<table>
<tr><td>**5.11**</td><td>**Perception**</td><td>John D. Mollon</td></tr>
</table>

## OVERVIEW

Many psychiatric conditions are characterised by hallucinations or by subtler alterations in perception [1–3]. To understand such symptoms, it is useful to have some understanding of normal perception. What are the intervening processes between an image falling on our retina and our recognition of a familiar face in a complex scene? Or between the variations in air pressure reaching our cochlea and our recognition of the specific words that are being spoken – as well as our recognition that they are being spoken by our youngest cousin?

### 5.11.1 Feature Analysis
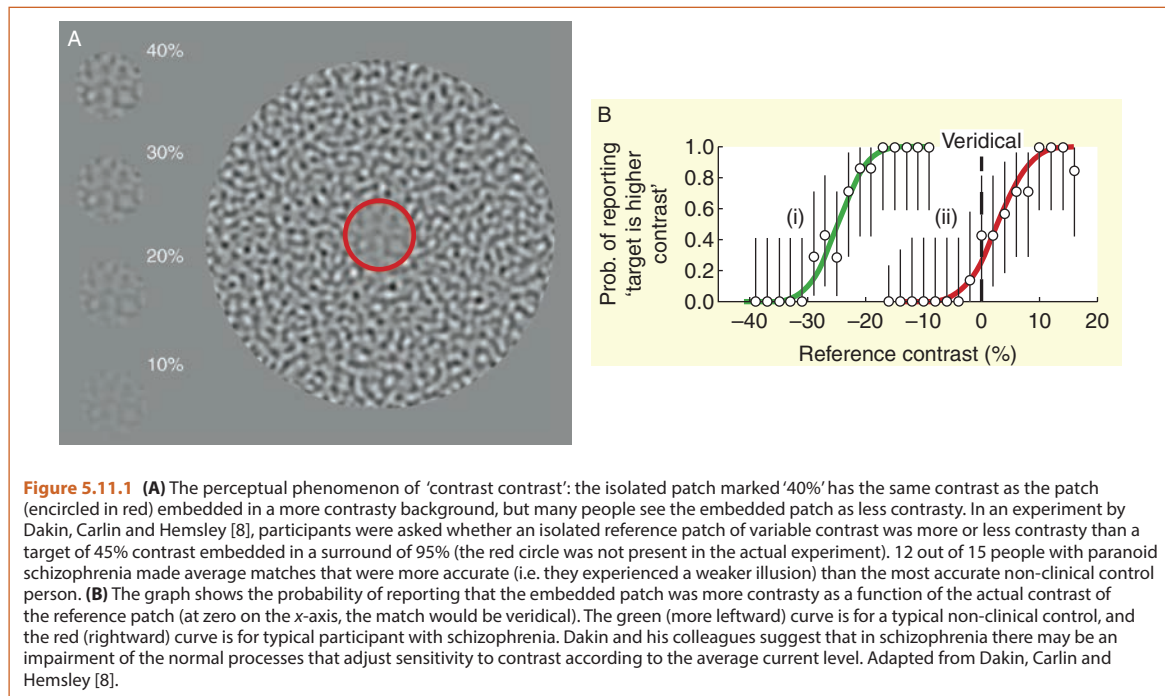
#### 5.11.1.1 The Visual System

The first stage in the remarkable train of processes involved in perception (and the best understood) is the extraction of particular features from the physical stimulus. The retina, for example, does not transmit a passive image to the brain, pixel by pixel. Within the retina, there are at least 19 different types of ganglion cell, which act as 'pre-processors', extracting different attributes from the ever-shifting pattern of light that falls on the rods and cones – attributes such as lightness, colour, temporal change, spatial detail, fine and coarse texture, and motion [4, 5]. The different types of ganglion cell are distinguished by their morphologies, by their immuno-chemistry (i.e. the proteins they express), by the strata of the retina in which their dendritic fields extend – and by the different sites within the visual brain to which their reports are delivered.

The main destination of signals from the retina (though certainly not the only one) is the primary visual cortex, the 'striate' cortex, which lies at the very back of the brain, in the occipital lobe. Here, the specialisations of individual neurons for particular features are maintained, but a new type of tuning famously emerges: many neurons are selective for edges of a particular orientation, and often this preference is combined with tuning for other attributes, such as colour or direction of motion [5–7]. Particular regions of 'pre-striate cortex' – the parts of the occipital lobe anterior to the primary visual cortex – appear to be specialised for the analysis of particular attributes, such as motion or colour or the disparities between the images on the retinas of the two eyes that give rise to stereoscopic perception of depth.

The cortical neurons that detect a particular visual feature (e.g. colour, orientation) typically have lateral, and mutually inhibitory, connections with nearby neurons that detect the same feature in adjacent areas of the field – an arrangement that is thought to allow our perceptual system to adjust to the current average value of a given attribute and to emphasise local departures from that average. Sometimes, however, this useful adjustment gives rise to the class of illusions called 'simultaneous contrast'. An example is given in Figure 5.11.1.

Another class of perceptual distortions arises from temporary miscalibrations in the arrays of feature-detecting neurons [7, 9]. For example, if we look fixedly for some minutes at a waterfall or at a moving flow of traffic, then – for a few moments afterwards – stationary objects may appear to move in the opposite direction. Such 'negative motion after-effects' are thought to arise from reductions in the sensitivity of a selective subset of the directionally selective neurons that are present at early stages of the visual system.

**Figure 5.11.1 (A)** The perceptual phenomenon of 'contrast contrast': the isolated patch marked '40%' has the same contrast as the patch (encircled in red) embedded in a more contrasty background, but many people see the embedded patch as less contrasty. In an experiment by Dakin, Carlin and Hemsley [8], participants were asked whether an isolated reference patch of variable contrast was more or less contrasty than a target of 45% contrast embedded in a surround of 95% (the red circle was not present in the actual experiment). 12 out of 15 people with paranoid schizophrenia made average matches that were more accurate (i.e. they experienced a weaker illusion) than the most accurate non-clinical control person. **(B)** The graph shows the probability of reporting that the embedded patch was more contrasty as a function of the actual contrast of the reference patch (at zero on the *x*-axis, the match would be veridical). The green (more leftward) curve is for a typical non-clinical control, and the red (rightward) curve is for typical participant with schizophrenia. Dakin and his colleagues suggest that in schizophrenia there may be an impairment of the normal processes that adjust sensitivity to contrast according to the average current level. Adapted from Dakin, Carlin and Hemsley [8].

Pathologies may affect one visual attribute more than another. Indeed, the first suggestion that the optic nerve might contain different fibres for different attributes came from the former Surgeon to the Confederate Army J. J. Chisholm, who observed in 1869 that spatial vision recovered before colour vision in a case of optic neuritis [10]. Conversely, colour recognition may be relatively preserved in cases where carbon monoxide poisoning leads to 'visual form agnosia', an inability to recognise objects by their shape [11].
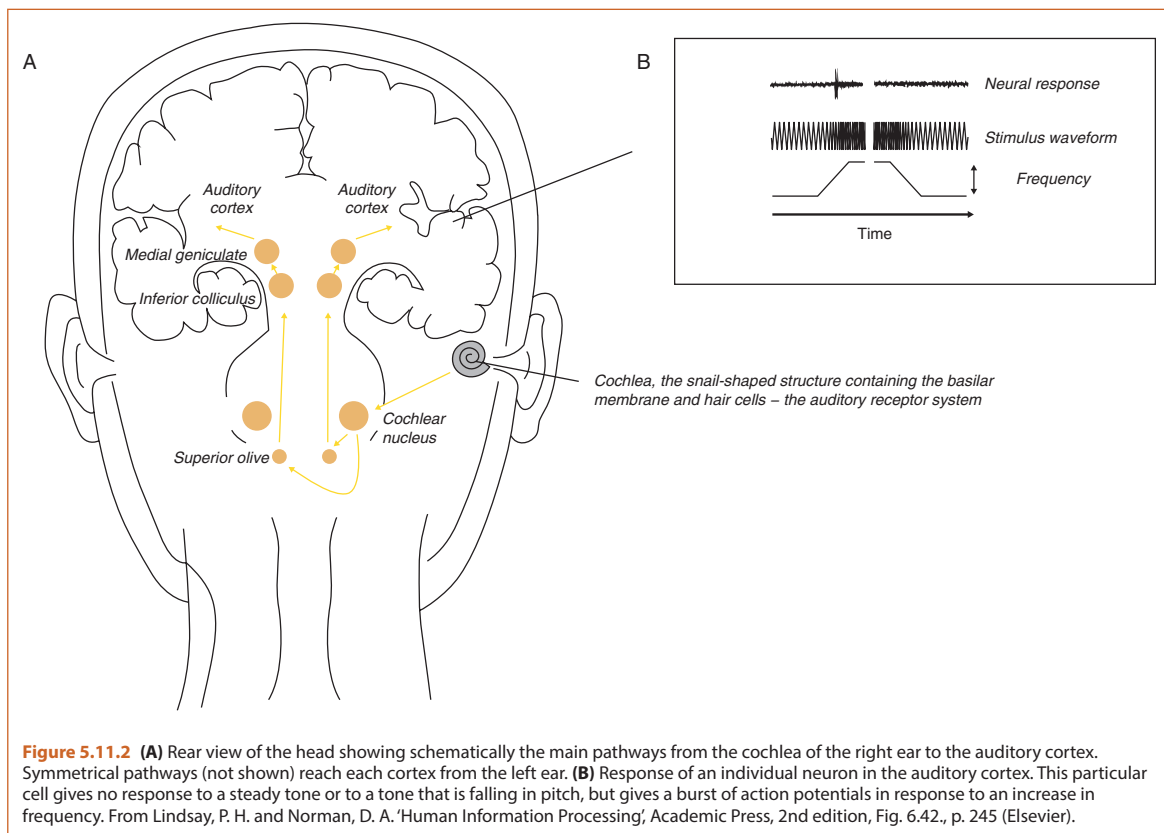
### 5.11.1.2 The Auditory System

Neurons specific for particular features of the stimulus are also found in the auditory system. The individual fibres that leave the cochlea of the ear are tuned just to particular frequencies of sound, but at subsequent stages (including the cochlear nucleus, the inferior colliculus and the auditory cortex; Figure 5.11.2A) there are, for example, neurons that respond to a particular direction of change of frequency – to a rising pitch or to a falling one [12]. The response of such a cell is illustrated in Figure 5.11.2B. A change in the frequency of sound can be seen as analogous to motion across the visual field: in the one case there is motion along the basilar membrane of the ear's

cochlea and in the other there is motion across the retina. So is there an auditory analogue of the negative motion after-effect discussed above? Indeed there is: if listeners are exposed to a band of frequencies that repeatedly increase in a sawtooth fashion, then afterwards a steady sound appears to fall in pitch [13].

### 5.11.1.3 Beyond Feature Analysis

In sum, at early stages of our sensory systems there are many parallel neural channels, often morphologically or anatomically distinct, that identify particular features of the external stimulus. But there is much more to our perception than this. First, there is the large matter of 'perceptual organisation'. The brain must parse the visual or auditory scene into distinct objects (e.g. faces, words), deciding what elements of the scene belong together, distinguishing figure from ground and estimating the three-dimensional arrangement of the scene. Second, the *identity* of each object must be recognised. In addition, corrections must be made to the estimated size, shape, colour, loudness, etc. of objects according to the conditions of viewing or listening – a process often termed 'perceptual constancy'. These several processes, described in the sections that follow, are not independent

**Figure 5.11.2 (A)** Rear view of the head showing schematically the main pathways from the cochlea of the right ear to the auditory cortex. Symmetrical pathways (not shown) reach each cortex from the left ear. **(B)** Response of an individual neuron in the auditory cortex. This particular cell gives no response to a steady tone or to a tone that is falling in pitch, but gives a burst of action potentials in response to an increase in frequency. From Lindsay, P. H. and Norman, D. A. 'Human Information Processing', Academic Press, 2nd edition, Fig. 6.42., p. 245 (Elsevier).

and successive. For example, some independent cue or some sudden recollection may cause a switch in our identification of an ambiguous object; and then our entire three-dimensional interpretation of the scene may change. What we initially perceived as an object may be reinterpreted as a shadow, and then its apparent colour may change. If a sound is reinterpreted as coming from a nearby insect rather than from distant farm machinery, its apparent loudness may change.

## 5.11.2 Perceptual Organisation

In the first half of the twentieth century, the Gestalt psychologists identified several rules that describe how the elements of a visual or auditory scene will be perceptually organised, that is, how the elements will be grouped to form segregated objects [14]:

(1) **Proximity in space**. Our ability to localise sounds is valuable not only in itself, but also because it allows the segregation of one auditory stream from another. For example, it allows us to follow one source of speech out of several at a busy drinks party (and not necessarily the person we are nominally listening to!). Similarly, in binocular stereo vision, elements that lie in different planes are displaced relative to one another on the two retinas and so emerge as distinct objects at different depths – allowing us to penetrate camouflage that defeats the monocular eye.

(2) **Similarity**. The elements in an auditory or visual scene tend to be grouped, and perceived as one object or sound source, if they are similar in some quality. In vision, this might be similarity in size, in colour, in orientation of contour, in texture or in shape. In hearing, similarity of frequency in successive tones determines whether a sequence will be heard as two streams or one.

(3) '**Common fate**'. If a subset of elements in the scene change in a correlated way, they tend to be grouped. Thus, elements in a visual scene that move together are grouped together and are segregated from a background that is static or is moving in a

different way. For example, 'motion parallax' allows us to break camouflage by moving our head: elements in different planes move relative to one another (the geometric basis is formally the same as for as binocular stereo vision; see above). In hearing, shared onset times, or shared fluctuation in amplitude, or shared direction of change of frequency promote grouping of different components [14]. Note that the 'law of common fate' could often be taken as a special case of the law of similarity (2) – since nowadays we know that the visual and auditory systems contain detectors for dynamic features of the stimulus, such as direction of movement or direction of change of pitch (see above and Figure 5.11.2B).

(4) **Good continuation**. If a subset of elements in an auditory or visual scene form a simple pattern (e.g. if the elements in a visual scene fall on a straight line or a smooth curve), they tend to be perceived as one object. In hearing, smooth trajectories of frequency (glides) often lead to grouping.

It is often held that the universality of these rules implies that they are innate, inbuilt in our perceptual systems. However, this does not necessarily follow: the rules of organisation reflect the statistical properties of the physical world that we all share. Thus it is logically possible that the rules are learnt.

## 5.11.3 The Identification of Objects

What neural processes allow us to identify a stimulus as a member of the generic class of chairs or of faces or of voices? And do similar processes allow us to identify our favourite chair or the face or voice of our lover?

Many neuroscientists, implicitly or explicitly, subscribe to the doctrine of 'gnostic units', the hypothesis that words, objects, faces, voices are represented in the brain by the activity of individual neurons [14]. The doctrine was explicit in the writings of the eighteenth-century theorist Charles Bonnet [15], and was developed in the twentieth century by Jerzy Konorski [16] and by Horace Barlow [17]. To build a gnostic unit, outputs of feature detectors (see above) at earlier levels might be connected to ever-more specific neurons at successive levels, giving, at the apex of the pyramid, a 'grandmother cell', a cell that responds only when your grandmother is present. Closely linked to this hypothesis is the idea that such cells constitute the 'engram', the (still elusive) physical basis of memory (see Section 5.14).

The problem with the gnostic unit hypothesis is that it implies that the information is frozen in place. In order to distribute to other parts of the brain the information that your grandmother is present, many dedicated fibres would be required, adding to the bulk of the brain's white matter. A conventional alternative is 'ensemble coding', where a given object is represented by the pattern of activity in a population of neurons [14]; but then there is the problem of how that pattern is identified and how the information is distributed to other sites in the brain. What is needed is an abstract representation that can move freely over a 'cerebral bus', where the same neurons carry different information at different times, rather than being dedicated to a particular object or concept [18].

## 5.11.4 The Influence of Context and Experience: Illusions and Hallucinations

Consider the two spoken phrases 'I scream, I yell' and 'ice cream sundae'. The initial acoustic sequences are identical (or, at any rate, they could be made so in a computer-generated stimulus). Yet the after-coming information changes the way that we segment the stimulus (Subsection 5.11.2 above) and changes the words that we hear. In everyday life, in constructing our percepts, we unconsciously draw both on the current context and on our long experience of the physical and social worlds. And often the physical input is the minor contributor to our internal reconstruction of a spoken message or of a three-dimensional scene. Most of the time, our brain gets it right and we are blissfully unaware of the many unconscious assumptions that are being made. But occasionally our expectations mislead us, and we are suddenly aware of an illusion that arises from our mistaken interpretation of the visual or auditory input [19]. If our personal history is different from that of most people, we may perceive malevolence in a facial expression, or in spoken words, that seem neutral to others [3].

When the external contribution to visual analysis is badly degraded (e.g. in macular degeneration or in optic nerve disease) and central visual processes are unconstrained by the input, then vivid hallucinations of people, animals or objects may occur in the absence of psychosis or of cognitive impairment. Charles Bonnet, who first described this rare syndrome in his own grandfather [15], held that the hallucinations reflected activity in the same

fibres in the brain that were normally dedicated to the corresponding percepts (see above, Subsection 5.11.3). Bonnet assures us that his grandfather was entertained by the hallucinations, recognising them to be no more than that: 'His brain is a theatre whose machines perform scenes that surprise the spectator all the more in that they are quite unforeseen'.

# Conclusions and Outstanding Questions

The core processes of perception across the different senses include feature analysis, perceptual organisation and object recognition. These processes are highly interactive, rather than independent and successive. Our resulting percept of the outside world depends as much on context and on our experience as on the current input from our eyes and ears. Illusions and hallucinations may arise from errors at different stages of perceptual analysis, but what is remarkable is how efficiently our senses serve most of us most of the time.

## *Outstanding Questions*

- What neural processes underlie the perceptual organisation of visual and auditory scenes?
- Are words, faces, objects and concepts represented in the brain by the activity of single neurons?
- In what format is information about perceptual objects transmitted to other parts of the brain?
- How are the normal processes of perception hijacked to generate hallucinations?

**5**

### REFERENCES

1. Silverstein S, Keane BP, Blake R et al. (2015). Vision in schizophrenia: why it matters. *Front Psychol* 6: 41.

2. Dakin S, Frith U (2005). Vagaries of visual perception in autism. *Neuron* 48: 497–507.

3. Fletcher P (2021). Visions. In Fabian A, Gibson J, Sheppard M, Weyland S (eds.), *Vision*. Cambridge University Press.

4. Grunert U, Martin PR (2021). Morphology, molecular characterization, and connections of ganglion cells in primate retina. *Annu Rev Vis Sci* 7: 73–103.

5. Masland R (2020). *We Know It When We See It: What the Neurobiology of Vision Tells Us about How We Think*. Basic Books.

6. Snowden R, Thompson P, Troscianko T (2012). *Basic Vision: An Introduction to Visual Perception*, 2nd ed. Oxford University Press.

7. Mollon JD (1977). Neural analysis. In von Fieandt K, Moustgaard IK (eds.), *The Perceptual World*. Academic Press, pp. 71–97.

8. Dakin S, Carlin P, Hemsley D (2005). Weak suppression of visual context in chronic schizophrenia. *Curr Biol* 15: R822–R824.

9. Mollon JD (1974). After-effects and the brain. *New Scientist* 61: 479–482.

10. Chisholm JJ (1869). Colour blindness, an effect of neuritis. *Ophthalmic Hospital Reports*, 214–215.

11. Milner AD, Perrett DI, Johnston RS et al. (1991). Perception and action in 'visual form agnosia'. *Brain* 114 (Pt 1B): 405–428.

12. Paraouty N, Stasiak A, Lorenzi C, Varnet L, Winter IM (2018). Dual coding of frequency modulation in the ventral cochlear nucleus. *J Neurosci* 38: 4123–4137.

13. Shu ZJ, Swindale NV, Cynader MS (1993). Spectral motion produces an auditory after-effect. *Nature* 364: 721–723.

14. Bizley JK, Cohen YE (2013). The what, where and how of auditory-object perception. *Nat Rev Neurosci* 14: 693–707.

15. Bonnet C (1769). *Essai Analytique sur les Facultés de l'Ame*, Volume 2, 2nd ed, Cl. Philibert.

16. Konorski J (1967). Some new ideas concerning the physiological mechanisms of perception. *Acta Neurobiol Exp* 27: 147–161.

17. Barlow HB (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1: 371–394.

18. Danilova MV, Takahashi C, Mollon JD (2020). How does the human visual system compare the speeds of spatially separated objects? *PLoS One* 15: e0231959.

19. Gregory RL (1973). The confounded eye. In Gregory RL Gombrich EH (eds.), *Illusion in Nature and Art*. Duckworth, pp. 49–95.

## National Neuroscience Curriculum Initiative online resources



This "stuff" is really cool: Halide Bilge Turkozer, *"Windows to The Soul" on visual. perception and psychosis*