1 The perceptual basis of spatial representation

Vyvyan Evans

Overview

The human experience of space includes knowledge relating to the size, shape, location and distribution of entities in a stable three-dimensional environment. In this introductory chapter I address the perceptual systems and processes that facilitate this: the sense-perceptory and brain mechanisms that process perceptual information giving rise to spatial experience. I also examine the processes whereby perceptual experience is redescribed into rudimentary representations of space. That is, I examine primitive concepts which form the bedrock of our ability to think, reason and talk about space and, indeed, more abstract realms. Thus, this chapter is concerned primarily with i) the perception of space, and the way in which spatial experience is 'constructed' by virtue of our sense-perceptory systems and brain mechanisms, and ii) how spatial experience is 'redescribed', giving rise to foundational spatial concepts prior to the emergence of language from around one year onwards.

The chapter begins by examining the distinction between spatial representations, exploring the difference between percepts and concepts. I then examine the human perceptual systems which facilitate the detection of sensory stimuli from the external environment. I then look at perceptual theories which attempt to explain how the brain constructs spatial experience from this sensory input. I then turn to the human mapping ability: an innate mechanism that allows us to construct spatial or cognitive 'maps' based on locational information. This ability is essential for wayfaring, which is to say navigating in space. I then examine how percepts are redescribed as the basic spatial primitives, known as image schemas.

1 Introduction: perception vs conception

My main concern in this chapter is to review the way in which space is experienced and constructed by the human *sensory* (or *sense-perceptory*) *systems*, and the brain**.** I also review the way in which these objects of spatial perception known as *percepts* give rise to rudimentary spatial representations (or *concepts*) known as *image schemas*. Accordingly, at this point I briefly review the distinction between perception (and percepts), and conception (and concepts).

Perception consists of three stages: i) *sensation* ii) *perceptual organisation* and iii) *identification and recognition*. Sensation concerns the way in which external energy, such as light, heat, or (sound) vibrations are converted into the *neural codes* which the brain recognises. Perceptual organisation concerns the way in which this sensory information is organised and formed into a perceptual object, a percept. Identification and recognition relates to the stage in the process whereby past experiences and conceptual knowledge is brought to bear in order to interpret the percept. For instance, a spherical object might be identified and recognised as a football or a coin, or a wheel, or some other object. That is, this stage involves meaning, which is to say understanding the nature, function and significance of the percept. As such, a previously-formed concept is employed in order to identify and categorise the percept.

Table 1: Three stages in perception

Sensation	external energy stimuli are detected and converted into neural codes
perceptual organisation	integration of neural codes by the brain to form a percept
identification and recognition	the percept is categorised, which involves matching with stored experiences

The distinction between percepts and concepts relates to distinctions in *representational formats*: how experience is presented at the cognitive level and how it is stored. Percepts constitute coherent representations which derive from sensory experience, and arise from multiple *modalities*. That is, they derive from information which is integrated from a number of different sensory systems, discussed in more detail in the next section. Percepts are typically available to conscious experience. That is, they are the product of *on-line processing*, resulting from a stimulus array perceived in the 'here-and-now'. A consequence of this is that they consist of specific information relating to the specific stimulus array that they are derived from. Thus, they are *episodic* in nature.

Concepts, on the other hand, represent *schematisations*, formed by abstracting away points of differences in order to produce representations which generalise over points of similarity. Thus, the concept CAR, for instance, is a schematisation derived by generalising across many different sorts of specific (episodic) experiences relating to automobiles in order to form a single representation. Of course, this greatly simplifies things, and I emphasise that concepts, while stable schematisations are not static and unchanging. Indeed, they continue to be updated and thus evolve as the human perceiver continues to be exposed to new experiences. A consequence of the schematic nature of concepts is that, unlike percepts, concepts are representations in the sense of re-presentations. That is, they are stored in memory and can be activated during *off-line processing*. That is, they can be recalled in the absence of the percept(s) which may have given rise to them.

A further important point is that while percepts relate primarily to the sensory details of a given entity, concepts include a much greater range of information types, including the nature and function of the entity which is being represented, as well as how it relates to other concepts. Thus, concepts are related to one another in a systematic way, and form a structured knowledge 'inventory, what I will refer to as the human *conceptual system*. Thus, concepts constitute 'theories' concerning a particular entity, and as such bring meaning to bear with respect to any given percept (for discussion see Mandler 2004).

This said, how do percepts and concepts arise? Percepts arise from a process termed *scene analysis* (e.g., Bregman 1990). Scene analysis is the process whereby the perceptual stimulus array is segregated into coherent percepts. This is achieved by both *bottom-up processing* and *top-down processing***.**

Bottom-up processing relates to the processing and integration of perceptual 'details' that make up, for instance, object percepts, such as a vase or a ball. I will consider two sorts of perceptual details later in the chapter which are termed *textons* and *geons*. Top-down processing relates to the integration of perceptual information which is guided by global principles. Such principles have been proposed, for instance by Gestalt psychology, an important and influential movement that I will consider in detail below.

Bottom-up and top-down processing cross-cut another important distinction which relates to *primitive segregation* versus *schema-based segregation*. That is, scene analysis proceeds by making use of both innate and learned constraints. Primitive segregation is segregation of the stimulus array based on innate, which is to say, pre-given, primitives. Such primitives, which include, for instance *figure-ground segregation*, discussed below, derive from invariants in the stimulus array which have, through evolutionary processes come to be 'hard-wired' in the human brain. In contrast, schema-based segregation involves scene analysis which employs learned constraints.

Before concluding this section, it is necessary to briefly say something about the relationship between spatial concepts and percepts. In fact, this is an issue I address in greater detail when I present the work of developmental psychologist Jean Mandler later in the chapter. However, for now I note that spatial concepts derive from, in the sense of being 'redescribed' from, perceptual experience. This process, which Mandler refers to as *perceptual meaning analysis*, uses spatial percepts as the basis for the formation of rudimentary spatial concepts: image schemas. I will have more to say about these basic spatial concepts later.

2 Sensory systems

In this section I review the mechanisms that facilitate the processing of energy signals from the environment, the *stimulus array*, and how this information is detected by our sensory systems, and processed. I begin by examining the sensory organs and systems which serve as our windows on our spatial environment.

2.1 The visual system

The crucial organ for the visual system is the eye. The brain and the eye work together to produce vision. Light enters the eye and is changed into nerve signals that travel along the optic nerve to the brain. As light enters the eye it is brought into focus on the rear surface of the eyeball. Light enters at the cornea (see Figure 1), which helps to bend light directing it through the pupil: the small dark circle at the centre of your eye. The amount of light that enters the pupil is controlled by the iris – often coloured brown or blue and encircles the pupil – which expands or contracts making the iris larger of smaller. Behind the pupil is a lens, a spherical body, bringing light waves into focus on the retina, the rear of the eyeball. The retina consists of a thin layer of light receptors known as photoreceptors. There are two kinds of photoreceptors: cones and rods. Cones allow us to see in colour and provide our perception in daylight. Rods facilitate vision under dim conditions and allow only black and white perception. That part of the retina which is most sensitive is called the *macula*, and is responsible for detailed central vision. The part of the macula which produces clearest vision is the *fovea*. It is a tiny area densely packed with cone cells. Accordingly, when we look ahead, light reflected from objects in our 'line of sight' is directed onto our fovea, and objects occupying this area of the macula are perceived by virtue of what is termed *foveal vision*. Objects at the edge of the visual field are perceived less clearly. Vision of this kind is known as peripheral vision.

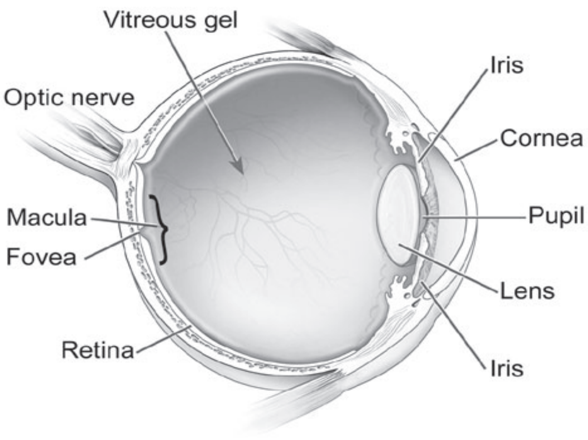

Figure 1. The eye.

'What' and 'where' visual systems

The photoreceptor cells on the retina convert light energy into neural information. However, this information from different parts of the retina is carried along two different pathways or 'streams', connecting different parts of the *visual cortex* – that part of the brain responsible for vision – and providing distinct sorts of information. The visual cortex occupies about a third of the (*cerebral*) *cortex*, the outer layer of the *cerebrum* (consisting of four lobes, see Figure 2).

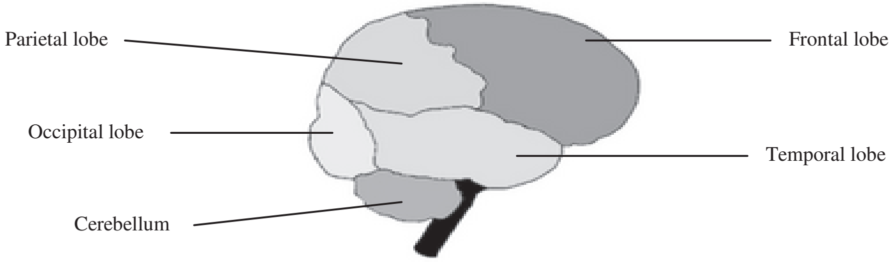

Figure 2. Diagram showing the four lobes of the cerebrum, and the cerebellum.

(The cebral cortex is the outer layer of the cerebrum. Note: The brain is seen from the right side, the front of the brain, above the eyes, is to the right.)

The visual cortex is divided into approximately thirty interconnected visual areas. The first cortical visual area is known as the primary visual cortex or V1. V1 sends information along two separate pathways or 'streams' through different parts of the visual cortex, giving rise to two separate visual systems each giving rise to different kinds of information (Ungerleider and Mishkin 1982). The primary visual system, known as the *focal system* sends information from the macula along the pathway known as the *ventral stream* (ventral means 'lower'). This system, often referred to as the 'what' system, provides information relating to form recognition and object representation. That is, it allows us to identify and recognise objects, including the recognition of attributes such as colour, for instance.

The second system, known as the *ambient system* sends information from both the macula and more peripheral locations on the retina along a pathway known as the *dorsal stream* (dorsal means 'upper'). This system, also known as the 'where' system, provides information relating to where an object is located in body-centred space, rather than with details of the object itself. Thus, light signals in the eye are transformed by the brain providing two distinct sorts of information relating to 'what' and 'where'.

More recently Milner and Goodale (1995) have demonstrated that the distinction between the two 'streams' does not strictly relate to the type of percept ('what' versus 'where') that visual processing provides, in the way conceived by Ungerleider and Mishkin. Rather, while the ventral stream provides information that allows humans to perceive particular objects ('what'), the dorsal stream provides functional information which facilitates readiness for action in order to interact with objects and other entities in the world. In other words, the ventral stream provides information leading to the conscious understanding of objects and other entities in the physical environment, while the dorsal stream serves to facilitate motor programming.

Important evidence for these two distinct visual systems comes from the phenomenon known as *blindsight*. Some blind individuals appear to be able to localise and orient to objects without actually being able to see them. In other words, some blind people appear to be able to locate objects without knowing what the objects are, that is, without being able to identify the object. This suggests that in such cases while the focal system is damaged, the ambient system, mediated by the dorsal stream allows them to make correct orientation judgments and responses, providing compelling evidence for two distinct kinds of visual information.

Recent work on spatial representation in language suggests that the 'what' and 'where' systems may have linguistic reflexes. For instance, Landau and Jackendoff (1993) argue that spatial relations, as encoded by prepositions, and objects as encoded by count nouns roughly approximate the pre-linguistic representations deriving from the 'where' and 'what' systems respectively. Similarly, Hurford (2003) argues that the 'where' and 'what' systems provide neurological antecedents for predicate-argument structure in language.

2.2 The vestibular system

The *vestibular system*, or *orienting sense* is the sensory system that provides information relating to our sense of balance, and is the dominant system with respect to sensory input about our movement and orientation in space. Together with the cochlea, the auditory organ, discussed below, the vestibular system, is situated in the *vestibulum* in the inner ear (Figure 3).

As our movements in space consist of *rotations* – circular motion, as when we turn around – and *translations* – linear motion, as when we walk along a path (horizontal motion), or climb a ladder (vertical motion or gravity) – the vestibular system comprises two components. The first component consists of semicircular canals which detect rotations. These are interconnected fluid-filled tubes which are located in three planes at right angles to one another. The inner surface of the canals also contain hairs. As the fluid moves in response to rotational movement the hairs detect motion of the fluid and transduce this into neural code. The three distinct canals serve to provide rotational information from three axes.

The second component consists of two fluid-filled sacs, the *utricle* and the *saccule*. These chambers contain *otoliths* – literally 'ear stones' – which are heavier than the fluid in the sacs and respond to linear and vertical motion, including both left-right, forward-back motion and gravity (vertical motion). As before both the utricle and saccule contain hairs which detect movement of the otoliths in response to linear motion. This information is transduced into neural code which is transmitted to the brain for processing.

The vestibular system sends signals primarily to the neural structures that control our eye movements, and to the muscles that keep us upright. One important function of the vestibular system is to coordinate body and head movement with the detection of motion by the visual system. This is referred to as the *vestibuo-couclar reflex (VOR)*, which is necessary for vision to remain clear. This works during head movement by producing an eye movement in the direction opposite to head movement, thus preserving the image on the centre of the visual field. For example, when the head moves to the right, the eyes move to the left, and vice versa. Since slight head movements are present all the time, the VOR is very important for stabilising vision.

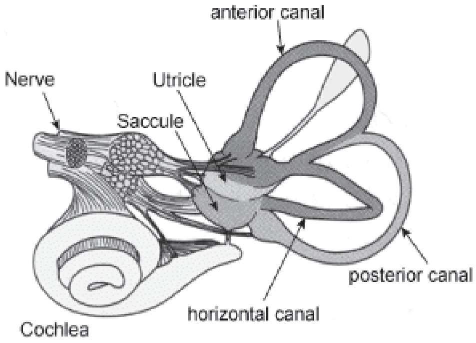

Figure 3. The vestibular system and cochlea.

The vestibular system is, in *phylogenetic* (i.e., evolutionary) terms, one of the first systems to have developed. In *ontogenetic* (i.e., developmental) terms it is the first to fully develop, by six months after conception.

2.3 The auditory system

The vestibular system, and the key auditory organ, the cochlea, are closely linked, both occupying the ear bone. It is widely believed that the cochlea evolved from the phylogenetically earlier sensory structures responsible for detecting bodily orientation.

The auditory system works by transforming sensory information first from air to fluid and then to electrical signals that are relayed to the brain. One important function of the ear is to amplify sound vibrations, in preparation for the transformation from air to fluid. The folds of cartilage that comprise the outer ear on the side of the head are called the *pinna* (see Figure 4). The sound waves enter the ear canal, a simple tube which starts to ampify the sound vibrations. At the far end of the ear canal is the eardrum which marks the beginning of the middle ear.

The middle ear includes the *ossicles* – three very small bones shaped like a hammer, an anvil, and a stirrup. The ossicles further amplify the sounds by converting the lower-pressure eardrum sound vibrations into higher-pressure sound vibrations. Higher pressure is necessary because the inner ear contains fluid rather than air. The signal in the inner ear is then converted to neural code which travels up the auditory nerve.

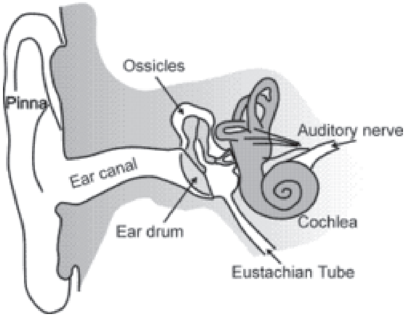

Figure 4. Anatomy of the ear.

The auditory nerve takes the neural code to that part of the brainstem known as the cochlear nucleus. From the cochlear nucleus, auditory information is split into two streams, similar to the way in which the visual signal is split into 'where' and 'what' streams. Auditory nerve fibres going to the ventral cochlear nucleus preserve the timing of the auditory signal in the order of milliseconds. Minute differences in the timing of signals received by both ears allow the brain to determine the direction of the sound.

The second, dorsal, stream analyses the quality of sound. It does this by virtue of detecting differences in frequencies and thus allows differentiation of phonemes, such as the distinction between *set* versus *sat*.

2.4 The haptic system

The haptic system includes the combined sensory input from the receptors for touch in the skin and *proprioception receptor***s** in the body's muscles and joints. Together sense-perception from the haptic system gives rise to perceptual information from a broad range of contact encounters between the body and environment that are sent to, and processed by, a region of the cerebral cortex known as the *somatosensory area*. The haptic system – deriving from *hapsis* which is Greek for 'to grasp' – provides perception of geometric properties including the shape, dimension, and proportions of objects. It also gives rise, through the proprioceptive receptors, to the felt sense of co-ordinated movement, and thus is responsible, in part, for our perception of being distinct from the environment which surrounds us. I review in more detail below the two key components that make up the haptic system, the skin, and proprioception.

The skin

The skin is the largest organ, covering the entire body. It contains specialised nerve endings which can be stimulated in different ways providing different sensations and thus different sorts of sensory information. The sensory effect resulting from stimulation of the skin is known as *cutaneous sensitivity*. There are three main *cutaneous qualities*: pressure (also known as touch), temperature and pain. The somatesensory cortex in the brain represents different skin regions as well as different cutaneous qualities. Thus, the brain is provided with information relating to where on the skin a particular stimulus is being received and what sort of quality is associated with it.

In terms of touch there is an important distinction to be made between *active touch*, and *passive touch*. In active touch, the experiencer actively controls sensory stimulus activation by virtue of picking up an object, for instance. By contrast, passive touch occurs without the reception of the stimulus being controlled by the experiencer, as when an object is placed in contact with the skin. Although the entire surface of the skin responds to touch, the most sensitive receptors are the 'exploratory' parts of the body. These include the fingers and hands, parts of the mouth and the tip of the tongue, as well as the genitalia.

Proprioception

Proprioception – from the Latin *proprius* which means 'one's own' – relates to the sense of body part position and movement. That is, it concerns the posture, location and movement of the arms, legs and other parts of the human skeleton. Another commonly-used term for proprioception is *kinaesthesis* – or kinaesthesia, from the Greek *kineo*, 'to move'. Proprioception is essential for a whole range of coordinated movements. To get a sense of how it functions close your eyes and then touch your nose with a finger tip. Your ability to do this comes from proprioception.

Proprioceptive receptors are known as *mechanoreceptors*. There are two types. The first type provides sensory stimuli for joint information. The second provides information deriving from mechanoreceptors founds in muscles and tendons. The mechanoreceptors for joint information are stimulated by contact between the joint surfaces. This occurs when the angles at which bones are held with respect to one another change, due to movement. The mechanoreceptors in the muscles and tendons respond to changes in the tension of muscle fibres when movement occurs.

3 Spatial perception: how we experience space

In this section I review the perception of objects, form, movement and three-dimensional space. Perhaps unsurprisingly, given the importance of the visual modality for primates in general and humans in particular, much of the work on various aspects of spatial perception has traditionally focused on visual cues. Indeed, visual perception is perhaps the best studied of the sensory systems. Accordingly, in this section I will primarily focus on the role of visual perception in the experience and construction of spatial percepts.

3.1 Texture and object perception

Before objects can be identified visual details must be processed and integrated by the visual system. Variations in visual scenes, in terms of i) light intensity, i.e., adjacent regions of light and dark areas – known as *contrast phenomena* – ii) patterns and iii) colour, form repeated patterns known as *visual texture*. The patterns, for instance, curly versus straight hair, or a tiger's stripes versus a leopard's spots, are often the result of the physical surface properties such as differentially oriented strands, and direction of light and direction of motion.

One important bottom-up theory of visual texture perception is known as *Feature Integration theory*. This theory assumes that there are two major stages involved in the perception of visual texture. The first stage, known as the *preattentive stage*, involves the unconscious processing of visual texture. In a seminal paper, psychologist Bela Julesz (1981) proposed that the preattentive stage serves to process textural primitives, the fundamental components of visual texture. These he labelled *textons*.

Textons are distinct and distinguishable characteristics of any given visual display. For instance, textons include straight lines, line segments, curvature, widths, lengths, intersections of lines, and so on. According to Julesz, the first stage of visual texture perception involves discriminating between the range of textons in a visual display. The second stage in visual texture perception is the *focused attention stage*. This involves conscious processing in order to integrate the textons into complex unitary objects.

Just as textons have been proposed as the primitive elements of visual texture perception, a related bottom-up theory has been proposed to account for object identification. This theory, associated with the work of Biederman (1987) is called *recognition by components*. Biederman's essential insight is that the identification of objects involves the combination of a set of primitive three-dimensional geometric components which he labels *geons*, short for 'geometric icons'. Geons are simple volumes such as cubes, spheres, cylinders, and wedges (see Figure 5). Biederman has proposed 36 geons which can be combined in a range of ways giving rise to complex objects. Biederman argues that object perception crucially relies upon recognising the components which make up an object, the geons. Figure 6 illustrates how a perceived object is comprised of a range of constituent geons. The image on the left corresponds to the perceived object (a desk lamp), and the image on the right to the constituent geons.

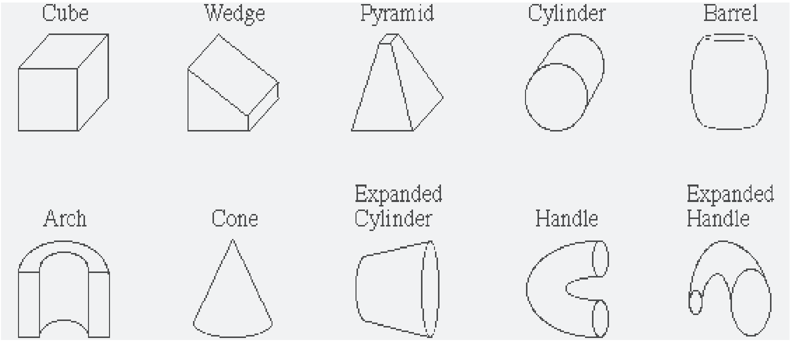

Figure 5. Some examples of geons (After Biederman 1987).

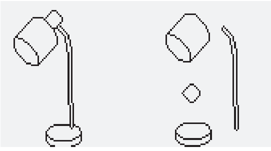

Figure 6. Geons in object perception.

3.2 Form perception

In the previous section I briefly looked at primitive elements that have been proposed for textual perception and the identification of objects. However, in addition to identifiable components of images and objects, there are also higher-level processes involved that are essential for the perception of forms and the grouping of objects. Moreover, these appear to be innate. I discuss two sorts of such organising principles below, figure-ground segregation, and the Gestalt grouping principles.

Figure-ground perception

A fundamental way in which we segregate entities in our environment, thereby perceiving distinct objects and surfaces, comes from the our ability to perceive certain aspects of any given spatial scene as 'standing out' from other parts of the scene. This is known as *figure-ground organisation*.

The phenomenon of figure-ground organisation was pointed out by the Danish psychologist Edgar Rubin in 1915. He observed that in visual perception we see parts of a given spatial scene as being made up of well-defined objects, which 'stand out' from the background. That is, we see objects as three-dimensional entities which stand out from the terrain in which they are located. For instance, in Figure 7, the image of the lighthouse, the figure, stands out from the grey horizontal lines, the ground, as a recognisable and distinct image.

Figure 7. Figure-ground segregation.

Rubin proposed a number of perceptual differences between the figure and ground. These are summarised in table 2.

Table 2: Distinctions between figure and ground.

Figure	**Ground**
Appears to be thing-like	appears to be substance-like
a contour appears at edge of figure's shape	relatively formless

Figure	**Ground**
appears closer to the viewer, and in front of the ground	appears further away and extends behind the figure
Appears more dominant	less dominant
better remembered	less well remembered
more associations with meaningful shapes	suggests fewer associations with meaningful shapes

In addition, figure-ground perception appears to be innate. For instance, photographs which lack depth, being two-dimensional surfaces, are perceived in three-dimensional terms. That is, the figure-ground organisation associated with photographs is an illusion. A particularly well-known illusion made famous by Rubin is the vase-profile illusion (Figure 8).

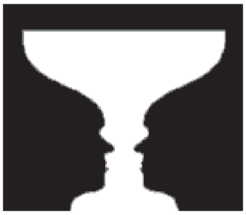

Figure 8. The vase/profile illusion.

The vase/profile illusion is an ambiguous figure-ground illusion. This is because it can be perceived either as two black faces looking at each other, on a white background, or as a white vase on a black background. In other words, it undergoes spontaneous **reversal.** This illusion shows that perception is not solely determined by an image formed on the retina. The spontaneous reversal illustrates the dynamic nature of the perceptual processes. These processes illustrate that how the brain organises its visual environment depends on our innate ability to segregate images on the basis of figure-ground organisation. As this image contains the same percentage of black and white, that part of the image which is assigned the role of figure determines whether a vase or faces are perceived.

Figure-ground organisation appears to be an evolutionary response to our physical environment. Our visual system, for instance, has evolved in order to be able to perceive three-dimensional objects as distinct from the surrounding terrain in which they are embedded. Figure-ground organisation thus constitutes a hard-wired response to this imperative.

Gestalt grouping principles

Gestalt psychology was a movement which emerged in the first decades of the twentieth century. Its primary concern, and those of its three leading proponents, the German psychologists Max Wertheimer, Kurt Koffka and Wolfgang Köhler was to investigate why some elements of the visual field form coherent figures, and others serve as the ground. Gestalt is the German term for 'form', or 'shape' or 'whole configuration'. The Gestalt psychologists proposed a number of innate grouping principles that enable us to perceive forms. Some of these, based on the work of Max Wertheimer (1923) are presented below.

Principle of Proximity (or nearness)

This principle states that the elements in a scene which are closer together will be seen as belonging together in a group. This is illustrated in Figure 9. The consequence of the greater proximity or nearness of the dots on the vertical axis is that we perceive the dots as being organised into columns rather than rows.

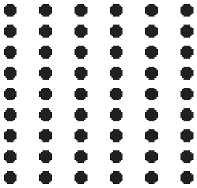

Figure 9. Column of dots.

If the scene is altered so that the dots are closer together on the horizontal axis, then we perceive a series of rows, as illustrated in Figure 10.

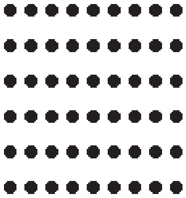

Figure 10. Rows of dots.

Principle of Similarity

This principle states that entities which share visual characteristics such as size, shape or colour will be perceived as belonging together in a group. For example, in Figure 11, we perceive columns of shapes (rather than rows). In fact, the shapes are equidistant on both the horizontal and vertical axes. It is due to our innate predisposition to organise based, here, on similarity that similar shapes (squares or circles) are grouped together and, consequently, are perceived as columns.

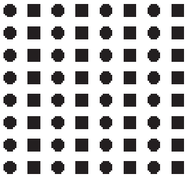

Figure 11. Columns of shapes

Principle of Closure

This principle holds that incomplete figures are often 'completed', even when part of the perceptual information is missing. For instance, in Figure 12 we perceive a circle, even though the 'circle' is incomplete. That is, there is a tendency to close simple figures, by extrapolating from information which is present.

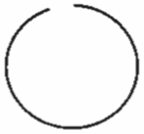

Figure 12. An incomplete figure subject to perceptual closure

A related perceptual process is illustrated by the following. In Figure 13, a white triangle is perceived as being overlaid on three black circles, even though the image could simply represent three incomplete circles. This phenomenon is known as the perception of *subjective* or *apparent contours*. It resembles closure, in so far as there is the appearance of edges across a blank area of the visual field.

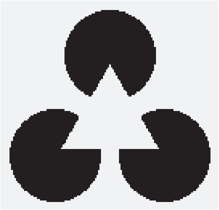

Figure 13. Subjective contour: A white triangle

Principle of Good Continuation

This principle states that human perception has a preference for continuous figures. This is illustrated in Figure 14. Here, we perceive two unbroken rectangles, one passing behind another, even though this is not what we actually see. In fact, the shaded rectangle is obscured by the first, so we have no direct evidence that the shaded area represents one continuous rectangle rather than two separate ones.

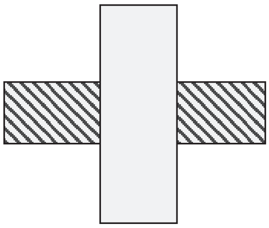

Figure 14. Two rectangles

Principle of Smallness

The Principle of Smallness states that smaller entities tend to be more readily perceived as figures than larger entities. This is illustrated in Figure 15. We are more likely to perceive a black cross than a white cross, because the black shading occupies a smaller proportion of the image.

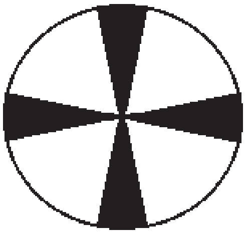

Figure 15. A black cross

Principle of common fate

The final principle I consider here is the Principle of Common Fate. This states that elements that move in the same direction are perceived as being related to one another. For instance, assume that we have two rows of 4 small squares. If the middle two squares from the bottom row begin to move down the page, as depicted by the arrows in Figure 16, they are perceived as belonging together and thus form a separate group from those that remain stationary.

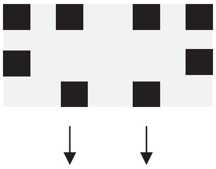

Figure 16. Motion in the same direction

The Gestalt grouping principles I have surveyed conform to the general Gestalt Principle of Good figure, also known as the *Law of Prägnanz*. This states that we tend to perceive the simplest and most stable of the various perceptual possibilities.

While I have primarily focused in this section on visual perception, it is important to emphasise that the principles I have discussed, both figure-ground and grouping principles manifest themselves in other modalities. For instance, Kennedy (1983; Kennedy and Domander 1984), present evidence that figure-ground perception, including the analogues to the ambiguous profile/vase illusion occur in the tactile (touch) modality, based on experiments involving raised-line drawings of reversible figures. Similarly, Bregman (1990) has argued that the Gestalt principles apply equally to auditory scene analysis. He makes the point, for instance, that the ability to perceive a series of musical notes as forming a tune is an instance of a gestalt par excellence.

3.3 The perception of movement

Our ability to detect movement is essential for the survival of the species. Below I discuss a number of different systems for the detection of motion, and different kinds of motion. I begin with the visual detection of motion.

Two visual systems

Motion detection appears to have evolutionary priority over shape detection (Gregory 1998). Indeed, as observed by Gregory, the evolution of the eye emerged in the first place in order to detect motion. Indeed, only eyes relatively high up the evolutionary scale produce stimulus in the absence of motion. The evolutionary development of vision and the detection of motion are represented in the human eye:

> The edge of our retinas are sensitive only to movement. You can see this by getting someone to wave an object around at the side of your visual field where only the edge of the retina is stimulated. Movement is seen, but it is impossible to identify the object, and there is no colour. When movement stops the object becomes invisible. This is as close as we can come to experiencing primitive vision. The extreme edge of the retina is even more primitive: when it is stimulated by movement we experience nothing; but a reflex is initiated, rotating the eye to bring the moving object into central vision...'

> (Gregory, 1998: 98).

The human visual system involves eyes which can move in the head, as when we keep our heads stationary and move our eyes from side to side or up and down. Consequently, our visual system has two distinct ways of detecting motion.

The first involves *image-retina movement*. This involves the eye ball remaining stationary. In this situation the image of moving objects run sequentially across adjacent photoreceptors on the retina. That is, the detection of movement occurs as different photoreceptors are understood by the brain as relating to different locations in space. The second method involves *eye-head movement*. This relates to movement of the eyes in the eye-ball socket when we follow an object in motion. In this situation, an object is not running across different photoreceptors as the eye moves in order to track the object. Rather, information from the eye muscles, which stretch in response to the movement of the eye is understood by the brain as relating to motion of the tracked object.

Optic flow

In normal vision when we move our eyes, the world remains stable. That is the visual world doesn't spin around. This follows as during normal eye movements, signals from the image-retina and eye-head systems cancel each other out, such that the world is perceived as stable. While the two visual systems just described relate to the detection of the movement of objects, another source of movement detection comes from the way in which the human experiencer moves about during the world. As we move around the location from which we view our environment changes. The consequence or this is that there is a continuous change in the light stimulus which is projected on the retina. Following the pioneering work of psychologist James Gibson (e.g., 1986), this changing stimulus is known as *optic flow*.

Optic flow relates to a radial pattern which specifies the observer's direction of self-motion and is essential for successful navigation through the environment. As we travel through the world, and as we approach objects, they appear to move towards us, flowing past behind us as we move beyond them. Moreover, different objects at different points in the visual field appear to move towards and past us at different rates. For instance, imagine sitting on a train and travelling through the countryside. Distant objects such as clouds or mountains appear to move so slowly that they are stationary. Closer objects such as trees appear to move more quickly while very close objects appear to whiz by in a blur. This motion, the optic flow pattern, provides important cues as to distance. Moreover, the optic flow varies depending on the relationship between viewing angle and direction of travel. For instance, objects which are dead-ahead and thus centred in the visual field will appear to remain stationary, while objects which are more peripheral in the visual field will appear to move more rapidly. However, because the edges of centred objects will not be in foveal vision, the edges will have optic flow associated with them. Thus, optic flow patterns provide important information about both distance and direction of travel.

Biological motion

The requirement of being able to rapidly detect the motor activities of humans and other organisms is essential for survival. Indeed, under certain lighting conditions, such as at dusk, details relating to the precise nature of the animal in question may not be readily discernable, especially if the animal is distant. Accordingly, humans have evolved an ability to detect what Johansson (1973) terms *biological motion*. Based purely on movement cues, we can quickly distinguish biological from non-biological motion. Moreover, humans can readily distinguish between different types of biological motion based solely on movement cues, for example, running versus jogging versus walking versus jumping, and so on. Each gait represents a gestalt constructed from a sequence of pendulum-like motions, specific to each activity type.

Evidence for this ability comes from the work of the visual psychologist Gunnar Johansson. He videotaped actors in complete darkness. The actors had *point-light displays* (points of light) fixed at ten main body joints which served as the only illumination. This eliminated all non-movement cues, such as the body contours of the actors. Subjects were then asked to identify biological motion and the motor activities engaged in by the actors. Johansson found that in the absence of motion subjects failed to recognise the point light displays as representing a human form. However, with movement the subjects vividly perceived human motion. In other words, subjects related moving lights in order to perceive human movement, and moreover, were able to identify the *pattern of movement*, that is, the kind of movement being engaged in.

3.4 The perception of three-dimensional space

In this section I briefly review how the brain constructs (three dimensional) space, that is, depth, when the retina is a two-dimensional surface. In other words, where does the third dimension come from? I consider below a number of cues that the brain extracts from the visual stimuli in order to construct our experience of (three-dimensional) space.

While depth and distance can be constructed on the basis of a range of visual (and other) stimuli, including auditory cues, and the optic flow patterns described above, an important means of obtaining depth information comes from *binocular cues*. This relates to the spatial stimuli provided by virtue of having two eyes.

The eyes are separated by about 6.5 cm (Gregory 1998). The consequence of this is that each eye sees a different view. As Gregory observes, '[t]his can be seen clearly if each eye is closed alternately. Any near object will appear to shift sideways in relation to more distant objects and to rotate slightly when each eye receives its view.' (Ibid.: 60). The difference between the two retinal images is known as *binocular disparity*, and gives rise to the perception of depth or *stereoscopic vision*. However, stereoscopic vision only applies to objects which are quite near. This follows as binocular disparity reduces the further away an object is. As Gregory notes, '[w]e are effectively one-eyed for objects further than about 100 metres.' (Ibid.: 60). In other words, depth is a consequence of binocular rather than *monocular* (one-eyed) vision.

4 Cognitive maps

In this section I review in more detail the sorts of spatial representations that the brain constructs from the sensory systems and perceptual stimuli described in previous sections. While I have examined distinct sensory systems, in practice perceptual information from a range of modalities is integrated in order to form spatial or cognitive maps. These are complex mental representations which facilitate navigation and moreover, are necessary for the emergence of the concepts of place and location. The concepts of place and location are independent of the entities and objects which occupy specific places or locations. That is, without a cognitive mapping ability which allows us to perceive places and locations independent of the objects which occupy them we would have no means of understanding these concepts. Accordingly, the concepts PLACE and LOCATION are a consequence not of such notions being an inherent aspect of an objective reality, but rather derive from innate cognitive mapping abilities, and particularly our ability to construct spatial maps independently of our egocentric spatial location, as discussed below.

4.1 Egocentric versus allocentric representations

There are two main sorts of spatial cognitive reference frames manifested by humans and many other species. These are *egocentric* representations and *allocentric representations*. In this section I briefly introduce cognitive reference frames of both these sorts.

There is good neurobiological evidence that humans, along with other mammals, maintain multimodal cognitive spatial 'maps' in the parietal cortex (recall Figure 2). The distinguishing feature of egocentric 'maps' is that they represent objects in space with respect to the organism, or part of the organism, such as the organism's hand, body or head. This follows as cognitive 'maps' of this kind represent space in *topographic fashion*. That is, neighbouring areas of neural space represent neighbouring regions of space in the world of the perceiving organism, with respect to the organism which serves as reference point or *deictic centre* for organising the location of the represented objects and regions of space. As regions of space are organised with respect to the organism, spatial maps of this kind are termed *egocentric* representations.

In addition, there is a second kind of spatial representation which is *allocentric* (or other-focused) in nature. These representations, which are more appropriately thought of in terms of maps (for reasons I shall discuss below), integrate information derived from the egocentric spatial representations. Crucially, however, the allocentric mapping ability represents space, and spatial regions independently of the momentary location of the organism. That is, entities and objects, and the locations of objects are related to one another independently of the ego. This system, which is located in the hippocampal region of the brain (O'Keefe and Nadel 1978) represents place, direction and distance information, rather than object details.

4.2 The hippocampus and the human cognitive mapping ability

In now classic work, neurobiologists John O'Keefe and Lynn Nadel (1978) show not only that i) humans have an objective or absolutive spatial framework in which the entities of our experience are located, but also that, ii) this ability is innate, and along with other mammals is associated with the brain region often implicated in motor function: the hippocampus. According to O'Keefe and Nadel, this allocentric mapping system provides 'the basis for an integrated model of the environment. This system underlies the notion of absolute, unitary space, which is a non-centred stationary framework through which the organism and its egocentric spaces move.' (Ibid.: 2). This hippocampal mapping system consists of two major subsystems, a *place system* and a *misplace system*.

The place subsystem is a memory system that allows the organism to represent places in its environment and crucially to relate different locations with respect to each other. That is, the place system allows the organism to represent relationships between different locations without having to physically experience the spatial relations holding between distinct places. In other words, humans, like many other organisms, can compute distances, and other spatial relations between distinct places such as directions, without having to physically experience the spatial relationships in question. Such a cognitive mapping ability is a consequence of the allocentric place subsystem.

The second subsystem to make up the allocentric cognitive mapping ability, the misplace system, facilitates and responds to exploration. That is, it allows new information experienced as a consequence of exploration to be incorporated into the allocentric map of the organism's environment. It thereby allows the organism to relate specific objects and entities to specific locations, and to update the cognitive map held in the place system based on particular inputs (cues) and outputs (responses). Thus, O'Keefe and Nadel demonstrate two things. Firstly, three-dimensional Euclidean space is, in a non-trivial sense, imposed on perceptual experience by the human mind. Secondly, the notion of all-embracing continuous space, 'out there', which 'contains' objects and other entities, as maintained by the misplace system, is in fact a consequence of first being able to represent locations in an allocentric (i.e., a non-egocentric) fashion, as captured by the place subsystem. In other words, our innate ability to form absolutive cognitive maps of our spatial environment is a prerequisite to experiencing objects and the motions they undergo.

4.3 Maps versus routes

In order to illustrate the distinction between egocentric and allocentric spatial mapping abilities, O'Keefe and Nadel provide an analogy which I briefly discuss here. The analogy relates to the geographic distinction between routes versus maps. In geographic terms, a route constitutes a set of instructions which directs attention to particular objects in egocentric space. That is, routes are inflexible, identifying landmarks in order to guide the traveller, and thus do not allow the traveller freedom of choice. Put another way, routes are *guide-post based*. Moreover, routes are goal-oriented, focused on facilitating travel from a specific, pre-specified location to another. In this, routes correspond to egocentric cognitive representations.

In contrast, maps are, in geographic terms, representations of part of space. A map is constituted of places, and the places which the map represents are systematically connected and thus related to each other. Moreover, and crucially, the places captured by the map are not defined in terms of the objects which may occupy a particular location. That is, and unlike routes, maps are not guide-post based. Thus, maps capture space that is held to exist independently of the objects which may be located at particular points in space. Crucially, a map is a flexible representation, which can be used for a range of purposes. In related fashion, this notion of a map is presented as an analogy of the allocentric cognitive mapping ability that many organisms, including humans, possess.

While then map-like representations of the environment are constructed by humans, as well as by other species, it is far from clear what the nature of these representations are. Nevertheless, it is by now well established that humans do possess complex information structures which can be used to generate highly-detailed map-like representations, which can be used for a range of behaviours. Indeed, an important finding to have emerged is that place memory has a high information capacity, and can be permanently modified by a single experience. Moreover, experiments reported on by O'Keefe and Nadel reveal that this mapping ability can be used to construct maps in a highly flexible and efficient manner.

Finally, I reiterate that the ability to represent space in an allocentric fashion, i.e., map-like representations, is a trait common to a wide variety of organisms. As O'Keefe and Nadel observe, 'The ability of many animals to find their way back to their nests over large distances would appear to be based on some type of mapping system' (Ibid.: 63). Obvious examples include the migratory and homing behaviour exhibited by many kinds of birds. Indeed, a robust finding from studies on homing pigeons is that

they are able to find their way 'home' using novel routes from new release sites. Such abilities would appear to require a cognitive mapping ability.

5 Primitive spatial concepts

In this section I turn to an examination of spatial concepts and the way in which spatial concepts are derived (or redescribed) from spatial experience. I focus here on the notion of the image schema. Image schemas were first proposed by cognitive linguists (e.g., Johnson 1987, 2007; Lakoff 1987; see Evans and Green 2006 for a review), and represent a rudimentary conceptual building block derived from *embodied experience* (discussed further below). This notion has been subsequently adopted by a range of other cognitive scientists in their work (see papers and references in Hampe 2005). In particular, the notion of the image schema has been developed in the influential work of developmental psychologist Jean Mandler (e.g., 2004) in her work on how conceptual development takes place.

5.1 Embodiment and experience

I begin this brief overview of the image schema by first introducing the role of *embodiment* in the formation of concepts. Due to the nature of our bodies, including our neuro-anatomical architecture, we have a species-specific view of the world. In other words, our construal of 'reality' is mediated, in large measure, by the nature of our embodiment. One obvious example of the way in which embodiment affects the nature of experience relates to biological morphology (i.e., body parts). This, together with the nature of the physical environment with which we interact, determines other aspects of our experience. For instance, while gravity is an objective feature of the world, our experience of gravity is determined by our bodies and by the ecological niche we have adapted to. For instance, hummingbirds – which can flap their wings up to fifty times per second – respond to gravity in a very different way from humans. They are able to rise directly into the air without pushing off from the ground, due to the rapid movement of their wings.

The fact that our experience is embodied – that is, structured in part by the nature of the bodies we have and by our neurological organisation – has consequences for cognition. In other words, the concepts we have access to and the nature of the 'reality' we think and talk about are a function of our embodiment – the phenomenon of *variable embodiment*. That is, we can only talk about what we can perceive and conceive, and the things that we can perceive and conceive derive from embodied experience. From this point of view, the human mind must bear the imprint of embodied experience. This thesis is known as the *thesis of embodied cognition*. This position holds that conceptual structure – the nature of human concepts – is a consequence of the nature of our embodiment and thus is embodied.

5.2 Image schemas

The theoretical construct of the image schema was developed by Mark Johnson in his now classic 1987 book, *The Body in the Mind*. Johnson proposed that one way in which embodied experience manifests itself at the cognitive level is in terms of image schemas. These are rudimentary concepts like CONTACT, CONTAINER and BALANCE, which are meaningful because they derive from and are linked to human *pre-conceptual experience*. This is experience of the world directly mediated and structured by the human body.

The term 'image' in 'image schema' is equivalent to the use of this term in psychology, where *imagistic* experience relates to and derives from our experience of the external world. Another term for this type of experience is sensory experience, because it comes from sensory-perceptual mechanisms that include, but are not restricted to, the visual system.

According to Johnson (1987) there are a number of properties associated with image schemas which I briefly review below.

Image schemas are pre-conceptual in origin

Image schemas such as the CONTAINER schema are directly grounded in embodied experience. This means that they are pre-conceptual in origin. Mandler (2004) argues, discussed further in the next section, that they arise from sensory experiences in the early stages of human development that precede the formation of concepts. However, once the recurrent patterns of sensory information have been extracted and stored as an image schema, sensory experience gives rise to a conceptual representation. This means that image schemas are concepts, but of a special kind: they are the foundations of the conceptual system, because they are the first concepts to emerge in the human mind, and precisely because they relate to sensory-perceptual experience, they are particularly schematic. Johnson argues that image schemas are so fundamental to our way of thinking that we are not consciously aware of them: we take our awareness of what it means to be a physical being in a physical world very much for granted because we acquire this knowledge so early in life, certainly before the emergence of language.

Image schemas form the basis of word senses

Concepts lexicalised by words such as prepositions, for instance, *in*, *into, out, out of* and *out from* are all thought to relate to the CONTAINER schema: an abstract image schematic concept that underlies all these much more specific *senses* – the semantic pole associated with lexical forms (see Tyler and Evans 2003).

The CONTAINER image schema is diagrammed in Figure 17. This image schema consists of the structural elements interior, boundary and exterior: these are the minimum requirements for a CONTAINER (Lakoff 1987). The *landmark* (LM), represented by the circle, consists of two structural elements, the interior – the area within the boundary – and the boundary itself. The exterior is the area outside the landmark, contained within the square. The container is represented as the landmark because the boundary and the exterior together possess sufficient Gestalt properties (e.g., closure and continuity) to make it the figure, while the exterior is the ground (recall my discussion of Gestalt principles above).

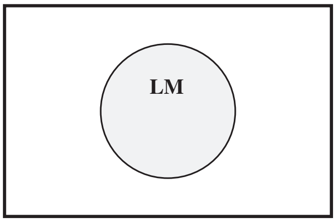

Figure 17. CONTAINER image schema

Although Figure 17 represents the basic CONTAINER schema, there are a number of other image schemas that are related to this schema, which give rise to distinct concepts related to containment. For instance, let's consider one variant of the CONTAINER schema lexicalised by *out*. This image schema is diagrammed in Figure 18 and is illustrated with a linguistic example. The diagram in Figure 18 corresponds to example (1). The *trajector* (TR) *Fred*, which is the entity that undergoes motion, moves from a position inside the LM to occupy a location outside the LM. The terms 'TR' and 'LM' derive from the work of Langacker (e.g., 1987), and relate to the Gestalt notions of figure and ground respectively.

(1) Fred went out of the room

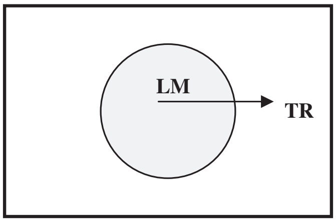

Figure 18: Image schema for OUT

The image schema shown in Figure18 represents a concept that is more specific and detailed than the image schema diagrammed in Figure 17, because it involves motion as well as containment. This shows that image schemas can possess varying degrees of schematicity, where more specific image schemas arise from more fundamental or schematic ones.

Image schemas derive from interaction

As image schemas derive from embodied experience, they derive from the way in which we interact with the world. To illustrate this idea, consider the image schema for FORCE. This image schema arises from our experience of acting upon other entities, or being acted upon by other entities, resulting in the transfer of motion energy. Johnson illustrates the *interactional derivation* of this image schema – how it arises from experience – as follows:

> [F]orce is always experienced through interaction. We become aware of force as it affects us or some object in our perceptual field. When you enter an unfamiliar dark room and bump into the edge of the table, you are experiencing the interactional character of force. When you eat too much the ingested food presses outwards on your taughtly stretched stomach. There is no schema for force that does not involve interaction or potential interaction. (Johnson 1987: 43).

Image schemas are inherently meaningful

As image schemas derive from interaction with the world, they are inherently meaningful. Embodied experience is inherently meaningful in the sense that embodied experiences have predictable consequences. To illustrate, imagine a cup of coffee in your hand. If you move the cup slowly up and down, or from side to side, you expect the coffee to move with it. This is because a consequence of containment, given that it is defined by boundaries, is that it constrains the location of any entity within these boundaries. In other words, the cup exerts force-dynamic control over the coffee. This kind of knowledge, which we take for granted, is acquired as a consequence of our interaction with our physical environment. For example, walking across a room holding a cup of coffee without spilling it actually involves highly sophisticated motor control that we also acquire from experience. This experience gives rise to knowledge structures that enable us to make predictions: if we tip the coffee cup upside-down, the coffee will pour out.

Image schemas are analogue representations

Image schemas are *analogue* representations deriving from experience. The term 'analogue' means image schemas take a form in the conceptual system that mirrors the sensory experience being represented. Because image schemas derive from sensory experience, they are represented as summaries of *perceptual states*, which are recorded in memory. However, what makes them conceptual rather than purely perceptual in nature is that they give rise to concepts that are consciously accessible (Mandler 2004). In other words, image schemas structure (more complex) lexical concepts.

Image schemas can be internally complex

Image schemas are often, perhaps typically, comprised of more complex aspects that can be analysed separately. For example, the CONTAINER schema is a concept that consists of interior, boundary and exterior elements. Another example of a complex image schema is the SOURCE-PATH-GOAL or simply PATH schema. Because a path is a means of moving from one location to another, it consists of a starting point or SOURCE, a destination or GOAL and a series of contiguous locations in between, which relate the source and goal. Like all complex image schemas, the PATH schema constitutes an *experiential gestalt*: it has internal structure, but emerges as a coherent whole.

One consequence of internal complexity is that different components of the PATH schema can be referred to. This is illustrated in example (2), where the relevant linguistic units are bracketed. In each of these examples, different components of the path are profiled by the use of different lexical items.

(2) a. SOURCE
John left [England]

b. GOAL
John travelled [to France]

c. SOURCE-GOAL
John travelled [from England] [to France]

d. PATH-GOAL
John travelled [through the Chunnel] [to France]

e. SOURCE-PATH-GOAL
John travelled [from England] [through the Chunnel] [to France]

Image schemas are not mental images

If you close your eyes and imagine the face of your mother or father, partner or lover, what results is a mental image. Image schemas are not the same as mental images. Mental images are detailed, and result from an effortful and partly conscious cognitive process that involves recalling visual memory. Image schemas are schematic, and therefore more abstract in nature, emerging from ongoing embodied experience. This means that you can't close your eyes and 'think up' an image schema in the same way that you can 'think up' the sight of someone's face or the feeling of a particular object in your hand.

Image schemas are multi-modal

Image schemas derive from experiences across different modalities (different types of sensory experience), and hence are not specific to a particular sense. In other words, image schemas are abstract patterns arising from a range of perceptual experiences, and as such are not available to conscious introspection. For instance, blind people have access to image schemas for CONTAINERS, PATHS, and so on, precisely because the kinds of experiences that give rise to these image schemas rely on a range of sensory-perceptual experiences in addition to vision, including hearing, touch, and our experience of movement and balance.

Image schemas form the basis for abstract thought

Lakoff (1987, 1990, 1993) and Johnson (1987) have argued that rudimentary embodied concepts of this kind provide the conceptual building blocks for more complex concepts, and can be systematically extended to provide more abstract concepts and conceptual domains with structure. According to this view, the reason we can talk about being *in* states like love or trouble (3) is because abstract concepts like LOVE are structured and therefore understood by virtue of the fundamental concept CONTAINER. In this way, image schematic concepts serve to structure more complex concepts and ideas.

(3) a. John is in love.
b. Jane is in trouble.
c. The government is in a deep crisis.

According to Johnson, it is precisely because containers constrain activity that it makes sense to conceptualise POWER and all-encompassing states like LOVE or CRISIS in terms of CONTAINMENT.

5.3 Perceptual meaning analysis

The developmental psychologist Jean Mandler (e.g. 1992, 1996, 2004) has made a number of proposals concerning how image schemas might arise from embodied experience. Starting at an early age infants attend to objects and spatial displays in their environment. Mandler suggests that by attending closely to such spatial experiences, children are able to abstract across similar kinds of experiences, finding meaningful patterns in the process. For instance, the CONTAINER image schema is more than simply a spatio-geometric representation. It is a 'theory' about a particular kind of configuration in which one entity is supported by another entity that contains it. In other words, the CONTAINER schema is meaningful because containers are meaningful in our everyday experience.

Mandler (2004) describes the process of forming image schemas in terms of a redescription of spatial experience via a process she labels *perceptual meaning analysis* (Mandler 2004). This process results from children associating functional consequences with spatial displays. That is, image schemas emerge by virtue of analysing spatial displays of various sorts as relating to the functional consequences with which they are correlated. For example, we saw above that a consequence of coffee being located in a coffee cup is that the coffee moves with the cup. That is, containment has functional consequences in terms of containing, supporting and constraining the location of the entity contained. Thus, the distinction between percepts and concepts such as image schemas is that image schemas encode functional information, that is meaning. As Mandler observes, '[O]ne of the foundations of the conceptualizing capacity is the image schema, in which spatial structure is mapped into conceptual structure' (Mandler 1992: 591). She further suggests that 'Basic, recurrent experiences with the world form the bedrock of the child's semantic architecture, which is already established well before the child begins producing language' (Mandler 1992: 597). In other words, it is experience, meaningful to us by virtue of our embodiment, that forms the basis of many of our most fundamental concepts.

References

Biederman, Irving. (1987) Recognition-by-components: A theory of human image understanding. *Psychological Review* 94: 115–147.

Bregman, Albert. (1990) *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Evans, Vyvyan and Melanie Green (2006) *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Gibson, James. (1986) *The Ecological Approach to Visual Perception*. Mahwah, NJ: Lawrence Erlbaum.

Milner, David and Melvyn Goodale. (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

Gregory, Richard. (1998) *Eye and Brain* (5th edition) Oxford University Press.

Hampe, Beate. (ed.) (2005) *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlin: Mouton de Gruyter.

Hurford, James. (2003) The Neural Basis of Predicate-Argument Structure. *Behavioral and Brain Sciences* 26: 261–283.

Johansson, Gunnar. (1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* 14: 201–211.

Johnson, Mark. (1987) *The Body in the Mind*. Chicago: University of Chicago Press.

Johnson, Mark. (2007) *The Meaning of the Body*. Chicago: University of Chicago Press.

Julesz, Bela. (1981) Textons: The elements of texture perception and their interactions. *Nature* 290: 91–97.

Kennedy, John. (1983) What can we learn about pictures from the blind? *American Scientist* 71: 19–26.

Kennedy, John and Ramona Domander. (1984) Pictorial foreground-background reversal reduces tactual recognition by the blind. *Journal of Visual Impairment and Blindness* 78: 215–216.

Lakoff, George. (1987) *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, George. (1990) The invariance hypothesis: Is abstract reason based on image schemas? *Cognitive Linguistics* 1(1): 39–74.

Lakoff, George. (1993) The contemporary theory of metaphor. In A. Ortony (ed.) *Metaphor and Thought* 202–251 (second edition) Cambridge University Press.

Landau, Barbara and Ray Jackendoff. (1993) 'What' and 'where' in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16: 217–265.

Langacker, Ronald. (1987) *Foundations of Cognitive Grammar.* (Vol. I) Stanford, CA: Stanford University Press.

Mandler, Jean. (1992) How to build a baby II. Conceptual primitives. *Psychological Review* 99: 567–604.

Mandler, Jean. (1996) Preverbal representation and language. In P. Bloom, M. Peterson, L. Nadel and M. Garrett (eds) *Language and Space* 365–384. Cambridge, MA: MIT Press.

Mandler, Jean (2004) *The Foundations of Mind: Origins of Conceptual Thought*. Oxford: OUP.

O'Keefe, John and Lynn Nadel (1978) *The Hippocampus as a Cognitive Map*. Oxford: OUP. [Available on-line at www.cognitivemap.net/] Tyler, Andrea and Vyvyan Evans. (2003) *The Semantics of English Prepositions: Spatial Scenes, Embodied Experience and Cognition*. Cambridge: Cambridge University Press.

Ungerleider, Leslie and Mortimer Mishkin (1982) Two cortical visual systems. In D. Ingle, M. Goodale, and R. Mansfield (eds) *Analysis of Visual Behavior* 549–586. Cambridge, MA: MIT Press.

Wertheimer, Max. (1923 [1938]) Untersuchungen zur Lehre von der Gestalt II, in *Psycologische Forschung*, 4: 301–350. Published as: *Laws of Organization in Perceptual Forms*. In W. Ellis (ed.) (1938) *A source book of Gestalt psychology* 71–88. London: Routledge.