.

Motion processing

V5 (or MT, middle temporal) is involved in motion processing. When TMS is applied to V5/MT, it produced a subjective slowing of stimulus speed and impaired observers’ ability to discriminate between different speeds (McKeefry et al., 2008). Brain-damaged patients who suffer from akinetopsia find that objects in motion become invisible. [Zihl et al. \(1983\)](#) studied patient LM who has bilateral V5 damage. Another area that is involved in motion processing is area MST (medial superior temporal), which is adjacent to V5 ([Vaina, 1998](#)). This area is thought to be involved in the visual guidance of walking. Different mechanisms may underlie perception of first-order motion (luminance difference between moving shape and background) and second-order motion (no luminance difference). [Rizzo et al. \(2008\)](#) reported that patients with damage to the visual cortex could have deficits limited to either first- or second-order motion perception.

Binding problem

If visual processing is widely distributed across areas of the brain, information about motion, colour and form will need to be combined into a coherent percept for object recognition to occur (the binding problem). Solutions proposed for the binding problem are as follows:

- Assuming that there is less functional specialisation than Zeki claimed (Seymour et al., 2009).

- Feldman (2013) argued that there are actually several binding problems.
- Binding-by-synchrony (e.g., Singer & Gray, 1995).
- Visual perception depends on patterns of neural activity over time rather than on precise synchrony ([Guttman et al., 2007](#)).

Zeki's theory is a simple overview of a complex reality. Limitations with this approach are as follows:

- Brain areas are not nearly as specialised in their processing as implied by the theory.
- Early visual processing in V1 and V2 is more extensive than suggested.
- The binding problem is not satisfactorily addressed.

- **Two visual systems: perception and action**

Milner and Goodale (1995, 2008) proposed that there are two visual systems with four characteristics (Schenk & McIntosh, 2010):

- a vision-for-perception system:
 - based on the ventral pathway
 - allocentric
 - long-term representations
 - usually conscious processing;
- a vision-for action system:
 - based on the dorsal pathway
 - egocentric
 - short-term representations
 - unconscious processing.

There is convincing evidence from brain-damaged patients, notably the presence of the predicted double dissociation. Patients with optic ataxia ([Perenin & Vighetto, 1988](#)) have damage to the dorsal pathway. They have problems with production of visually guided motions. Patients with visual agnosia ([Milner et al., 1991](#); [James et al., 2003](#); patient DF) have damage to the ventral pathway. They have problems with object recognition but are able to perform visually guided movements normally.

According to Milner and Goodale (1995, 2008), most visual illusions involve the ventral vision-for-perception system. In their meta-analysis, [Bruno et al. \(2008\)](#) found that illusory effects were four times greater in the Müller-Lyer illusion studies involving the vision-for-perception system than studies involving the vision-for-action system. [Króliczak et al. \(2006\)](#) found that the hollow-face illusion was reduced when participants made rapid movements involving the dorsal stream.

INTERACTIVE EXERCISE: [Müller-Lyer](#)

WEBLINK: [Hollow-face illusion](#)

Action

Milner and Goodale (2008) argued that most tasks in which observers grasp an object involve some processing in the ventral stream as well as the dorsal stream. Involvement of the ventral stream is especially likely in the following circumstances:

- Memory is required (e.g., there is a time lag between the offset of the stimulus and the start of the grasping movement).
- Time is available to plan the forthcoming movement (e.g., Kroliczak et al., 2006).
- Planning which movement to make is necessary.

- The action is unpractised or awkward.

[Creem and Proffitt's study \(2001\)](#) suggests that perception for action sometimes depends on the ventral pathway as well as the dorsal pathway. [Milner and Goodale \(2008\)](#) suggested that the ventral pathway is involved in planning for actions that are not automatic. Evidence for this comes from patients with optic ataxia who have damage to the dorsal stream. They perform better when making delayed (memory) rather than immediate movements to a target ([Milner et al., 2003](#)). Visually guided action can occur in the absence of conscious awareness and with the probable use of the dorsal stream (Roseboom & Arnold, 2011).

The central assumption that there are somewhat separate visual systems underlying perception for recognition and perception for action is probably broadly correct. There is support for this view from studies on brain-damaged patients and studies involving visual illusions. However, both processing streams are able to influence reaching and grasping, and the two visual systems typically interact with each other.

Milner and Goodale's influential model of visual perception posits two separate visual systems fulfilling different functions: vision for perception (ventral pathway) and vision for action (dorsal pathway). There is support for this model from studies on patients with damage to dorsal or ventral pathways, demonstrating a double dissociation in function between the two pathways. Studies on visual illusions have also shown that illusory effects depend more on ventral processing and may be diminished when the dorsal stream is involved. However, the functions of the two pathways cannot be dichotomised as both processing streams are involved in planning and executing motor actions, and there are numerous interactions and connections between dorsal and ventral visual streams.

- **Colour vision**

Colour is important to make an object stand out from its background. It also helps us recognise and distinguish between objects (e.g., finding ripe fruit). Colour has three main qualities: hue, brightness and saturation.

Trichromacy theory

Cone receptors contain light-sensitive photopigment. Trichromatic theory describes three kinds of cone receptors. One responds to short-wavelength light, perceived as “blue”, one to medium-length (yellow–green) and the last to long-wavelength light (orange–red). Other colours are perceived according to the relative amount of stimulation of each cone type. If all three types of cones are activated, we see white.

All three cone types are distributed quite randomly, except that there are few cones responding to short-wavelength light within the fovea. The most common type of colour deficiency is dichromacy, in which one cone class is missing. The trichromatic theory fails to account for negative afterimages.

WEBLINK: [Colour perception](#)

WEBLINK: [Colour-blindness tests](#)

Opponent processes

Hering's (1878) key assumption was that there are three types of opponent processes or channels in the visual system:

- a red–green channel, which will perceive green when responding in one way and red when responding in the opposite way;

- a blue–yellow channel;
- an achromatic (white–black) channel.

DeValois and DeValois (1975) found evidence of opponent cells in monkeys' lateral geniculate nucleus (LGN). The theory predicts that it is impossible to see blue and yellow, or red and green, together. Abramov and Gordon (1994) found evidence for this. Opponent-process theory helps to explain colour deficiency and negative afterimages.

WEBLINK: [Colour after-effect](#)

WEBLINK: [Opponent-process theory](#)

Dual-process theory

[Hurvich and Jameson \(1957\)](#) developed this theory as a synthesis of the trichromacy and opponent-process theories. According to this theory, signals from the three cone types are sent to opponent cells. The difference in the activity of types of cones is processed along three channels: achromatic, blue–yellow and red–green. There is support for the dual-process theory, however it is probably an oversimplification of colour perception.

Colour constancy

Colour constancy is the tendency for an object to appear the same colour when the wavelength of light illuminating it changes. The phenomenon indicates that colour vision does not depend only on wavelength of reflected light. Reeves et al. (2008) argued that it is important to distinguish between our subjective experience and our judgements about the world. According to Land's retinex theory (1977, [1986](#)), we decide on the colour of a surface by comparing wavelength reflection against that of adjacent surfaces (context). [Foster and Nascimento \(1994\)](#) proposed that cone-excitation ratios underlie colour constancy. [Reeves et al. \(2008\)](#) argued that subjective experiences and judgements can affect our perception of colour.

CASE STUDY: Does colour constancy exist?

WEBLINK: [Colour constancy example](#)

Chromatic adaptation is when sensitivity to light of any given colour decreases over time. This reduces the distorting effects of any given illumination on colour constancy. [Kraft and Brainard \(1999\)](#) found that the most important factor in colour constancy was local contrast, followed by global contrast. Top-down influences and knowledge of familiar colours of objects can also have a strong effect on colour constancy. [Zeki \(1983\)](#) found that certain cells in monkey area V4 respond to the actual colour of a surface rather than simply to wavelength, exhibiting colour constancy. Colour constancy may also affect other kinds of object processing, such as perceived shape. Many factors make a contribution to colour constancy:

- local contrast (Kraft & Brainard, 1999);
- global contrast;
- cone-excitation ratios;
- top-down factors;
- chromatic adaptation (Lee et al., 2012).

However, we still lack a comprehensive theory indicating how the various factors combine to produce colour constancy.

Colour vision helps us to detect objects and to make fine discriminations among objects. According to the trichromatic theory, there are three types of cone receptors that differ in the light wavelengths to which they respond most strongly. The opponent-process theory argues for three types of opponent processes in the

visual system: green–red, blue–yellow and white–black. A dual-process theory synthesises the earlier two theories and accounts reasonably well for colour perception. Colour constancy occurs when a surface seems to have the same colour when there is a change in the illuminant. Chromatic adaptation and familiar colour are two factors involved in colour constancy, but there are several others.

- **Depth**

A major accomplishment of vision is the transformation of a 2-D retinal image into a 3-D perception. Judgements of relative distance are more accurate than actual distance judgements. In real life, depth cues are often provided by movement. With static objects, depth cues are monocular cues, binocular cues and oculomotor cues:

- Monocular cues require only one eye.
- Binocular cues need both eyes to work together.
- Oculomotor cues rely on sensations from muscles around the eyes.

Monocular cues

These are sometimes called pictorial cues:

- Linear perspective is when the convergence of lines creates a powerful impression of depth in a 2-D drawing.
- Aerial perspective is when distant objects look hazy.
- Texture gradient (Gibson, 1979) runs from the front to the back of a slanting object.
- Interposition is a cue in which a nearer object hides a more distant one. This is seen in Kanizsa's (1976) illusory square.
- Shading provides good evidence of a 3-D object ([Ramachandran, 1988](#)).
- Familiar size is another cue. We can use the retinal image size to provide an estimate of distance when we know its actual size.
- Motion parallax refers to movement of an object's image over the retina due to movement of the observer's head.

WEBLINK: [Cues to depth \(visual descriptions\)](#)

Oculomotor and binocular cues

Oculomotor cues:

- Convergence is when eyes turn inwards to focus on an object more for a close object than for one farther away.
- The usefulness of convergence is disputed and it is only useful up to a few metres.
- Accommodation is when the lens thickens to focus on a close object.
- Accommodation is also only a useful cue at very close quarters.

Binocular cues:

- Stereopsis depends on the difference in the images projected on the retinas of the two eyes.
- Bruce et al. (2003) found that stereopsis rapidly becomes less effective at greater distances.
- Stereopsis involves two stages. Matched features in the two eyes are identified and then retinal disparities are calculated.
- Mistakes in stereopsis can lead to visual illusions such as the wallpaper illusion. An autostereogram is a 2-D image containing depth information so that it appears 3-D.
- Most regions of the visual cortex have neurons that respond strongly to binocular disparity.

- Both dorsal and ventral processing streams are involved in stereopsis. Both streams process absolute and relative disparity, but there is more complete processing of relative disparity in the ventral stream.

INTERACTIVE EXERCISE: [Depth perception test](#)

Information from several depth cues may be combined either additively, selectively or in more complex ways. [Jacobs \(2002\)](#) argued that we combine information from multiple visual cues by assigning more weight to reliable cues. [Bruno and Cutting \(1988\)](#) found support for the notion of additivity when participants viewed untextured parallel flat surfaces monocularly. However, when two or more cues provide conflicting depth information, selection may be used, as in Gregory's (1973) "hollow-face" illusion in which stereoscopic information is being ignored. [Triesch et al. \(2002\)](#) found evidence that less ambiguous cues are regarded as more reliable and [Atkins et al. \(2001\)](#) found evidence that a cue is regarded as reliable if it is consistent with other available cues.

Information from different depth cues is typically combined to produce accurate depth perception, often in an additive fashion. Depth perception is most likely to depend on only one cue if different cues give very conflicting evidence. There is also much support for the view that we attach more weight to cues that are reliable, and that we assign this weighting flexibly.

WEBLINK: [Ambiguous depth cues](#)

Size constancy

Size constancy is the tendency for objects to appear the same size whether their size in the retinal image is large or small – for example, someone walking towards you seems to remain the same size. Perceived size and size constancy generally depend in part on perceived distance.

WEBLINK: [Size constancy](#)

An Ames room (Ames, 1952) creates an illusory effect in which the perceived distance of an object drives its perceived size, but perceived distance differs considerably from actual distance. Haber and Levin (2001) argued strongly that size perception typically depends on memory of objects' familiar size rather than on perceptual information. Their findings support this, however they cannot account for the fairly high accuracy of size judgements of unfamiliar objects. Witt et al. (2008) found that objects look larger when we have the ability to act effectively with respect to them.

WEBLINK: [Ramachandran explains the Ames room](#)

Size perception and size constancy depend mainly on perceived distance and familiarity. We do not yet have a coherent account of how these factors combine to produce size judgements.

WEBLINK: [Hundreds of visual illusions](#)

Monocular cues to depth include linear perspective, aerial perspective, texture, shading, shadows, familiar size and motor parallax. Convergence and accommodation are oculomotor cues of limited usefulness. Stereopsis involves binocular cues, and is based on establishing correspondences between the information presented to one eye and that presented to the other eye. Information from depth cues is often combined, but combination is less common when there are gross differences in the depth information supplied by

different cues. Size constancy depends mainly on perceived distance, but familiar size and horizon information can both be used to estimate size.

- **Perception without awareness**

Blindsight

Patients with blindsight have severe brain damage to V1 and lack conscious perception in the damaged visual field, but surprisingly still show residual visual abilities (e.g., motion detection) in the blind visual field (Ko & Lau, 2012). It is thought that blindsight vision relies on pathways that bypass V1. Three subtypes of blindsight have been proposed (see [Dancikert & Rossetti, 2005](#)):

- Action-blindsight: patients can make use of dorsal stream processing to grasp or point at objects in the blind field.
- Attention-blindsight: patients make use of the dorsal stream and motor areas to detect objects and motion.
- Agnosopsia: patients can discriminate form and wavelength and use the ventral stream.

The most thoroughly studied patient is DB (e.g., Weiskrantz, 2010). He could identify the location of a stimulus presented to his “blind” area but reported no conscious experience. He stated that he was just guessing. Interestingly, DB experienced negative afterimages without conscious perception of the original object. Blindsight can be very unlike normal conscious vision. [Persaud and Cowey \(2008\)](#) reported that patient GY could detect the location of a visual stimulus but lacked conscious awareness of that information. Overgaard et al. (2008) demonstrated that, when more sensitive self-reporting criteria were used, blindsight patients showed evidence of having degraded conscious vision.

WEBLINK: [A neat demonstration of blindsight](#)

There is solid evidence that blindsight is a genuine phenomenon. However, some blindsight patients may actually possess some conscious visual awareness in the blind field. Blindsight patients also show considerable differences, and may have abnormal visual processing pathways ([Bridge et al., 2008](#)).

Subliminal perception

Visual stimuli may be presented below the level of conscious awareness if they are weak, presented very briefly or immediately followed by a masking stimulus. Observers often show “awareness” of a stimulus assessed by the objective threshold even when the stimulus does not exceed the subjective threshold. We can have more confidence that unconscious perception has been demonstrated in studies using objective measures. [Naccache et al. \(2002\)](#) reported that, in a masked digit paradigm, masked digits received some unconscious perceptual processing when participants were cued to attend to it. [Persaud and McLeod \(2008\)](#) found that whether participants consciously perceived a stimulus or not had effects on their behaviour.

WEBLINK: [Subliminal perception: facts and fallacies](#)

The notion of unconscious perception is controversial. However, there is now reasonable evidence from behavioural and neuroimaging studies for some level of visual processing for stimuli that are not consciously perceived. Windey et al. (2014) found that perceptual awareness was graded with the lower-level tasks. However, it was all-or-none with high-level tasks.

WEBLINK: [Kazdin: subliminal perception](#)

Blindsight is a genuine phenomenon in which patients with severe damage to the visual cortex retain some unconscious perception of visual stimuli. These residual visual abilities may depend on processing in the dorsal visual stream or on subcortical mechanisms. Unconscious perception can be assessed using a subjective threshold or a more stringent objective threshold. There is evidence for unconscious perception in neuroimaging and behavioural studies in which stimuli that are not consciously perceived nevertheless receive some visual processing and may even influence behaviour.

Additional references

- Abramov, I. & Gordon, J. (1994). Colour appearance: On seeing red, or yellow, or green, or blue. *Annual Review of Psychology*, 36: 715–29.
- Gregory, R.L. (1973). The confounded eye. In R.L. Gregory & E.H. Gombrich (eds), *Illusion in nature and art*. London: Duckworth.
- Land, E.H. (1977). The retinex theory of colour vision. *Scientific American*, 237: 108–28.