CHAPTER 1: INTRODUCTION TO THE SENSES

Introduction

From a subjective standpoint, there seems to be little to explain about perception. Our perception of the world is direct, immediate, and effortless, and there is no hint of any intervening operations taking place in the brain. The apparent simplicity of perception is reinforced by the fact that our perceptions are almost always accurate. We rarely make mistakes when identifying people by their face or voice, or in judging how hot a cup of tea is, or in navigating a flight of steps. Moreover, our own perceptions nearly always agree with those of other people. Sounds, sights, and smells seem to be ‘out there’ in the world, not constructed in our head.

Yet our perceptual world is constructed in the brain, by a huge mass of neurons performing complex, but hidden operations. Three observations hint at the complexity of the brain processes involved in perception. First, a large proportion of the brain’s most highly developed structure, the cerebral cortex, is devoted entirely to perception. Vision alone consumes over half of the neurons in the cortex. Second, despite the complexity and power of modern computer technology, computer scientists have not yet succeeded in building general-purpose systems with the perceptual proficiency of even an infant. Relatively confined problems, such as detecting abnormalities in medical images, or identifying a face or a voice, have proven to be formidable problems to solve by computer. Third, as a result of brain damage through injury or disease, a small number of unfortunate individuals suffer deficits in their perceptual capabilities. These deficits can be very specific and debilitating, but also dramatic and perplexing to other people. It seems difficult to believe that someone can fail to recognise their own face reflected in a mirror (prosopagnosia), or cannot judge the position of their limbs without looking directly at them. Such cases remind us of the sophisticated brain processes serving perceptual abilities that most of us take for granted.

Spectator sports provide a very clear example of the reliability, and occasional fallibility of the information extracted by our perceptual systems. Everyone involved - participants, referees/umpires and spectators - must make perceptual judgements in order to interpret events on the sports field, and to decide what should happen next. Did the tennis ball bounce out of court? Did the soccer ball enter the goal net? All those involved nearly always agree on what happened, because their perceptual systems arrive at the same decisions. Sporting activities would not be viable either for participants or for spectators without reliable perceptual systems. Certain critical judgements do require special skills and observation conditions. For instance, the judge who decides whether a tennis ball strikes the top edge of the net during a serve often uses a combination of three senses – sight (deflection of the ball in flight), sound (the impact of the ball on the net), and touch (vibration of the net). As a result, the net judge can detect the slightest contacts between ball and net that are missed by the most or all of the spectators. Disagreements between participants or observers can and do arise, and can offer hints about the nature of the underlying perceptual processes (as well as providing additional entertainment. See figure 1.1).
Common sources of disagreement involve decisions about whether a ball crossed a line on the sports field, such as whether a tennis ball bounced inside a court line, or whether a football crossed a goal line. Participants often reach opposite decisions in ‘close’ calls. This disagreement is not simply a reflection of differences in skill or concentration level, but a natural consequence of the inherent variability in our perceptual decisions. Psychophysical research has taught us that in marginal conditions, perceptual responses are probabilistic. There is a zone of uncertainty spanning the border, so that when a ball bounces slightly to the left of a line, the response of the perceptual system itself will sometimes lead to a “left” response, and other times lead to a “right” response. As a result, different observers are likely to disagree a certain proportion of the time. Perceptual research aims to estimate the precise degree of uncertainty attached to perceptual judgements, and to identify its likely causes.

Classification of the Senses
The senses can be divided into five major groups, as shown in the table below, on the basis of the particular form of environmental stimulation they detect.

<table>
<thead>
<tr>
<th>Sense</th>
<th>Stimulus</th>
<th>Receptor</th>
<th>Sensory Structure</th>
<th>Cortex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vision</td>
<td>Electromagnetic energy</td>
<td>Photoreceptors</td>
<td>Eye</td>
<td>Primary Visual Cortex</td>
</tr>
<tr>
<td>Hearing</td>
<td>Air Pressure waves</td>
<td>Mechanoreceptors</td>
<td>Ear</td>
<td>Auditory Cortex</td>
</tr>
<tr>
<td>Touch</td>
<td>Tissue distortion</td>
<td>Mechanoreceptors,</td>
<td>Skin, muscle, etc.</td>
<td>Somatosensory Cortex</td>
</tr>
<tr>
<td>Balance</td>
<td>Gravity, acceleration</td>
<td>Mechanoreceptors</td>
<td>Vestibular organs</td>
<td>Temporal cortex (?)</td>
</tr>
<tr>
<td>Taste/Smell</td>
<td>Chemical composition</td>
<td>Chemoreceptors</td>
<td>Nose, mouth</td>
<td>Primary Taste Cortex,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Olfactory Cortex</td>
</tr>
</tbody>
</table>

Receptor cells convert environmental energy into electrical nerve impulses. A variety of methods are used to achieve this ‘transduction’, including molecular changes in photoreceptors triggered by light absorption, and mechanical deflection of tiny hairs by fluid currents in the inner ear. Receptors in each sense are connected to cells in different specialised areas of the cerebral cortex of the brain, as shown in the right hand column of the table. The cortex is a crumpled sheet of cells 2.5 mm thick and 1000cm² in surface area (Braitenberg and Schutz, 1991). It contains approximately 20,000,000,000 cells. Figure 1.2 (top) shows a drawing of the human brain, identifying the receiving area for each sense. Activity of cells in these cortical areas is thought to lead to conscious perceptual experience. It is important to note that the figure shows only the receiving areas. Many other cortical areas are also devoted to the senses, by virtue of connections between cortical cells. There are interesting species differences in the total extent of cortical surface devoted to different senses. In primates, including humans, the visual cortex is the largest sensory area in the brain. Figure 1.2 (bottom) shows the relative area of cortex devoted to vision, hearing, and touch in two other species as well as primates.
Auditory cortex is dominant in bats, and somatosensory cortex is dominant in moles. The relative area of cortex devoted to different senses is indicative of their relative importance to the survival of each animal.

**FIGURE 1.2 (as is)**

**Methods used to study perception**

A number of techniques have been used to study perception over the last two hundred years. Each technique has particular advantages and limitations, but no one technique is to be preferred over the others. Different techniques complement each other, so that when considered together they allow us to construct a very detailed picture of how we perceive the world.

**Lesion experiments**

We now know that the cortex can be sub-divided into many areas that are specialised for certain functions, as figure 1.2 has already shown. But in the mid-1800’s, scientific opinion held that the cortex could not be sub-divided in this way. Many believed that sensation, perception, and action were represented diffusely throughout the cortex. Lesion experiments provided some of the earliest evidence against this view, and in favour of localisation of function in the brain. The procedure in such experiments is to surgically remove or destroy a specific area of an animal’s brain, and then observe the consequences for behaviour. If a specific behavioural function is impaired or removed following surgery, then one may infer that the relevant brain area is crucial for the maintenance of that function. However, care is needed if one is to avoid drawing erroneous conclusions from lesion experiments. For example, one of the earliest experiments was performed by David Ferrier (1876). He examined monkeys after removal of an area on each side of the cortex known as the angular gyrus (see figure 1.3).

**FIGURE 1.3 (as is)**

Ferrier concluded from his observations that the animals were completely blind following surgery. One monkey, for instance, was very fond of tea. Ferrier noted that:

“On placing a cup of tea close to its lips it began to drink eagerly. The cup was then removed from immediate contact, and the animal though intensely eager to drink further, as indicated by its gestures, was unable to find the cup, though its eyes were looking straight towards it”.

Later experiments, some of which are described below, indicate that Ferrier was mistaken in concluding from his observations that the monkeys were blinded by the lesion. Blindness is associated with damage to the occipital cortex, not the angular gyrus (occipital cortex is at the very back of the brain). According to Glickstein (1985), Ferrier’s lesions had disrupted visually guided action, not vision itself. The monkey he described above could probably see the cup, but could not perform the actions needed to drink from it. Despite such early mistakes, lesion studies have played an important part in establishing localisation of function as a basic principle of cortical organisation.
**Clinical Studies**

Research on localisation of function in humans has relied largely on clinical investigation into the consequences of accidental damage or disease to specific brain areas. The usefulness of these studies is very similar to that of lesion experiments, in that they allow inferences to be drawn about localisation of function. Some of the earliest work to establish the importance of the occipital cortex for vision was undertaken by Tatsuji Inouye in the early 1900’s. Inouye was a Japanese army physician, who studied soldiers wounded during combat in the Russo-Japanese war. His job was to assess their degree of blindness following bullet wounds to the head, as this determined the size of their pension (see Glickstein and Whitteridge, 1987). Inouye devised an instrument to locate precisely in three-dimensions the position of entry and exit wounds (see figure 1.4).

**FIGURE 1.4 (re-draw)**

Assuming a straight path for the bullet, he was then able to identify the brain areas damaged, and relate them to the impairments observed in the soldiers. Inouye was among the first to show that the visual field is mapped in a highly ordered way on the surface of human occipital cortex (see below).

Clinical studies of the consequences of brain damage are necessarily more untidy than lesion studies, since the researcher has no control over the location and extent of the damage. As a result, the inferences that can be drawn from clinical studies are limited. However, clinical studies have led to many important discoveries concerning localisation of function.

**Single-unit recordings**

Although a great deal was known about anatomy and about localisation of function prior to the 1950's, nothing was known for certain about how individual nerve cells contributed to sensory processing. As David Hubel (1988) remarked:

“...I can well remember, in the 1950’s, looking at a microscopic slide of visual cortex, showing the millions of cells packed like eggs in a crate, and wondering what they all could conceivably be doing…”

Theories of perception were inspired largely by anatomy. The brain was known to contain huge numbers of cells, massively interconnected (but only over short distances) in circuits that are similar over the whole cortex. As we have seen, studies of localised brain damage showed that the visual cortex was mapped topographically. These facts inspired the Electrical Field Theory of perception. Visual patterns were thought to set up corresponding fields of electrical activity across the surface of the cortex. Perceptual organisation in complex displays was said to be governed by interactions between fields of current extending across the cortical surface. Experimental tests of the theory included attempts to short-circuit the electrical fields by pinning metallic strips across the surface of the cortex in rhesus monkeys, and then performing tests of visual functioning (eg. Lashley et al., 1951).

In the early 1950’s, Stephen Kuffler was among the first to use a new microelectrode recording technique to monitor the activity of single sensory cells. He inserted electrodes (very fine insulated wires) through the white of the eye in an awake, anaesthesised cat,
and was able to record activity generated in individual retinal ganglion cells by simple visual stimuli placed in front of the animal. Kuffler’s (1953) work on the cat retina, along with work by Barlow (1953) on the frog retina, and by Hubel and Wiesel (1959) on the cat visual cortex, provided the first detailed information on the stimulus preferences of individual sensory cells. We now know that, despite anatomical uniformity, functional properties vary hugely from cell to cell. For example, some retinal cells prefer small light spots of light, while others prefer large dark spots. In the cortex, individual cells are highly selective for line orientation, movement direction, colour, size, and so on (see figure 1.5).

**FIGURE 1.5 (as is)**
The key word is specialisation rather than uniformity of function. These discoveries led to theories of pattern recognition based on neural ‘feature detectors’. As we shall see in later chapters, this view of single cells as feature detectors is rather too simple. One must also be wary of drawing conclusions about the functioning of a huge mass of neurons on the basis of responses in single units. Nevertheless, single-cell recording data have had a profound influence on theories of perception.

[Incidentally In partnership with Torsten Wiesel, David Hubel performed a series of ground-breaking experiments based on single-cell recordings from cells in the visual system of the cat. They were later awarded a Nobel prize for these discoveries]

**Brain Imaging**

Brain imaging techniques were developed in the 1970’s, primarily for use in medicine. The earliest technique to be developed was computerised tomography (CT). The subject is placed bodily in a long thin cylindrical tube (see figure 1.6).

**FIGURE 1.6 (as is + Parkin, 2000, bottom p. 23 or similar)**
X-ray emitters and detectors are positioned around the circumference of the tube. A highly focussed X-ray beam is emitted from one side of the cylinder so that it passes through the subject’s body before being collected by detectors at the opposite side. X-rays are passed through the head from many directions around the tube. From the resulting pattern of X-ray transmission, sophisticated data analysis procedures can build up a detailed picture of the different structures inside the head, as shown in the figure. CT scans reveal areas of brain damage, and are therefore particularly useful in combination with clinical investigations into the behavioural consequences of brain damage.
Magnetic resonance imaging (MRI) scanners detect the magnetic properties of brain molecules, revealed by passing radio waves through the head in all directions. Functional MRI (fMRI) scanners detect minute magnetic changes in haemoglobin induced by variation in blood oxygen concentration. Since variation in blood oxygen concentration is related to neural activity (activity consumes energy) fMRI scans can inform us about brain function. The primary inferences from brain scanning data concern localisation of function. Studies using fMRI scans often compare scans obtained while the subject is performing different tasks, in order to identify the brain areas that are associated with those tasks. Brain imaging is expensive and technically complex, and the data obtained require careful interpretation. However, the technique is likely to grow in importance as its use becomes more widespread, and data analysis techniques become more sophisticated still.
**Psychophysics**

Psychophysics is the scientific study of relationships between physical stimuli and perceptual phenomena. A typical psychophysical experiment involves carefully controlled stimuli, usually presented by a computer, and highly constrained responses from adult human observers. Figure 1.7 shows a typical psychophysical stimulus, presented on a computer monitor.

**FIGURE 1.7 (as is)**

In this example, the stimulus is designed to study the subject’s ability to discriminate small differences in stimulus contrast (the difference in intensity between bright and dark bars). The subject is given a response pad containing two buttons, labelled ‘left’ and ‘right’. He/she is instructed to press the button corresponding to the stimulus that appears to have higher contrast. The contrast difference between the stimuli is manipulated to find the difference at which the subject achieves the required level of accuracy. A number of experimental techniques have been developed over the last hundred years to ensure that data obtained in psychophysical experiments are not contaminated by uncontrolled variables such as subject expectations or desires. In the example, the position of the higher contrast stimulus would be varied randomly between left and right from trial to trial, and the contrast difference may also vary randomly, without the subject’s knowledge. The tutorial at the end of this chapter provides an introduction to the major psychophysical techniques, and their theoretical background. Psychophysical experiments are particularly useful for testing predictions from theories of perception. However, inferences about the neural structures mediating performance must be treated with some caution, and require cross-referencing against physiological data.

**Artificial Intelligence (AI)**

In the 1930’s the mathematician Alan Turing developed the notion of universal computation, according to which all sufficiently powerful computing devices are essentially identical. Any one device can emulate the operation of any other device. If we accept that the brain is a form of computational device, then it follows that it can be emulated by other such devices, namely computers. This is the conceptual basis for AI approaches to brain function. But what does it mean to say that the brain is a computational device? The senses send information about the external world to the brain. This information can be measured and specified very precisely in mathematical terms eg. patterns of light and dark in optical images, or sound-pressure waves entering the ear. The response of the brain can also be described in mathematical terms eg. patterns of neural activity, or consistent patterns of behaviour. Since both the input to the brain and its output can be expressed mathematically, AI researchers attempt to develop formal mathematical rules (computations) that transform one into the other. The computations are assumed to emulate brain processing. A more detailed discussion of the concept of computation as it is applied to the brain can be found later in the chapter.

**FIGURE 1.8 (as is)**

In the example in figure 1.8, a visual scene is captured by a camera and converted into a matrix of numbers. Each number represents the intensity of a specific point in the image. The computer performs operations on these numbers to create an output matrix that is converted back into an image. The computation in the example attempts to find the edges
of any objects present in the scene. Any edges found appear as very light or dark marks in the output image.

AI research has made a major contribution to our understanding of the nature of sensory information, and of the kinds of operations likely to be performed by the brain.

**General principles of sensation and perception**

All of these techniques have been used to study the five major perceptual systems. A number of general principles have emerged from these studies. A brief review of them will serve as a useful introduction to later chapters that cover each sense in much more detail.

**Physiological principles**

**Neural impulses and transduction**

Information in the nervous system is conveyed by trains of electrical signals (neural impulses) passed from one cell to another through the system. These impulses travel from a cell’s dendrites and body to its terminal buttons, typically via an axon. The terminal buttons connect to the dendrites of another cell or cells at synapses. When the impulse reaches a synapse, it causes the release of neurotransmitter chemicals that affect the electrical state of the receiving neuron. The neurotransmitter can be excitatory (eg. acetylcholine, ACh), increasing the probability that the receiving neuron will generate an impulse, or inhibitory (eg. gamma amino butyric acid, GABA), decreasing the probability that the receiving neuron will fire an impulse.

Environmental energy takes a number of forms, as table 1 showed. Each sense requires specialised cells that receive one particular form of energy and convert or ‘transduce’ it into neural signals. The eye, for example, contains photoreceptors, each of which contains photopigments (two examples are shown in figure 1.9). The breakdown of these photopigments when struck by light results in the generation of a receptor voltage that is transmitted to neurons in the retina. The mechanoreceptors of the inner ear contain hairlike outgrowths (cilia). Vibrations initiated by sound pressure waves arriving at the outer ear deflect the cilia and trigger an electrical change in the receptor.

**FIGURE 1.9 (re-draw)**

**Hierarchical processing**

Neural signals generated during transduction are transmitted to several structures in the brain. A common feature of all the senses is that ultimately at least some of the signals arrive at a receiving area in the cortex of the brain, as described earlier and pictured in figure 1.2.

In between transduction and arrival at the cortex, signals from each sense organ pass through a series of synapses at successively higher levels of neural processing. In the case of hearing, for example, there are five synapses on the route from hair cells to cortex. In the case of vision there are three levels of synapse between photoreceptors and brain. In all the senses except olfaction, one of the synapses on the route from sense organ to brain is located in the thalamus (olfactory signals are an exception because they pass directly from olfactory bulb to cortex). After the sensory signals arrive at a receiving area in the
cortex, they are passed on to other cortical areas, often called association areas. Figure 1.10 summarises the successive hierarchical stages characteristic of sensory processing.

**FIGURE 1.10 (as is)**

Arrows in the figure identify the direction of flow of neural signals through the system. Signal flow is unidirectional up to the thalamus, and bi-directional thereafter. Each stage of processing (each box in the figure) contains a large population of cells, often extensively interconnected. The input signal that arrives at each stage is modified by interactions that take place between the cells in that stage. As a result, the output signal that is passed on to the next stage differs in some way from the input signal – it has undergone a transformation during its passage through the processing stage. The successive transformations that occur as the sensory signal progresses through the hierarchy of processing serve to refine the information it contains. For example, useful information is selectively retained and elaborated, while less useful information is lost.

**Selectivity**

Each sensory system responds only to a particular range of stimuli. The human auditory system, for example, responds to sound pressure wave frequencies between 20 Hz and 16,000 Hz. Sounds outside this range are not detectable (though they may be detectable to other organisms; dogs, for example, can detect frequencies higher than 16,000 Hz). The range of effective stimuli for a particular system can be described as its sensory space. Within this sensory space, stimuli can vary along many different dimensions. A single spot of visible light can vary in, for example, its horizontal position in the visual field, its vertical position, its size, its intensity, and its wavelength characteristics. Single-unit recording techniques allow us to take an individual neuron at any one level of processing in a sensory system, and examine the particular range of stimuli within the system’s sensory space to which that cell responds. Single-unit recording data reveal that sensory cells are highly selective in their response. A specific cell in the visual system, for instance, may respond only when a spot of light is presented at a specific location in the visual field, and has a particular size and colour. A change in any one of these parameters causes a reduction in the cell’s response (see figure 1.11).

**FIGURE 1.11 (as is)**

Such selectivity is a universal property of sensory cells. Different cells have different stimulus preferences, so a stimulus change that results in a reduction in one cell’s response is likely to result in an increase in another cell’s response. The limited spatial area of a cell’s response (e.g. an area of the visual field for a cell in the visual system, or an area of the body surface for a cell in the somatosensory system) is usually called the cell’s the receptive field.

**Organisation**

Each neuron at a particular level of sensory processing, whether in the sense organ, the thalamus, or the cortex, has its own specific stimulus preferences. The physical location of cells at each processing level is highly organised with respect to their stimulus preferences. In general, cells that prefer similar stimuli tend to be located near to each other in the brain. The most dramatic examples of organisation are so-called topographic maps. Figure 1.12 shows an example from vision.

**FIGURE 1.12 (as is)**
The upper bull’s eye pattern was presented to a monkey so that the area outlined by the rectangle appeared in the animal’s right visual field (i.e. fixation at the centre of the pattern). The lower image is a map of the left hemisphere of the monkey’s cortex, showing only a portion at the rear of the hemisphere (the area where visual signals arrive – striate cortex). Tootell et al. (1982) used a physiological staining technique to identify which cells in this area of cortex were active while the animal viewed the bull’s eye.

Regions containing active cells are drawn in red. Neurons with receptive fields at nearby retinal positions are clearly located near to each other in the cortex, since the active regions are grouped together. The pattern of activity is so well ordered that it constitutes a map of the pattern of retinal stimulation (often called a topographical map). Notice that the cortical map is distorted. The small region of the image near to fixation (innermost ring of the bull’s eye) occupies a relatively large proportion of the cortical surface (left-hand third of the cortical map). This property of organisation is called ‘cortical magnification’, and is a common feature across the senses.

**Specific nerve energy**

All sense organs generate the same kind of electrical signals, as we have seen. After transduction, there is no feature of the neural signals that marks them as belonging to one of the senses rather than any of the others. How, then, can they evoke different experiences? Differences between the senses are not reflected in the nature of the sensory signals themselves, but in their destination in the brain. As table 1 and figure 1.2 showed, signals in different sensory systems arrive at different cortical receiving areas. It is the destination that marks a particular signal as arising from a specific sense, giving the signal a characteristic sensory quality. Johannes Muller introduced this idea in 1838, and described it as the law of specific nerve energy.

Dramatic and direct support for the idea of specific nerve energy can be drawn from observations made during neurosurgery. The neurosurgeon removes a section of skull to expose an area of the cortex. In order to navigate the dense folds of the cortical surface safely (avoiding damage to important functions), small electrical signals are often applied directly to the cortex while the patient is awake but anaesthetised. This stimulation evokes sensations associated with the particular sense organ connected to that part of the cortex, such as visual sensations.

**Plasticity**

The neural mechanisms that acquire and process sensory information are modifiable during development and during adulthood. As a human infant grows, the dimensions of his or her body change progressively. Limbs become longer and heavier, the eyes move apart. Sensory systems must be capable of adapting to these changes. Over much shorter time periods, each sensory system is also able to adapt itself to the specific sensory environment that the individual find him/herself in. For example, as the sun sets the visual system’s sensitivity changes progressively to match the prevailing illumination level. If one dresses in the morning with a particularly coarse-textured shirt, the initial feeling of itchiness conveyed by touch receptors in the skin soon subsides as the receptors adapt the their new environment.

**FIGURE 1.13 (author to re-draw)**
The graph in figure 1.13 shows the change in response of touch receptors to a steadily applied stimulus over a period of 40 seconds. Adaptability is a universal feature of sensory systems.

**Noise**

The activity level of a neuron can be measured in terms of the frequency with which it generates electrical impulses. Activity level can vary between zero (no impulses at all) to approximately 800 impulses per second, though the typical rate for a very active cell is 100-200 impulses per second. In the example shown in figure 1.13, the initial activity level of the touch receptor was about 100 impulses/sec. Neural signals show a certain degree of variability, even in the absence of adaptation or short-term plasticity. The response to repeated presentation of identical stimuli differs randomly for presentation to presentation. This kind of variability is usually called ‘noise’ because it bears no systematic relation to the incoming stimulation, or signal. There are two sources of variability (White, 2000). First, there are fluctuations in the electrical excitability of neurons, caused mainly by random opening and closing of ion channels. Second, there are fluctuations in synaptic transmission caused by, among other factors, the random nature of diffusion and chemical reaction across synapses. Detection of changes in neural response is crucial, since they reflect changes in the state of the outside world. However, any measure of change in neural response must be made in the face of the inherent variability in the sensory signal. Theories of sensory coding must, as we shall see, accommodate neural noise.

**Perceptual Principles**

**Sensation**

All the senses share one fundamental property – stimulation of the sense organ causes a conscious mental state. For example, we sense “sour” when a mild acid is in contact with the tongue; we sense “sound” when air pressure waves enter the ear; we sense “light” when electromagnetic radiation enters the eye. Such mental states have particular qualitative, experiential, or felt properties such as loudness, pain, or colour (sometimes called sensations or qualia). By their very nature, sensations are private, and accessible only to the person who has them. Most researchers believe that sensations can be regarded as identical to specific brain states or functions of brain states. For example, there is a specific brain state associated with the sensation of the colour red. If one’s sensation of colour changed to, say, green, there would be a corresponding change in brain state. The assumed link between sensations and brain states lies at the very foundation of modern theories of perception, as will become clear below. However, an ‘explanatory gap’ remains (eg. Young and Block, 1996) between the physical world (brain states) and the mental world (sensations). No one has been able to explain precisely how the qualitative nature of sensation can be explained by reference to neural activity.

**Detectability**

There is a very reliable relationship between the intensity of an environmental stimulus and the probability that it will evoke a sensory experience. A graph of the relation
between stimulus level and the probability of detection by an observer (a so-called ‘psychometric function’) always shows a smooth transition between no-detection and detection as stimulus level increases (see figure 1.14) - detection is probabilistic.

**FIGURE 1.14 (author to re-draw)**
This property of perceptual experience is almost certainly related to the ‘noisy’ nature of neural signals mentioned earlier, though there has been some debate as to the precise explanation (see the tutorial on psychophysical methods later in the chapter).

**Sensory Magnitude**
 Variation in intensity affects not only the detectability of a stimulus, but also its perceived magnitude. For example, the brightness of a light, the loudness of a sound, or the heaviness of a weight increases as stimulus magnitude increases. An experimental technique called magnitude estimation allows us to establish the precise relationship between physical stimulus magnitude and sensory magnitude. The subject is initially presented with a standard stimulus at moderate intensity, and asked to assign an arbitrary number to it, such as 100. Other stimuli are then presented one at a time, and the subject is asked to estimate the magnitude of each relative to the standard stimulus using the same numerical scale. If, for example, a stimulus appears twice as intense as the standard (twice as bright or twice as loud), then the subject should assign the number 200 to it. The technique has been applied to a wide range of sensory stimuli. Representative data are shown on the left of figure 1.15. The relationship between stimulus intensity and sensory magnitude is not linear.

**FIGURE 1.15 (author to re-draw)**
In some sensory judgements, such as brightness, sensory magnitude increases rapidly at low stimulus intensities, but flattens off or saturates at higher intensities. In others, sensory magnitude shows the opposite pattern. If the data are plotted on logarithmic axes rather than linear axes, they fall along straight lines in the graph (right of figure 1.15, right). This means that sensory magnitude data conform to a power law, in which sensory magnitude grows in proportion to stimulus intensity raised to a power (the slope of the each line in the logarithmic graph corresponds to the power to which intensity must be raised for that sensation). This property of sensory magnitude data is known as ‘Stevens’s power law’ (eg. Stevens, 1961). The power law relation between stimulus intensity and sensory magnitude means that equal ratios of intensity correspond to equal ratios of sensory magnitude. For example, each time light intensity increases by a factor of six, brightness increases by a factor of two, at all levels of intensity. It seems that the sensory systems provide information about changes in the level of stimulation rather than about the absolute level of stimulation.

**Adaptation**
The relation between sensory magnitude and stimulus level is not fixed, but varies so that all sensory systems can adapt to prevailing stimulus conditions. Continuous exposure to a relatively intense stimulus usually has three consequences for sensation. First, sensitivity changes so that a more intense stimulus is required to induce a perceptual response after adaptation than before adaptation. In figure 1.14, for example, the psychometric function shifts rightward along the intensity axis as background intensity increases. Second, the apparent intensity of the stimulus diminishes, as shown in figure 1.16.
FIGURE 1.16 (author to re-draw)

Third, the rate at which sensory magnitude increases with stimulus level steepens. Given that the response capacity of each sensory system is limited, in that it can only respond to a certain range of stimulus levels at any one time, adaptation ensures that this restricted response is well matched to the prevailing stimulation.

**Theoretical Principles**

**Representation**

We have seen that certain forms of environmental stimulation cause activation along specific neural pathways from sense organ to brain. At some point in this process, if we accept that there is a causal link between neural activity and consciousness, the activity evokes perceptual experience – we see, hear, feel, taste or smell something. Although the world appears to be ‘out there’, it is in fact a pattern of neural activity evoked in our head during the act of perceiving. As Boring (1950) noted:

“The immediate objects of the perception of our senses are merely particular states induced in the nerves.”

A specific internal state of the brain, in the form of a particular pattern of neural activity in some sense *represents* the state of the outside world. Perception must involve the formation of these representations in the brain. Most modern theories of perception are in essence theories about how the brain builds and uses representations of the world. Earlier in the chapter we discussed the idea that neural signals in sensory systems pass through a series of processing stages. According to the notion of representation, each of these stages must contain a representation of the state of the world. The transition through a series of neural processing stages can be viewed as a transition through a series of internal representations.

What do we mean by ‘representation’? The idea that the state of one physical system (eg. the brain) can in some sense represent the state of another system (eg. the world) is very general, and can be applied to many systems. For example, the reading on a thermometer represents the current temperature; the display on a wristwatch represents the current time. As time moves on, the watch’s display changes accordingly. The analogy with wristwatches is worth pursuing further, because it illustrates a distinction between two basic forms of representation, analogue and symbolic. There are two representational styles available in wristwatches, analogue and digital. In analogue watches the current time is represented in graphical form using a clock-face. The pattern displayed on the clock-face (the positions of the hands relative to hour and minute markers) provides a *pictorial* representation of the time. Marks around the clock face represent the division of each day into hours, and each hour into minutes, and so on. The continuous movement of the second hand is analogous to the passage of time, and the circular design of the display is analogous to the periodic nature of measured time. Digital watches, on the other hand, represent the current time in purely *symbolic* form using a sequence of digits. The digits 0 to 9 are used to represent the current hour, minute and second. There is no continuous change to represent the passage of time. Instead there is a discrete change from one digit
to another every second or every minute. The difference between analogue and digital forms of representation becomes clearer when one considers how to extract time information from them. To establish the time displayed on an analogue clock-face, one must know how to interpret the spatial pattern. For example, the angle between the minute hand and the twelve o’clock mark represents the number of minutes left in the hour. To understand the time displayed on a digital watch, one must be able to understand and manipulate symbols. For example, to obtain the number of minutes left in the hour, one must subtract the minute digits from sixty. Digital watches therefore use a discrete, symbolic representation of time.

As a second example of the distinction between analogue and symbolic representation, consider how a bird-spotter’s handbook might represent a specific species of bird. The entry for that species may contain a still image of the bird, and a text list of its attributes, as illustrated in figure 1.17.

**FIGURE 1.17 (re-draw?)**

A multimedia text may also contain an audio clip of the bird’s call, and a movie clip of the bird’s flight. The image, audio clip, and movie clip are clearly analogue representations, since they represent the bird in terms of its visual or audible pattern. The text list of attributes is an abstract symbolic representation of the bird.

There are two contrasting views about the nature of perceptual representation in the brain, that relate quite closely to the distinction between analogue and symbolic representations in the bird-spotter’s handbook. According to one view, perceptual representations are ‘analogue’ representations, in which the pattern of internal activity reflects in a very direct sense the pattern of external stimulation. For instance, the pattern of activity in the sheet of photoreceptors in the retina of the eye can be said to represent the pattern of light and dark in the image. The matrix of numbers in figure 1.8 depicts the pattern of activity generated by a small region of the image. Intensity at a specific point in the image is represented in terms of a numerical quantity varying between 0 and 255. The numerical quantity is equivalent to the level of activity in a specific sensory cell.

According to the alternative view, perception involves ‘symbolic’ representations, in which the state of the world is described using a limited set of abstract symbols. Each symbol would represent a perceptual object or a property of an object, in the same way that words themselves symbolise objects and properties of objects. For example, there may be a perceptual symbol for ‘creature’ that can be set to equal one of a restricted number of types such as biped, quadruped, bird, fish, etc. This symbol itself may posses a number of symbolic properties such as ‘size’ (small, medium, large), and ‘flight’ (undulating, direct, flitting, gliding), as shown in figure 1.17. The nature of the representation adopted by a sensory system determines what kind of computations the system can perform on the representation.

**Computation**

The concept of computation lies alongside the concept of representation at the heart of most present-day theories of perception. In an abstract sense, computation can be defined as the manipulation of quantities or symbols according to a set of formal rules. It follows from this abstract definition of computation that a neural process that produces a perceptual quantity such as brightness, or a perceptual symbol such as an object property, can be described as a computational process. The formal rules used in computations are
sometimes called algorithms. The idea that neural processing is a form of computation originated from the work of the mathematician Alan Turing in the 1930’s, as mentioned earlier. According to Turing, the brain can be considered as a computing device, in the sense that it manipulates quantities and symbols according to sets of rules. How exactly does the concept of computation apply to perception? We have seen that perceptual systems can be considered as *representational* systems – internal brain states represent the state of the outside world. Perceptual analysis proceeds through a series of representations, reflecting a series of neural processing stages. The representation at each level is transformed into a new representation at the next level by a computational operation, as depicted in figure 1.18 (a modification of figure 1.10).

**FIGURE 1.18 (as is)**

The nature of the computation that transforms one representation into the next depends on what form the two representations take, analogue or symbolic. Computations performed on analogue representations involve the creation and manipulation of *quantities* according to a set of rules, sometimes called signal processing. The computations involve mathematical manipulations of the values stored in the original representation. For example, given the analogue representation of the sculpture in figure 1.8, containing many different lightness levels, one might wish to produce a more economical representation that contains only the sharp edges present in the image. Appropriate computations can find the edges, as shown in figure 1.8. As we shall see in later chapters, this kind of computation captures the essential properties of many cortical cells. Computations performed on symbolic representations involve the creation and manipulation of *symbols* according to a set of rules. The computations involve comparisons between symbols to test for equality, and the combination of symbols to create new symbol structures. For example, the first representation may contain the symbols illustrated in figure 1.17 (Size = Small; Flight = Undulating; etc.). The perceptual system may contain a rule that states that (IF Size = small AND Flight=undulating AND… THEN Bird = Goldfinch). An instance of the symbol for ‘Goldfinch’ would be created at the next level of representation in the processing hierarchy.

Symbolic representations and computations have traditionally been associated with human cognition, such as problem solving (Newell and Simon, 1972), and seem a natural choice for high-level perceptual representations relating to object identity. Early perceptual representations, such as those in sense organs and cortical receiving areas, are probably best considered to be analogue in form. It is not yet clear where and how perceptual representations shift from analogue to symbolic. Some theorists have proposed that symbolic perceptual representations are used even at the earliest levels of analysis (Marr, 1980).

**Linking Propositions**

Contemporary theories of perception attempt to describe the relationship between the physical state of the brain (patterns of activity in certain groups of neurons) and the mental state of the perceiver (sensations or perceptions). Brindley (1960) recognised that any rigorous theory should express this relationship in terms of explicit propositions that he called psychophysical linking hypotheses. More recently, Teller (1984) defined a linking proposition as
“...a claim that a particular mapping occurs, or a particular mapping principle applies, between perceptual and physiological states. “ (Teller, 1984, p. 1235).

Rigourous theories of perception usually contain at least one linking proposition of this kind. An example of such proposition is that the loudness of a sound is coded by the rate of firing of certain cells in the auditory system. Linking propositions provide a bridge between the quantities and symbols that are computed by neural processes, and the perceptual experiences they evoke. Many perceptual theories do not spell out explicitly their linking propositions, but it is important to be aware that such propositions must form part of any theory that attempts to relate neural events to perceptual events.

**Decision Rules**

Perceptual judgements usually involve a decision of some kind. For example: Is the cup within reach? Did I hear the telephone ring? Many of these decisions must be made in the face of unreliable, incomplete, or inconsistent information. Judgements of distance, for example, can make use of a range of cues including perspective, stereo vision, and relative movement. Each cue may supply a different estimate, but the perceptual judgement must combine them in some way to arrive at a single estimate. To take a second example, a high level of ambient background noise in the environment may mask the sound of the telephone. In addition to such external factors, we have already seen that the internal neural signals mediating perception are subject to fluctuation, so the computations performed during sensory processing must accommodate this variability. The most complete theories of perception therefore incorporate rules governing how decisions are made in the presence of ambiguity, noise, and multiple sources of information. These rules are often based on the mathematics of probability and statistics.

**Chapter Summary**

Perception involves highly complex neural processes that consume a substantial proportion of the brain's cerebral cortex.

**Classification of the senses**

There are five major groups of senses:

- Vision
- Hearing
- Touch
- Balance
- Taste/Smell

Senses differ in terms of the environmental stimuli that excites them, and the neural structures involved.

**Methods used to study perception**

Methods include:

- Lesion experiments
- Clinical studies
- Single-unit recordings
Each method has advantages and drawbacks. Some acquaintance with results obtained using different methodologies is essential for an adequate understanding of perception.

**General principles of sensation and perception**

A number of unifying principles have emerged from studies of sensory systems. Physiological principles include:

- Neural impulses and transduction
- Hierarchical processing
- Selectivity
- Organisation
- Specific nerve energy
- Plasticity
- Noise

Perceptual principles include:

- Sensation
- Detectability
- Sensory magnitude
- Adaptation
- Theoretical principles include:
  - Representation
  - Computation
  - Linking propositions
  - Decision rules

Specific examples of these principles can be found in the detailed coverage of the senses in later chapters.

**Tutorials**

**Psychophysical Methods**

As we saw earlier in the chapter, certain physical stimuli evoke perceptual experiences ranging from simple sensations such as ‘redness’ or ‘loudness’ to complex perceptions such as face recognition. How can we study the relationship between physical stimuli and perceptual experience? The simplest method is to use verbal reports, such as “it looks red” or “that is my grandmother”. This phenomenological approach is severely limited, for several reasons. First, it obviously requires subjects who can describe their experiences in words, so excludes infants and animals. Second, even when restricted to subjects who can talk, it is contaminated by differences in the way different people use words. Third, it is open to bias introduced by individual expectations and desires.

We need precise, accurate measures of perception that can be used to establish the limits of perceptual ability, to monitor how these limits change with stimulus conditions, and to test the predictions of perceptual theories. Ideally these measurement methods should be immune to the effects of verbal ability, expectation, and attitude. Over the last 100 years or so a body of experimental techniques have been developed to provide the required
measurements. Since these techniques provide quantitative, physical measures of psychological phenomena, they are called psychophysical methods.

**Psychometric functions**

Any plot relating a quantifiable response to a physical stimulus measure is known as a psychometric function. One might plot, for example, sound intensity against the probability that the subject will detect the presence of the sound. What is the typical shape of a psychometric function in a detection experiment? One might expect that below a certain stimulus level the sound is never heard, and above it, the sound is always heard. – a step function as shown in the figure. Real psychometric functions always show a gradual shift from no-detection to detection as stimulus level increases, rather than a sudden shift (as shown earlier in the chapter in figure 1.14). Why?

**Classical psychophysical theory and the psychometric function**

The concept of the threshold is crucial to classical psychophysical theory. A threshold marks a transition from one perceptual experience to another, usually as a result of a simple change in the physical stimulus. For example: How intense must a sound be for us to detect it? How fast must something move for us to see the movement? How different in distance must two objects be for us to tell that one is nearer? There are two kinds of threshold, the *absolute threshold* and the *differential threshold*. The absolute threshold marks the smallest amount of stimulus energy required for an observer to just detect its presence (e.g. the minimum sound intensity or movement velocity required for detection). The differential threshold marks the minimum change in stimulus energy that can be detected by an observer. This threshold is also known as the ‘just noticeable difference’, or JND (e.g. the small change in sound intensity required for the observer to notice a change in loudness). Classical psychophysical methods were basically developed to measure thresholds accurately and reliably.

Classical psychophysical theory explains smooth real-world psychometric functions (as in figure 1.14) with the following three assumptions. First, there is an ideal threshold function that relates the internal response of the sensory system (‘sensory magnitude’) to stimulus level. This function is a step function with two levels, ‘low, and ‘high’. Second, when the internal response is ‘high’, the observer always reports detection of the stimulus, and when the internal response is ‘low’, the observer never reports detection of the stimulus. Third, the exact position of the threshold in relation to stimulus level is subject to some random fluctuation, due to momentary variations in neural sensitivity, arousal level, and so on. Although the threshold tends, on average, to cluster around a specific stimulus level, it occasionally falls below or above this level, so that the probability that the threshold will fall at a particular stimulus level conforms to a bell-shaped curve or normal distribution, as in figure 1.19 (left).

**FIGURE 1.19 (as is)**

In the figure, at a low stimulus level (top-left graph), the probability that the threshold will be lower than this level is small (arrowed area), so detection rates are low. As stimulus level increases, the likelihood of detection improves because there is a much greater probability that the threshold will be lower than the stimulus level (lower-left graph). Consequently, if we plot probability of detection against stimulus level, a typical psychometric function is obtained (right-hand graph). At what stimulus level is threshold
reached? According to classical theory, the ‘true’ threshold coincides with the mean of the probability distribution in figure 1.19. Since, by definition, 50% of the distribution lies below the mean, and 50% lies above it, the most logical place on the psychometric function to locate the threshold is at the 50% point. This account of thresholds applies to both absolute and differential thresholds.

**Classical psychophysical methods**

All classical methods aim to measure the observer’s threshold. Some methods provide an estimate of the whole psychometric function. Others provide an estimate of just one point on the function, usually the 50% point. A number of classical methods were developed at the turn of the nineteenth century, but this tutorial will describe the only two methods that are still in use, the method of adjustment and the method of constant stimuli.

In the method of adjustment, the observer is given control of the stimulus (eg. a dial that controls stimulus intensity), and asked to adjust it until it is just detectable. This method is quick and easy to use, but rather unreliable. The observer has direct control of the stimulus, so is free to apply some degree of bias to their settings. Some observers may try to impress with their high sensitivity, and tend to bias dial settings toward low stimulus levels. Other observers may prefer to be cautious and careful, tending to bias their settings toward high stimulus levels.

In the method of constant stimuli, the experimenter selects a range of stimulus levels at the start of the experiment. These different levels are presented to the subject repeatedly in random order, in a series of experimental trials. After each presentation, the subject is required to respond “yes” if the stimulus (or a difference between stimuli) was detected in that trial, or “no” if it was not detected. This method is more trustworthy than adjustment, since the subject has no direct knowledge of the stimulus level presented. It constructs the full psychometric function, so is reliable but more labour-intensive than the method of adjustment. Computers can be used to take care of stimulus selection, increasing the efficiency of the method.

**The problem of bias in classical methods**

In classical psychophysics, the subject’s response to the stimulus is assumed to depend only on their sensitivity, the stimulus level at which the internal response shifts from low to high. However, responses are also likely to reflect uncontrolled bias effects. The problem is most severe using the method of adjustment, but may also intrude in the method of constant stimuli. Since a stimulus is presented in every trial, the observer is free to apply some degree of bias to their responses. The observer may, for example, be feeling uncooperative or lacking in confidence, and so unwilling to respond “yes” unless they are very confident of being correct. As a result, the measured threshold will not be a pure estimate of the subject’s sensitivity to the stimulus, but will reflect some unknown combination of sensitivity and bias. Signal detection theory was developed specifically to address the problem of bias effects.

**Signal detection theory (SDT)**

SDT acknowledges the importance of bias effects by assuming that stimulus detection is a two-stage process (figure 1.20, top). The first stage is a purely sensory process in which a specific stimulus level produces an internal sensory response that depends on the intensity
of the stimulus and the sensitivity of the sensory system. This internal response is subject to random internal ‘noise’ of the kind described earlier in the chapter. The second stage is a decision process in which the sensory response magnitude is compared to an internally set criterion. If the response magnitude exceeds this criterion, the decision process decides that a stimulus was present. If the internal response falls below the criterion, then the decision process decides that no stimulus was present. The position of the criterion is influenced by all the factors described earlier that affect bias. Highly motivated subjects may adopt a low criterion, reflecting a bias in favour of accepting rather weak stimuli. Subjects who lack confidence in their judgements may adopt a high criterion, because they are biased toward accepting only relatively intense stimuli. The experimenter is interested primarily in the sensitivity of the sensory system, rather than the subject’s bias, but SDT provides methods of estimating both sensitivity and bias.

FIGURE 1.20 (as is)

SDT methodology: Yes/no and forced-choice tasks

In classical psychophysical methods, every stimulus presentation in the experiment contains a stimulus. In SDT methods only half of the presentations contain stimuli, randomly selected. For example if the subject is required to detect the presence of a visual pattern against a uniform background, then only half of the presentations contain the pattern and background, while the other half contain only the background. Presentations containing a stimulus are called noise+signal presentations, for reasons that will become obvious, and presentations not containing a stimulus are called noise presentations. The subject must discriminate between noise+signal presentations and noise presentations. Two kinds of task are commonly used. In a yes/no task, the subject is presented with a single stimulus event in each experimental trial, which may or may not contain a signal. The subject must respond “yes” if he/she decides that a stimulus was presented in that trial, and “no” otherwise. In a forced-choice task, the subject is usually presented with two stimulus events in each trial, side-by-side or one after the other. In a vision experiment, for example, two stimulus patches may be presented side-by-side. In a hearing experiment, two sounds may be presented sequentially. Only one event contains the stimulus to be detected. The subject must decide which of the two events contained the stimulus, and respond “left” or “right”, or “one” or “two” as appropriate. Tasks of this kind are commonly called a two-alternative forced choice or 2AFC tasks. Notice that in SDT tasks the subject has no direct knowledge of which event contains the required stimulus. This reduces the possibility of bias, because when the stimulus is not detectable the subject is forced to guess as to which event contained the stimulus. However, in yes/no tasks there is a possibility of some bias in favour of “yes” responses, because of a social aversion to saying “no”. Many researchers prefer to use forced-choice tasks wherever possible, because the alternative responses are fairly neutral (Green and Swets, 1966).

SDT measures of sensitivity and bias

This brief description of SDT measures is based on a yes/no task, but also applies (with appropriate modifications) to forced-choice tasks. SDT theory assumes that both noise+signal and noise events generate an internal response in the sensory process of the detection system, because this process is subject to internal noise (figure 1.20). Noise
events reflect only the contribution of internal noise to the response. Noise+signal events reflect contributions from both internal noise and external stimulation. The probability distribution of the response to each event can be plotted, as shown in figure 1.20. Each distribution simply plots the relative probability of that event generating a specific response magnitude. The noise distribution reflects only the variable level of internal noise, that tends to cluster around a mean value (the peak of the distribution). The noise+signal distribution contains contributions from both internal noise and external stimulation. The effect of the external stimulus is to add a constant value to the noise distribution, displacing it towards higher response magnitudes. The distance over which noise+signal distribution is shifted relative to the noise distribution depends on the system’s sensitivity to the stimulation. The difference between the means of the two distributions is taken as a measure of the sensitivity of the system to the stimulus, and is known as d’ (“d-prime”).

In any one trial of a yes/no task, the decision process receives a response at a particular magnitude, and must decide whether that response was drawn from the noise distribution or from the noise+signal distribution. SDT assumes that the decision process selects a specific criterion level of response, shown by the arrow in figure 1.20. Response levels below this value are deemed to belong to the noise distribution, so are assigned a ‘no’ response. Response levels above this level are deemed to belong to the noise+signal distribution, and are assigned a ‘yes’ response. The level at which the criterion is set depends on biasing factors. It may be ‘unbiased’, or midway between the two distributions, or biased in one direction or the other.

SDT provides various methods for making precise estimates of sensitivity or d’ independent of criterion level or bias (also known as • ). However, for unbiased experiments such as those employed 2AFC tasks, a simple measure of sensitivity is given by the proportion of correct responses recorded by the subject. Readers interested in the mathematical details of SDT measures are referred to Stanislaw and Todorov (1999), who provide formulae and procedures for performing the calculations using general-purpose software such as spreadsheets.

**Evaluation**

SDT was first applied to psychophysical problems by Tanner and Swets in the mid-1950’s and, as we have seen, it discards the classical notion of the threshold in favour of d’. Fifty years later, despite the widespread acceptance in the scientific community of many of the ideas in SDT, much contemporary research still measures performance in terms of thresholds rather than d’ (Gordon, 1997). Why should this be so? Thresholds are still a very useful, and intuitively meaningful way of summarising the performance of a subject, reflecting the stimulus level that is just detectable by the subject. By contrast, d’ is a more abstract measure of sensitivity to a specific stimulus level, and is meaningful only if one appreciates the statistical concepts that underlie it. Despite the continuing attachment to thresholds, many researchers measure them using percentage correct responses in 2AFC tasks, having taken on board the concerns about bias effects raised by advocates of SDT.
Theoretical Approaches to Perception

The previous tutorial on psychophysical methods introduced some of the techniques that have been developed for collecting perceptual data, and the rationale behind them. This tutorial discusses the major theoretical movements that have motivated psychophysical experiments over the last 150 years. We must first define the essential properties of a theory are, and discuss how the adequacy of different theories can be assessed. Minimally, any scientific theory worthy of the name must have three properties (Popper, 1960). First, it must provide a framework for organising and understanding the known facts in an economical manner. Second, it must attempt to provide explanations for the facts, or at least suggest causal links between them. Third, it must be capable of generating predictions that can be tested experimentally. If there are two competing theories to account for a particular set of facts, how can one select the theory that is to be preferred? Several criteria can be applied:

1. **Empirical consistency** One can compare the two theories according to their ability to explain the known facts. A theory is not much use if it cannot account for the data.

2. **Logical consistency or computability** If both theories pass the first test, one can judge their relative merits on the basis of logical consistency. Is the reasoning behind each theory tight and logically consistent? If a theory involves computational operations, can these operations be performed successfully? The inclusion of arbitrary (ad hoc) propositions, or computations that are difficult or impossible to implement, diminishes a theory’s attractiveness.

3. **Occam’s Razor** If both theories pass the first two tests, then one can apply the principle of Occam’s Razor, which states that “Entities must not be multiplied beyond necessity”. What this means is that the more parsimonious theory of the two is to be preferred. If a simple theory can explain the data as convincingly as a more complex theory then, other things being equal, the additional complexity is superfluous.

4. **Generality** A final test of two competing theories concerns their generality. Some theories appear to exist in a vacuum, successfully accommodating the data they were devised to explain, but with no obvious connection to other phenomena or theories. Other theories attempt to place themselves in a wider context by, for example, making connections with other theories. In these circumstances, the better-connected theory is to be preferred. This criterion selects theories on the basis of higher order logical consistency. Are different theories invented ad hoc to explain phenomena in isolation, or is there some higher order rationale or structure that links different theories together? Examples of such higher order links would include energy efficiency, ecological validity.

If two competing theories cannot be separated on the basis of *any* of the four criteria, the only course of action is to return to the first criterion, empirical consistency. New predictions must be generated from each theory concerning the outcome an experiment, formulated in such a way that (ideally) the results are bound to falsify one of the theories. In principle, the aim of any new theory is to provide the only true explanation for a particular phenomenon. However, it is worth remembering that few theories stand the test of time. Most new theories are ultimately discarded either because of empirical
inconsistency, or because they prove to be unsatisfactory on the basis of one of the other
criteria. Most theorists accept that the best they can hope for a particular theory is that it
will provide a closer approximation to the truth than other available theories. Once a new
theory appears that offers a better way of understanding the facts, then the old theory must
be discarded. This does not mean that theorising is futile and doomed to failure, for two
reasons. First, it would be extremely difficult or impossible to arrive at the truth without
having first absorbed the insights offered by previous theories. As Isaac Newton
remarked: “If I have seen farther, it is by standing on the shoulders of giants” (letter to J.
Locke, 1675). Although Newton’s own theories provided the foundation stones for most
of the sciences, he acknowledged the debt he owed to predecessors such as Galileo and
Kepler. Second, much empirical research would be aimless and trivial unless it was
motivated by the need to test the predictions of new theories.
It should now be clear why it is important to understand some of the major theoretical
movements in the scientific study of perception. As we shall see, each movement has
made a valuable contribution to our understanding of perception. The major theoretical
movements were developed in the context of vision, but the ideas can be taken to apply to
all the senses. Modern theories of perception began with Structuralism 150 years ago.

Structuralist Approach
Structuralism drew inspiration from the chemical decomposition of complex substances
into elements. It proposed that each complex perceptual experience could be decomposed
into a large collection of elementary sensations. Structuralists used introspection to break
down a particular perceptual experience into its sensory components. For example,
Tichener (1901) decomposed the taste of lemonade thus:

“The taste of lemonade is made up of a sweet taste, an acid taste, a scent (the fragrance of
lemon), a sensation of temperature, and a pricking (cutaneous) sensation.” (p. xx)

Introspection proved to be an unsatisfactory basis for theories of perception for reasons
that, in retrospect, appear obvious. First, introspective data are inherently qualitative
rather than quantitative. Second, observers frequently disagree in their introspections.
Third, many important perceptual processes cannot be studied by introspection.

Gestalt Approach
Gestalt psychologists rejected the basic principles of Structuralism, and proposed instead
that when a collection of elementary sensations is combined together a new perceptual
entity emerges – a Gestalt. The major exponents of Gestaltism (Wertheimer, Kohler, and
Koffka) were German, and the German word ‘gestalt’ means form, figure, or
configuration. According to Gestalt Psychology, perceptual systems are not passive
recipients of isolated, elementary sensations, but dynamically organise these sensations
into meaningful ‘wholes’ or Gestalts. Gestaltism emphasised the importance of structure
and organisation in perception. It identified a number of organising principles or laws to
describe the variety of ways that perceptual systems achieve organisation. The general
theme of these laws is that isolated elements which share some property in common, such
as spots of the same colour, or shapes that move in the same direction, or notes of similar
pitch, tend to be grouped together perceptually. Elements that form a ‘good figure’
(phagnanz), such as dots falling along a smooth curve or forming an enclose regular shape, also tend to group together perceptually.
The main weakness of Gestalt Psychology was that its laws tended to be descriptive rather than explanatory. Its arguments tended to be circular. For example, Gestalt Psychologists would explain why certain pattern elements group together by invoking the principle of good figure or ‘phagnanz’. But what is the principle of phagnanz? It is the tendency of elements forming a good figure to group together. Despite its limitations, Gestalt Psychology made a valuable contribution to perceptual theory by emphasising the way that entirely new perceptual entities can emerge from the organisation of simpler elements. Gestaltism is no longer at the forefront of perceptual theorising, but is still influential, particularly in European psychology, and is relevant present-day computational theories.

**Constructivist Approach**
The German scientist Hermann von Helmholtz introduced the idea of “unconscious conclusion” in his monumental, three-volume ‘Treatise on Physiological Optics’ published between 1856 and 1866:

“The psychic activities that lead us to infer that there in front of us at a certain place there is a certain object of a certain character, are generally not conscious activities, but unconscious ones. In their result they are equivalent to a conclusion,...it may be permissible to speak of the psychic acts of ordinary perception as unconscious conclusions…” (1962 translation of volume III, page 4)

To expand on this idea, Helmholtz used the example of an astronomer “who computes the positions of the stars in space, their distances, etc.” from his conscious knowledge of the laws of optics. He argued that “there can be no doubt” that perception involves the same kind of computation as that used by the astronomer, but at an unconscious level. Helmholtz went further, stating confidently that:

“Our ideas of things cannot be anything but symbols, natural signs for things which we learn how to use in order to regulate our movements and actions.” (1962 translation of volume III, page 19)

Helmholtz therefore advocated the view that sensory systems construct some kind of internal representation of the world, and that this representation mediates perceptual experience. Related views on the indirect and inferential nature of perception have been promoted by, among others, Gregory (1980), and Rock (1983). It is fair to say that constructivism has had a profound impact on theories of perception. Most modern theoretical approaches rely heavily on the notions of representation and computation. Helmholtz’s ideas on symbolic representation were remarkably prescient, since they appeared 100 years before Turing laid the foundations of artificial intelligence.

**Ecological Approach**
Perception begins with physical stimulation and ends with perceptual experience. In between the two, according to the Gestalt Psychologists and constructivists, are
sophisticated processes that construct internal representations from the sensory information. Perceptual experience has only an indirect relationship to the sensory data. James J Gibson took the opposite view, in rejecting entirely the need for internal representation. He argued instead that there is sufficient information available in the visual image for unambiguous perception to be derived directly, without the need for intervening processes. He suggested that the brain as a whole “picks up” the relevant information by some kind of “resonance”. Gibson used an analogy with a radio set to explain this idea. Your immediate surroundings are almost certainly filled with low-energy electromagnetic radiation broadcast by TV and radio transmitters. A radio, properly tuned, will be able to pick up some of this information and produce intelligible sounds. Gibson would argue that in this situation all of the components of the radio resonate with the information available in the electromagnetic radiation. There is no need to assume that some internal representation is constructed by the radio. Gibson’s ideas were inspired by his work in aircraft pilot training during the second world war. He noticed that conventional treatments of depth cues were of little practical value, and became convinced that the highly structured patterns of movement pilots view from the cockpit were critical for aircraft control. As a plane comes in to land, surface details in the environment, such as markings on the runway, stream across the image projected into the pilot’s eyes. They radiate out from the point in the image toward which the aircraft is heading, creating an ‘optic flow field’. Gibson correctly deduced that this flow field contains sufficient information to specify precisely where and when the aircraft would make contact with the ground. He argued that this information is somehow picked up directly by the sensory system. Gibson identified other properties of natural images, such as texture gradients, that can be used to specify surface depth, slant, and size. Due to its emphasis on natural images, Gibson’s perspective became known as the ecological approach to perception. Its denial of the relevance of mediating processes also led the label ‘direct perception’.

Direct perception performed a valuable service in identifying some powerful sources of information in visual images. But it drastically underestimated the difficulty of the problem posed by picking up this information. Research on artificial intelligence has shown that the information available in visual images is usually not sufficient by itself to recover unambiguous information about the surfaces and objects that created the image.

Computational Approach

The computational approach was anticipated by Helmholtz, in his analogy between astronomical calculations and perceptual conclusions. As mentioned earlier, firm foundations for the computational approach to perception were later laid by Turing. His notion of universal computation led to the idea that the brain was an information-processing device that could be emulated by other such devices, namely computers. An information-processing device receives input data and performs some processing operation on the data to produce an output. An electronic calculator is a good example of an information-processing device. It receives input data in the form of a sequence of numbers and symbols, and processes this data to produce an output, usually the result of a calculation. In the case of perception, the input is environmental data such as a visual image. The output is perceptual data. Intervening processes transform one into the other. Computational theorists attempt to discover the nature of the intervening processes given
only the input to the system, and the output it produces. To continue the analogy with an
electronic calculator, the task is similar to trying to discover the rules of arithmetic given
only the sequence of numbers and symbols providing the input, and the numbers
produced by the calculator as output.
Modern computational theories of human cognition began with Newell and Simon’s
(1980) accepted the notion that vision is “exactly and precisely an information-processing
problem”. He emphasised the need to define at a very abstract level the nature of the
computations performed by the system. Computational theorists such as Marr test their
theories by attempting to implement them using computer programs. In a typical test, the
computer is given an image and attempts to produce the required output. Experience has
shown that many natural images are inherently ambiguous, in that they do not contain
sufficient information to allow the computer to arrive at a single unique output. Instead, a
range of possible outputs is consistent with the input information. In this situation it
becomes necessary to apply some kind of constraint on the output. The constraint may
involve restricting the test to only certain images. Alternatively, it may take the form of
additional information that allows the computer to eliminate most of the possible outputs.
The additional information may reflect an assumption based on the properties of real-
world scenes.

The computational approach has introduced a high degree of rigour into theories of
perception. The importance of constraints is both a strength and a weakness of the
approach. On the one hand, the need for constraints confirms the validity of the
constructivist and Gestalt perspectives, since it reinforces the need for mediating
processes that bring in their own sources of information. On the other hand, constraints
can limit the generality of a computational theory. If the constraint requires the use of
artificial images, or only a sub-set of natural images, this limits the range of situations in
which the theory can operate. If the constraint involves a specific set of assumptions
about the properties of natural scenes, there is a danger that they will be arbitrary and ad
hoc.

**Neurophysiological Approach**

The results of single-cell recording experiments have revealed a great deal about the
response properties of individual cortical cells, as mentioned earlier in this chapter.
Lennie (1998) remarked:

“Single-unit recording on its own is a weak instrument for discovering what visual cortex
really does, but when harnessed to a theory it is immensely valuable” (p. 924)

Early data from cells in the retina of the frog (Barlow, 1953; Lettvin et al., 1959) quickly
led to the idea that these cells were “bug” detectors – their job was to detect the presence
of small dark spots moving across the image, as would be created by bugs for the frog to
eat. Barlow (1972) developed this idea into a lucid theoretical position now commonly
known as the Neuron Doctrine. He argued that each individual neuron is tuned to detect
the presence of a particular stimulus that looks like a particular object (such as a bug) or
attribute of an object. He also proposed that activity in such neurons ‘directly and simply
causes the elements of our perception’. Recent research on cells in the cortex of monkeys has also been interpreted in terms of feature detection by individual sensory cells. However, it has become clear that the response of an individual cell is not driven uniquely by a single stimulus feature, but by a constellation of features. As a result there has been much dispute about the correct theoretical interpretation of single-cell data. Lennie (1998) argues that the neural representation of a stimulus property is distributed among the responses of a whole group of cells rather than assigned to just one cell. Although progress in relating cortical physiology to perception has been slow, the basic facts about cortical cell responses have exerted a profound influence on theories of perception. The causal link between cell activity and perception is a basic theoretical assumption accepted by most sensory scientists.

Evaluation

In its emphasis on a specific set of issues and ideas, each theoretical movement has made its own particular contribution to our understanding of human perception. The modern theoretical approach, as represented in this text, recognises the limitations of each movement in isolation, and attempts to draw elements from all of the movements. The origins of the four theoretical principles outlined earlier in the chapter should now be clear. The emphasis on representation and computation has its origins in Helmholtz’s constructivist approach. Linking propositions are central to Barlow’s neuron doctrine. Decision rules are an important feature of signal detection theory, originally applied to perception by Tanner and Swets. Yantis (2001) has collected together many of the key papers described in this tutorial, as well as other classic papers in visual perception, and offers an excellent opportunity to study the primary sources that laid the foundations of modern perceptual theories.

The modern amalgam of several previous theoretical approaches has evolved into a distinctive approach of its own that falls within the realm of neuroscience. Whether this new approach should be called ‘computational neuroscience’ or ‘cognitive neuroscience’ is still open to debate. Parkin (2000) advocates a broad definition of the latter, as incorporating “…any research which attempts to integrate issues regarding cognitive function with direct assessments of brain function”. He included sensory functions in this definition.
Figure 1.2 Cortical representation of the senses. Top: Cortical receiving areas in the human brain. Bottom: Total cortical area devoted to three senses in three different animals (re-drawn from Krubitzer, 1995). The broken line identifies the cortical receiving area in Macaque monkey.

Figure 1.3 Site of the lesion in Ferrier’s monkeys (re-drawn from Glickstein, 1985).

Figure 1.4 Inouye’s instrument for tracing the path of a bullet in head wounds suffered by Japanese soldiers (re-drawn from Glickstein and Witteridge, 1987).

Figure 1.5 Single-unit recording. A stimulus is presented to the animal (in this case a visual stimulus) while a fine electrode registers activity from cells in the sensory system. The activity is recorded and analysed by special-purpose equipment, in this case a computer equipped with appropriate hardware and software.

Figure 1.6 CT scanner. The patient lies on a table that can be slid inside the scanner (left). The walls of the scanner are lined with X-ray emitters and detectors. X-rays are emitted from one side of the scanning tube so that they pass through the patient’s body before being registered by detectors on the opposite side. A detailed image of the brain can be constructed from the pattern of X-ray transmission in all directions around the head.

Figure 1.7 A typical psychophysical stimulus, generated by a computer. The subject must select the grating that appears to have the higher contrast between its bright and dark bars. In this case the observer should select the left-hand grating. Responses are usually made by pressing one of two response buttons. The computer records each response before selecting the next stimulus to be displayed. The position of the higher contrast grating varies randomly between left and right from presentation to presentation.

Figure 1.8 Schematic illustration of a computer vision system. A visual scene is captured by a camera and fed into a computer. The scene is converted into a matrix of numbers. Each number represents the intensity of the light at a specific location in the scene. The computer performs computations on the matrix of numbers to create a new output matrix. The output matrix can be converted back into an image, again assuming that the magnitude of each number codes the intensity of light in the output image. In this example the computation has attempted to isolate all the edges of objects in the scene. Values in the output image that depart significantly from mid-grey represent the edges found.

Figure 1.9 Sensory receptors. Left: Visual photoreceptors (a rod on the left, and cone on the right). Middle: Auditory inner hair cell. Right: Somatosensory Pacinian corpuscle. Redrawn from Uttal ( ).

Figure 1.10 Hierarchical stages of sensory processing. Neural signals originate in sensory receptors and pass through a series of processing stages. Each stage consists of a large population of interconnected neurons. Arrows denote the direction of flow of the signals.

Figure 1.11 Selectivity in neural responses. The visual stimulus was a tilted bar that oscillated back and forth repeatedly (left). The upper trace on the right shows the neural
impulses (short vertical lines) recorded from a cat cortical cell by Hubel and Wiesel. Time is plotted horizontally, and arrows represent the two phases in the bar’s movement. The cell responded only when the bar moved up and to the right. When is moved out of the cell’s receptive field (lower trace), then no response at all is recorded.

Figure 1.12 Topographic map in the visual system. The bull’s eye pattern in the upper part of the figure was presented to a monkey so that the area enclosed by the rectangle appeared in its right-hand visual field. The lower part of the figure is a flattened view of the animal’s left cerebral hemisphere. A physiological staining technique highlights any cells that were active while the pattern was being viewed (red areas). The pattern of activity is highly organised, and demonstrates how the animal’s visual field is laid out topographically across the surface of the cortex. Re-drawn from Tootell et al. (1982).

Figure 1.13 Time-course of the response of a pressure receptor to stimuli at three different intensities. Response rate is initially high, but declines steadily over a period of 40 seconds. Re-drawn from .

Figure 1.14 Psychometric function plotting the probability of a ‘yes’ response to the presentation of a small light increment against a uniform background. The horizontal axis represents the magnitude of the light increment, and the three curves represent data obtained at three different background intensities. Re-plotted from Mueller (1951).

Figure 1.15 The relationship between stimulus intensity and sensory magnitude. The left-hand graph is plotted using linear axes – sensory magnitude increases nonlinearly at different rates in different senses. The right-hand graph shows the same data plotted on logarithmic axes – sensory magnitude now increases linearly, showing the power-law relationship between stimulus intensity and sensory magnitude. Exponents of the plots were taken from Stevens (1961).

Figure 1.16 Adaptation to an odour. Sensory magnitude of an odour sensation was measured at regular intervals during a 12-minute exposure, showing adaptation to a constant stimulus. Once the stimulus is removed, sensory magnitude gradually recovers to former levels. Re-plotted from Eckman (1967).

Figure 1.17 Entry for a Goldfinch in a birdspotter’s handbook. The pictorial image of the bird constitutes an analogue representation, while the list of attributes on the right constitutes a symbolic representation.

Figure 1.18 Representation and computation in relation to the hierarchical processing scheme depicted in figure 1.10. Each processing stage contains its own representation of the sensory stimulus. The modification that takes place from one processing stage to the next can be considered as a computational operation that transforms one representation into another.

Figure 1.19 Explanation of the empirical psychometric function, according to classical psychophysical theory. The stimulus level at which sensory response reaches threshold is subject to some degree of random variation (left-hand graphs). A low intensity stimulus
(eg. 1.0 in the upper-left graph) is unlikely to be detected (probability 0.1), because only rarely does the threshold drop to such a low stimulus level. A high intensity stimulus (eg. 4.0 in the lower-left graph) is very likely to be detected (probability 0.9), because most of the time the threshold is lower than this level. As a result detection rates improve gradually with stimulus level (right-hand graph).

Figure 1.20 Signal detection theory (SDT). Top: Two hypothetical stages in detection, according to SDT. Middle: According to SDT, both stimulus-absent (“Noise Only”) and stimulus-present (“Noise+Signal”) trials generate a response in the sensory process of the detection system. Each response is subject to some random variation due to internal noise, shown by the two distributions. The observer’s sensitivity to the stimulus is characterised by the difference between the means of the two distributions. Bottom: The decision process receives a response from the sensory process, and must decide whether the response came from the Noise Only distribution or from the Noise+Signal distribution. A specific response level is selected (‘criterion’), above which the decision is that the response came from the Noise+Signal distribution.