

## INVITED REVIEW

## Consciousness

Adam Zeman

Department of Clinical Neurosciences, Western General  
Hospital, Edinburgh, UK

Correspondence to: A. Zeman, Department of Clinical  
Neurosciences, Western General Hospital, Edinburgh  
EH4 2XU, UK  
E-mail: az@skull.dcn.ed.ac.uk

## Contents

**Summary**

I	Introduction	1264
II	Concepts of consciousness	1265
(a)	<i>The etymology of 'consciousness' and 'conscience'</i>	1265
(b)	<i>The meanings of 'consciousness'</i>	1265
(i)	Consciousness as the waking state	1265
(ii)	Consciousness as experience	1266
(iii)	Consciousness as mind	1266
(c)	<i>The meanings of 'self-consciousness'</i>	1266
(i)	Self-consciousness as proneness to embarrassment	1266
(ii)	Self-consciousness as self-detection	1266
(iii)	Self-consciousness as self-recognition	1266
(iv)	Self-consciousness as awareness of awareness	1266
(v)	Self-consciousness as self-knowledge	1266
III	The science of wakefulness	1267
(a)	<i>The electricity of the brain</i>	1267
(b)	<i>The control of conscious states</i>	1268
(i)	Anatomy	1268
(ii)	Physiology	1269
(iii)	Pharmacology	1270
(c)	<i>Pathologies of wakefulness</i>	1271
IV	The science of awareness	1272
(a)	<i>Exquisite correlations</i>	1273
(b)	<i>Implicit perception</i>	1275
(c)	<i>Insights from the study of memory and action</i>	1277
V	Theories of consciousness	1278
(a)	<i>Neurobiological theories</i>	1278
(b)	<i>Information processing theories</i>	1280
(c)	<i>Social theories</i>	1281
VI	The philosophy of consciousness	1282
(a)	<i>Three intuitions about consciousness</i>	1282
(b)	<i>Identity theory</i>	1282
(c)	<i>Functionalism</i>	1283
(d)	<i>Dualism</i>	1283
VII	Conclusion	1284

## Summary

Consciousness is topical, for reasons including its renewed respectability among psychologists, rapid progress in the neuroscience of perception, memory and action, advances in artificial intelligence and dissatisfaction with the dualistic separation of mind and body. Consciousness is an ambiguous term. It can refer to (i) the waking state; (ii) experience; and (iii) the possession of any mental state. Self-consciousness is equally ambiguous, with senses including (i) proneness to embarrassment in social settings; (ii) the ability to detect our own sensations and recall our recent actions; (iii) self-recognition; (iv) the awareness of awareness; and (v) self-knowledge in the broadest sense. The understanding of states of consciousness has been transformed by the delineation of their electrical correlates, of structures in brainstem and diencephalon which regulate the sleep-wake cycle, and of these structures' cellular physiology and regional pharmacology. Clinical studies have defined pathologies of wakefulness: coma, the persistent vegetative state, the 'locked-in' syndrome, akinetic mutism and brain death. Interest in the neural basis of perceptual awareness has focused on vision. Increasingly detailed neuronal correlates of real and illusory visual experience are being defined. Experiments exploiting circumstances in which

visual experience changes while external stimulation is held constant are tightening the experimental link between consciousness and its neural correlates. Work on unconscious neural processes provides a complementary approach. 'Unperceived' stimuli have detectable effects on neural events and subsequent action in a range of circumstances: blindsight provides the classical example. Other areas of cognitive neuroscience also promise experimental insights into consciousness, in particular the distinctions between implicit and explicit memory and deliberate and automatic action. Overarching scientific theories of consciousness include neurobiological accounts which specify anatomical or physiological mechanisms for awareness, theories focusing on the role played by conscious processes in information processing and theories envisaging the functions of consciousness in a social context. Whether scientific observation and theory will yield a complete account of consciousness remains a live issue. Physicalism, functionalism, property dualism and dual aspect theories attempt to do justice to three central, but controversial, intuitions about experience: that it is a robust phenomenon which calls for explanation, that it is intimately related to the activity of the brain and that it has an important influence on behaviour.

**Keywords:** consciousness; awareness; perception; neuropsychology; neuroscience

**Abbreviations:** ARAS = ascending reticular activating system; NCC = neural correlate of consciousness; REM = rapid eye movement; SWS = slow wave sleep

## I. Introduction

The past decade has seen a rising tide of interest in consciousness, accompanied by a surge of publications, new journals and scientific meetings (Dennett, 1991; McGinn, 1991; Edelman, 1992; Flanagan, 1992; Milner and Rugg, 1992; Searle, 1992; Crick, 1994; Penrose, 1994; Metzinger, 1995; Chalmers, 1996; Velmans 1996; Weiskrantz, 1997; Hurley, 1998; Jasper *et al.*, 1998; Rose, 1998; Velmans, 2000). The 'problem of consciousness' has been identified as an outstanding intellectual challenge across disciplines ranging from basic neuroscience through psychology to philosophy, although opinions vary widely on the chances of achieving a solution. The subject is unusual in drawing together scholars from both sides of the gulfs which separate the sciences and the arts, the study of the body and the study of the mind.

Several factors help to explain the current fascination with consciousness. The techniques of physiological psychology, human neuropsychology and, recently, functional imaging are revealing exquisitely detailed correlations between neural processes and features of conscious experience. The

recognition of unconscious or 'implicit' capacities which can exert an influence on behaviour, such as blindsight, has opened up the way to distinguishing the neural substrates of conscious and unconscious activity in the brain. Advances in computational science and artificial intelligence hold out the prospect of engineering conscious systems. More generally, there is a deep dissatisfaction with the Cartesian separation of body and mind, and a desire to find a place for subjective experience in the scientist's world picture.

This is a timely moment to review progress in the field. As at least part of the problem of consciousness flows from the ambiguities of the term, I shall briefly consider its various senses, and its relationship to 'self-consciousness', in Section II, flagging up some philosophical problems for later discussion. Sections III and IV provide a necessarily selective review of empirical research bearing on mechanisms of arousal, visual awareness, memory and volitional, or conscious, action. Section V examines a number of overarching theories of consciousness. Section VI reflects on the philosophical 'problem of consciousness' in the light of the

empirical advances and recent contributions to the philosophical debate.

## II. Concepts of consciousness

### (a) *The etymology of 'consciousness' and 'conscience'*

The word 'consciousness' has its Latin root in *conscio*, formed by the coalescence of *cum*, meaning 'with', and *scio*, meaning 'know'. In its original Latin sense, to be conscious of something was to share knowledge of it, with someone else, or with oneself. The knowledge in question was often of something secret or shameful, the source of a bad *conscientia*, a bad conscience. A 'weakened' sense of *conscientia* coexisted in Latin with the stronger sense which implies shared knowledge: in this weak sense *conscientia* was, simply, knowledge. All three senses (knowledge shared with another, knowledge shared with oneself and, simply, knowledge) entered the English language with 'conscience', the first equivalent of *conscientia*. The words 'conscious' and 'consciousness' first appear early in the 17th century, rapidly followed by 'self-conscious' and 'self-consciousness' (Lewis, 1960).

### (b) *The meanings of 'consciousness'*

The Oxford English Dictionary distinguishes 12 senses of 'conscious' and eight of 'consciousness'. For our purposes it is helpful to distinguish three principal meanings (Zeman *et al.*, 1997).

#### (i) *Consciousness as the waking state*

In everyday neurological practice consciousness is generally equated with the waking state, and the abilities to perceive, interact and communicate with the environment and with others in the integrated manner which wakefulness normally implies. Consciousness in this sense is a matter of degree: a range of conscious states extends from waking through sleep into coma. These states can be defined objectively, using behavioural criteria like those supplied by the Glasgow Coma Scale (Teasdale and Jennett, 1974). Thus we speak of consciousness dwindling, waning, lapsing and recovering; it may be lost, depressed and regained. To be conscious in this sense is to be awake, aroused, alert or vigilant.

#### (ii) *Consciousness as experience*

Consciousness in its first sense is the behavioural expression of our normal waking state.<sup>1</sup> But when we are conscious in this first sense we are always conscious *of* something. In its second sense consciousness is the content of experience from moment to moment: what it feels like to be a certain person,

now, in a sense in which we suppose there is nothing it feels like to be a stone or lost in dreamless sleep (Nagel, 1979). This second sense of consciousness is more inward than the first. It highlights the qualitative, subjective dimension of experience. Philosophers sometimes use the technical (and controversial) term 'qualia' to refer to the subjective texture of experience which is the essence of this second sense of consciousness (Dennett, 1988; Chalmers, 1996).

Several authors have followed William James in seeking to characterize the general properties of consciousness which call for scientific explanation (James, 1890; Shallice, 1988; Searle, 1992; Crick, 1994; Chalmers, 1996; Greenfield, 1998; Tononi and Edelman, 1998a). There is a broad consensus that, in addition to its qualitative character, the following features are central: consciousness is personal, involving a conscious subject with a necessarily limited point of view; its contents are stable for short periods, lasting from hundreds of milliseconds to a few seconds, but characteristically vary over longer intervals; its contents are unified at any one time; they are continuous over time, in the sense that memory normally allows us to connect consciousness of the present with consciousness of the past; consciousness is selective, with a foreground and background, and a limited capacity at a given moment; over time, however, it ranges over innumerable contents, with potential contributions from each of the senses, and from all the major psychological processes, including thought, emotion, memory, imagination, language and action planning. Most states of consciousness are 'intentional', in the philosophical sense that they are directed at the world, consciousness *of* this or that, and these states, in turn, are 'aspectual': conditioned by the perspective which our conscious viewpoint affords (Searle, 1992). Finally, most commentators emphasize the centrality of consciousness to human values: the prolongation of human life, where one can be certain that consciousness has been lost forever, is generally regarded as a wasted effort.

Although we all tend to consider ourselves expert witnesses on the nature of our experiences, the thought that we may be misled by introspection, and that our experience is not as we usually take it to be, underlies several lines of recent work. For example, research on our sensitivity to change in our visual surroundings suggests that the focus of our visual attention is much narrower than we normally suppose (O'Regan, 2000; O'Regan and Noe, 2001); work requiring subjects to give instantaneous reports of their current experience, at the moment a random buzzer sounds, reveals a surprising preponderance of reports of 'inner thought' (Hurlburt, 2000); approaches inspired both by the phenomenological tradition in continental philosophy<sup>2</sup>, and by the practice of meditation, emphasize the potential value of disciplined observation of awareness in supplying first-person data for the scientific study of consciousness (Varela and Shear, 1999). These lines of research take the qualitative

<sup>1</sup>In this review I shall use 'consciousness' and 'awareness' synonymously: their content is 'experience'.

<sup>2</sup>The cornerstone of this tradition, associated particularly with the work of Husserl, is the painstaking description of 'experienced phenomena in order to exhibit their universal character' (Thines, 1987).

character of consciousness seriously while recognizing that our ordinary assumptions about it may be mistaken. This attentive but critical scrutiny of the ‘view from within’ (Varela and Shear, 1999) is a promising development.

### (iii) *Consciousness as mind*

Echoing the weakened Latin sense of *conscientia*, any mental state with a propositional content can be said to be conscious—anything that we believe, hope, fear, intend, expect, desire, etc. Thus we might accurately say that ‘the prime minister is conscious of the funding crisis in the health service’ at a time when his thoughts are quite otherwise occupied. Most of the recent interest in consciousness has centred on its first and second senses, rather than this third sense in which consciousness is synonymous with mind.

It may be helpful to give one example of the use of ‘conscious’ in each of these three main senses: (i) after a lucid interval, the injured soldier lapsed into unconsciousness; (ii) I became conscious of a feeling of dread, and an overpowering smell of burning rubber; (iii) I am conscious that I may be straining your patience.

## (c) *The meanings of ‘self-consciousness’*

‘Self-consciousness’ is also a multi-faceted concept.

### (i) *Self-consciousness as proneness to embarrassment*

The idiomatic sense of self-consciousness implies awkwardness in the company of others. Interestingly, we are self-conscious in this sense when we are excessively aware of others’ awareness of ourselves. This humdrum usage thus turns out to be rather sophisticated, hinting at a link between consciousness of self and consciousness of others which is a focus of current research in developmental psychology (Baron-Cohen, 1995; Frith and Frith, 1999).

### (ii) *Self-consciousness as self-detection*

We might speak of an organism as self-conscious if it can respond to stimuli which impinge upon it directly, or modify its behaviour in ways which imply an awareness of its own actions. Thus your awareness of an insect walking across your hand involves self-consciousness in this rather minimal sense. Rats, who can be trained to respond to a signal in a way that depends on what they were doing last, may be conscious of their own actions in a similar sense (Beninger *et al.*, 1974). But this variety of self-consciousness amounts to little more than perceptual awareness, directed towards events brought about by, or ones which impinge directly upon, the creature in question.

### (iii) *Self-consciousness as self-recognition*

Chimpanzees and orang-utans, but not monkeys, in common with children over ~18 months of age can recognize themselves in mirrors (Gallop, 1970). This ability implies the possession of a rudimentary concept of self. The flowering of the ‘idea of me’ in the human child over subsequent months is attested by the mastery of the first person pronoun and a growing interest in self-adornment (Parker *et al.*, 1995). But physical and verbal self-recognition falls short of the most distinctively human species of self-consciousness, which allows us to reflect upon the mental lives of others and ourselves.

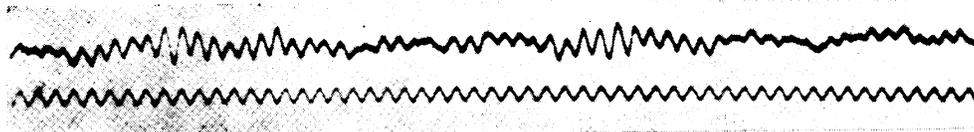
### (iv) *Self-consciousness as awareness of awareness*

We constantly attribute mental states in the everyday explanation and prediction of behaviour: talk of states of perception, desire and belief, for example, peppers our conversation. These have recently been described as evidence for an implicit ‘theory of mind’ (Frith and Frith, 1999). Thus by the age of 5 most children have discovered that they and others are fallible subjects of experience, who glimpse the world from eccentric points of view and are prey to deception and misapprehension. An influential account of autism suggests that the core impairment in this condition stems from the failure to acquire such a ‘theory of mind’ (Baron-Cohen, 1995). This sense of self-consciousness echoes the idiomatic use of the term, and although we tend to regard self-consciousness in its colloquial sense as a social disadvantage, we would not really want to be without it: only a nuance separates the valuable ability to inform ourselves about the impression we are making on others from the awkward encumbrance of ‘self-consciousness’.

### (v) *Self-consciousness as self-knowledge*

Like consciousness, self-consciousness has an extended final sense. It can refer to our knowledge of the broad social and cultural background which shapes us: thus my ‘idea of me’ takes in not just a body and a mind but membership of a cultural and linguistic community, a profession, a family group. In this extended sense our self-consciousness evolves throughout our lives, as it has done through the course of history. It finds its richest expression in self-portraiture and autobiography, activities of which most human children, but no other animals, are enthusiastic practitioners from an early age.

How do consciousness and self-consciousness in their various senses interrelate? Superficially they appear to be independent. There is no compelling reason to regard self-consciousness in any of the important senses we have distinguished (iii–v) as an absolute prerequisite for the three varieties of consciousness which we have teased apart. But we need to beware of prejudging difficult issues with



**Fig. 1** One of the illustrations from Berger's first paper 'On the Electroencephalogram of Man' (Berger, 1929). The recording (upper trace) was made from his close-shaven son Klaus, using electrodes on forehead and occiput. The lower trace represent time in 1/10ths of a second. Berger described the regular high amplitude oscillations seen in this trace as 'alpha' waves in his second paper (Berger, 1930).

premature definition. There are at least two connections in which my definitions may prove misleading. I will highlight these here, and return to the difficulties they point towards later.

First, I have taken the existence of 'qualia', the supposed subjective qualities of experience, on trust, but in the philosophical literature their existence is highly contentious (Dennett, 1988; O'Regan and Noe, 2001). Reductive analyses of consciousness hold out the hope that experience, in all its subjective richness, can be exhaustively described in the objective terms of function and physiology (Dennett, 1991). If it can, the characterization of consciousness which I have offered in its second sense, as the inescapably subjective contents of awareness, may lure us into the pursuit of an illusion. Reductionists seek to dispel this illusion, and in so doing to cure us of the misunderstandings which, allegedly, give rise to the 'problem of consciousness'. I shall examine the merits of these views in Section VI.

Secondly, although I have argued that consciousness is independent of self-consciousness, possible interconnections between them compose a recurring theme in the literature, echoing consciousness' connotation of knowledge *shared*. The idea is in play, for example, in Mead's description of consciousness as 'self-address using significant symbols' (Natsoulas, 1983), in Rosenthal's conception of consciousness as involving 'the thought that one is in [a certain] mental state' (Rosenthal, 1986) and in Nicholas Humphrey's suggestion that consciousness evolved to provide Machiavellian insight into the minds of others by way of awareness of one's own mental states (Humphrey, 1978). Most, but probably not all, of these links between consciousness and self-consciousness, imply a social origin for consciousness. We will take this issue up again in Section V.

### III. The science of wakefulness

A full biological account of consciousness would specify its mechanisms, functions, phylogeny and ontogeny. I will concentrate on its mechanisms in the following two sections, reviewing first our knowledge of the neural basis of sleep and wakefulness (Section III), and then our understanding of the neural processes which underly the contents of awareness, particularly visual awareness (Section IV).

Two interlacing strands of enquiry have informed the study of wakefulness over the past century: the investigation of the

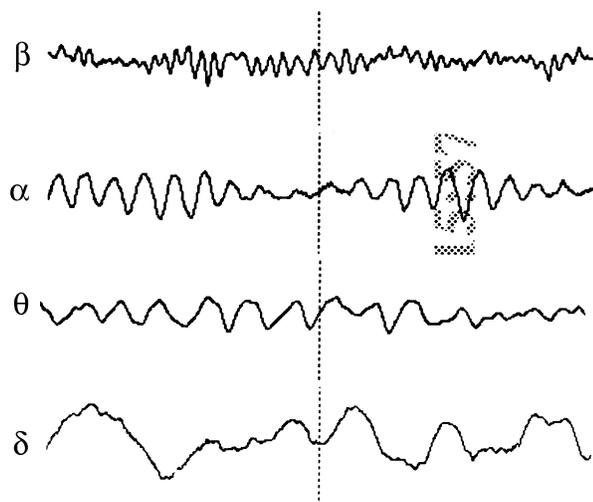
electrical correlates of states of consciousness, and the discovery that critical structures in the brainstem, thalamus and basal forebrain regulate conscious states.

#### (a) *The electricity of the brain*

Sensory evoked potentials were recorded from the brains of experimental animals by physiologists in Britain and continental Europe during the 19th century (Brazier, 1961). Several of these scientists also noted the occurrence of spontaneous electrical activity over the cortex of unstimulated animals, but it was not until 1929 that Hans Berger, a psychiatrist working in Jena, published his landmark observations 'On the Electroencephalogram of Man' (Berger, 1929) (Fig. 1). Although Berger's foremost achievement was to demonstrate that spontaneous activity could be recorded from the human brain with extracranial electrodes, his underlying purpose was to elucidate the physical basis of consciousness. His first paper closed with a series of questions which were to launch a fertile, continuing programme of research: how is the EEG affected by sensory stimulation, by sleep, by drugs which alter mental state and by intellectual activity?

Berger distinguished two contrasting rhythms of wakefulness: 'alpha' at 8–13 Hz which characterizes the 'passive EEG', typically recorded from occipital electrodes in a waking subject with eyes closed; and 'beta' rhythm, occurring at frequencies >13 Hz, the 'active EEG' which accompanies mental exertion. It was soon appreciated that slower rhythms ('theta' waves, at 4–7 Hz, and 'delta' at <3.5 Hz) at higher amplitudes characterize states of reduced arousal in adults. Their cyclical involvement in sleep became apparent in the 1950s, particularly from the work of Kleitman and his collaborators (Fig. 2).

In 1955 Aserinsky and Kleitman reported the repeated occurrence of periods of 'rapid eye movement sleep' in the course of the night; sleepers woken during these periods were likely to report concurrent dreams (Aserinsky and Kleitman, 1955). Two years later Dement and Kleitman demonstrated the cyclical structure of sleep on the basis of observations of eye movements, body movements and EEG appearances in normal sleepers (Dement and Kleitman, 1957). This work helped to define the distinction between 'slow wave sleep' (SWS), associated with a high proportion of delta activity in the EEG, and 'rapid eye movement' (REM) or 'paradoxical' sleep, during which the features of the EEG

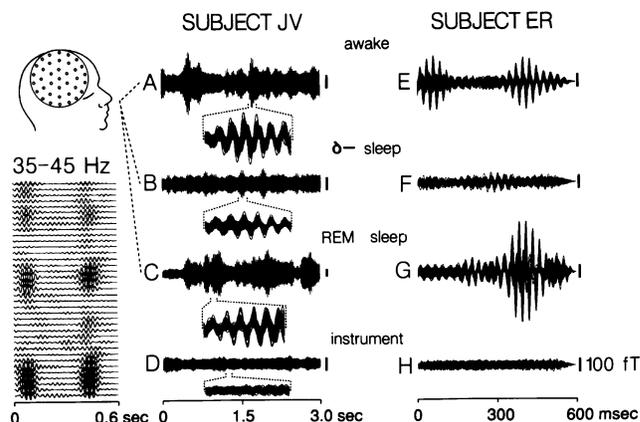


**Fig. 2** Records from diagnostic encephalograms performed in four different patients exemplifying beta ( $\beta$ ) rhythm ( $>14$  Hz); alpha ( $\alpha$ ) rhythm (8–13 Hz); theta ( $\theta$ ) rhythm (4–7 Hz) and delta ( $\delta$ ) rhythm ( $<4$  Hz). The theta trace was obtained from a sleep encephalogram; the delta trace from a wakeful, but confused, patient with an encephalopathy. In each case the dotted line bisects a 2 s sample.

resemble those in the waking state, although subjects are paradoxically difficult to arouse. Predictable cycles of descent into SWS, followed by ascent into REM sleep, recur four or five times each night, with decreasing proportions of SWS and increasing proportions of REM sleep as the night proceeds (Chokroverty, 1999).

These behavioural and electrical observations have helped to define three principle states of consciousness in health: wakefulness, REM sleep and SWS. This view has historical precedent: the Upanishads, dating from around 2000 BC, define the same three basic states (Jones, 1998). While SWS has sometimes been characterized as a state of electrical synchronization in contrast to the ‘desynchronized’ EEG of wakefulness and REM sleep, interest has been aroused recently by the possibility that activity synchronized across the brain in the gamma frequency range, at 35–45 Hz, may be a signature of the waking state and REM sleep (Llinas and Ribary, 1993; John, 2000) (Fig. 3). If so, such activity might provide an explanation for the unity of the contents of awareness. We shall return to this topic.

More generally, Berger’s discovery and its subsequent developments suggest a tendency to widespread synchronization of brain activity, whose functional significance has not yet been fully unravelled. The existence of mechanisms for synchronization helps to explain why epilepsy—resulting from the pathological synchronization of cerebral activity—should be such a common disorder of consciousness. In its applications to the topics which Berger thought promising at the end of his first paper, the EEG has proved hugely fruitful: sensory evoked potentials, the growing variety of cognitive potentials and the Bereitschaftspotential (a ‘readiness potential’ which can be recorded over motor areas up to 1 s before the execution of a movement) have demonstrated that



**Fig. 3** Recordings of rapid 35–45 (gamma) oscillations in wakefulness, delta, or slow wave, sleep and rapid eye movement (REM) sleep made using magnetoencephalography (Llinas and Ribary, 1993; with permission—copyright 1993, National Academy of Sciences, USA). The diagram at top left indicates the distribution of sensors over the head; recordings from these sensors, filtered at 35–45 Hz, are shown below. The figures at right show superimpositions of these oscillations in two subjects, during wakefulness, slow wave sleep and REM sleep. Note the differing time bases of the two recordings. The amplitude of synchronized gamma oscillations is markedly diminished in slow wave sleep in comparison with wakefulness and REM sleep.

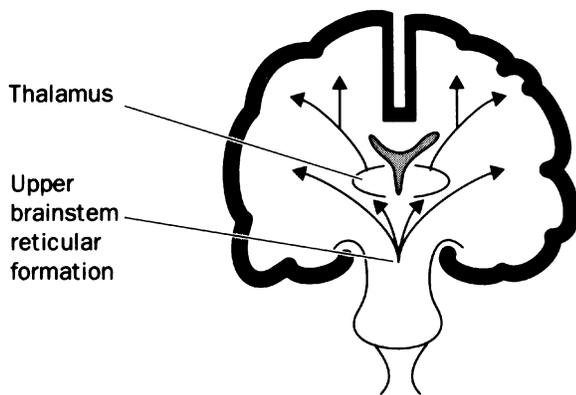
sensation, attention, intellectual activity and the planning of movement all have distinctive electrical correlates at the surface of the skull (Kutas and Dale, 1997).

## (b) *The control of conscious states*

### (i) *Anatomy*

Clinico-pathological studies made at the time of the epidemic of encephalitis lethargica which occurred during and after the First World War suggested to the Viennese pathologist Constantin Von Economo that structures in the upper brainstem and posterior hypothalamus mediate arousal (Von Economo, 1931). Frederic Bremer later confirmed this suggestion experimentally by showing that transection of the cat’s brain at the cervicomedullary junction had no effect on arousal, or on the sleep–wake cycle, while transection through the midbrain brought about a state resembling deep sleep (Bremer, 1929).

Bremer hypothesized that this impairment of arousal resulted from interruption of ascending sensory pathways in the midbrain. His student Giuseppe Moruzzi, working with Horace Magoun, later showed that the critical areas were not, in fact, in the sensory pathways but rather in the reticular core of the upper brainstem and, probably, their thalamic targets (Moruzzi and Magoun, 1949). Moreover, electrical stimulation of this region in a drowsy animal ‘activated’ the EEG. These observations gave birth to the concept of the ‘ascending reticular activating system’ (ARAS) (Fig. 4). While the central insight of this concept, that structures in the brainstem regulate our states of consciousness, still holds true, a much more complex picture has emerged since the



**Fig. 4** A schematic diagram of the reticular activating system, indicating pathways of activation which involve and those which bypass the thalamus. (Zeman, 1997; by permission of *The Lancet*, London, UK.)

pioneering work of Moruzzi and Magoun. The ARAS is no longer regarded as a monolithic unit, nor as a system restricted to the classically defined 'reticular' nuclei of the brainstem; indeed, activating structures are not confined to the brainstem at all, and their influence descends to the spinal cord as well as ascending to the cerebral hemispheres.

Rather than revealing any single 'place where consciousness dwells', the exploration of these structures is identifying a series of somewhat specialized nodes in a complex network controlling aspects of arousal (Fig. 5). It would be surprising if functions as fundamental as the maintenance of wakefulness or the control of the sleep-wake cycle depended exclusively and unalterably on any single region of the brain (Steriade, 1999). Experimental work in animals suggests that the following structures play key roles in the maintenance and modulation of wakefulness: cholinergic nuclei in the upper brainstem and basal forebrain; noradrenergic nuclei, in particular the locus coeruleus; a histaminergic projection from the posterior hypothalamus; and probably dopaminergic and serotonergic pathways arising from the brainstem (McCarley, 1999) (Fig. 6). Much, but not all, of the influence exerted by these pathways is mediated by the thalamus, which can be regarded as the apex of the ARAS, as well as a critical synaptic relay for most sensory and many intracerebral pathways (Jones, 1998) (Fig. 7). The function of these activating structures is not, of course, confined to the maintenance of wakefulness: they are of profound importance to a wide range of interrelated functions including mood, motivation, attention, learning, memory and movement (Robbins and Everitt, 1993; Marrocco *et al.*, 1994).

Some specific contributions made by these and other structures to the regulation of conscious states have been defined. For example, the suprachiasmatic nucleus of the hypothalamus has emerged as the timekeeper of consciousness (Kilduff and Kushida, 1999). It normally entrains the sleep-wake cycle to the alternation of night and day under the influence of the direct retinohypothalamic projection. Transection experiments by Jouvet and subsequent work have established the key importance of cholinergic nuclei at the

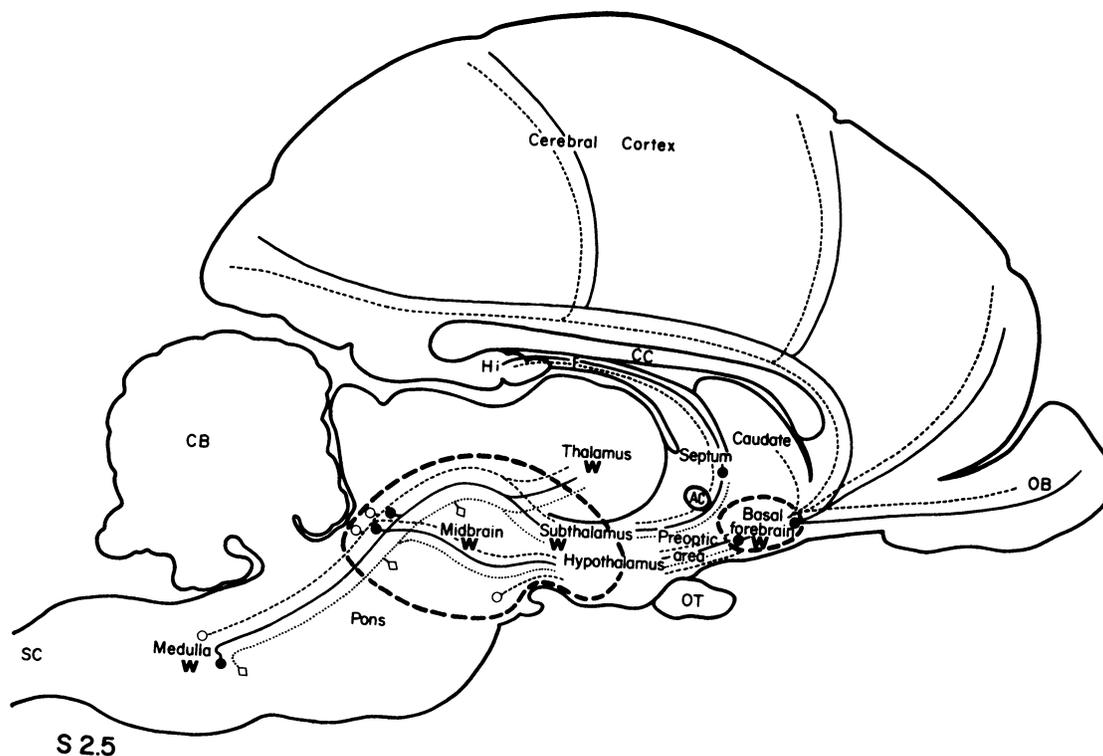
pontomesencephalic junction, the laterodorsal tegmental and pedunculopontine nuclei, in orchestrating the phenomena of REM sleep, i.e. activation of the EEG, pontogeniculoccipital waves, rapid eye movements and muscle atonia (McCarley, 1999). The location of structures crucial for the induction of SWS remains uncertain (McCarley, 1999). A reduction of activity in the cholinergic, noradrenergic and histaminergic nuclei which maintain wakefulness is clearly relevant (Zolotowski *et al.*, 1999); candidates for a critical role in sleep induction include the anterior hypothalamus (Sherin *et al.*, 1996) and basal forebrain (McCarley, 1999).

Recent functional imaging studies of regional brain activity in the three major behavioural states in man have corroborated and extended the conclusions drawn from animal experiment and clinical observations. Global cerebral glucose metabolism falls in SWS by ~20%, rising back to, or even above, waking levels in REM (Heiss *et al.*, 1985; Buchsbaum *et al.*, 1989). During SWS regional blood flow declines, in proportion to the amount of slow wave activity on EEG, in the rostral brainstem, thalamus, prefrontal and cingulate cortex (Hofle *et al.*, 1997; Macquet *et al.*, 1997). In REM sleep regional blood flow increases in the rostral brainstem, thalamus and limbic regions, in keeping with the electrical and subjective features of dreaming sleep, but declines in prefrontal and posterior cingulate cortex, and in some regions of parietal cortex (Macquet *et al.*, 1996). Variations in the level of arousal during wakefulness also appear to correlate with levels of activity in the structures of midbrain and thalamus which regulate conscious states: Kinomura and colleagues reported activation of the midbrain tegmentum and intralaminar nuclei of the thalamus by the transition from a resting state to the performance of visual and somatosensory reaction time tasks (Kinomura *et al.*, 1996); Paus and colleagues have described a decrease in midbrain and thalamic activation during a tedious 1 h auditory detection task, associated with declining performance and increasing slow wave activity on EEG (Paus *et al.*, 1997). Finally, there is evidence that the loss of consciousness induced by some anaesthetics is associated with selective depression of thalamic function, linking the mechanisms of anaesthesia and sleep (Fiset *et al.*, 1999; Alkire, 2000).

## (ii) Physiology

It should, in principle, be possible to explain the behavioural and electrical features of the three major states of consciousness in terms of the characteristics of relevant neuronal types and the networks in which they are organized. Substantial progress has been made in this direction. This is well illustrated by way of the contrast between neuronal activity during sleep and during wakefulness within the thalamus.

In the waking state thalamocortical neurones are tonically depolarized by cholinergic, noradrenergic and histaminergic inputs from the brainstem and hypothalamus, which block a hyperpolarizing potassium conductance (Steriade *et al.*, 1990,



**Fig. 5** A sagittal drawing of a cat brain indicating the structures implicated in generating and maintaining the waking state (from Jones, 1998; with permission—copyright 1998; Lippincott Williams and Wilkins, Philadelphia, USA). Areas marked with a W are those from which electrical stimulation elicits, and where cells are maximally active during, wakefulness. Areas encircled by dashed lines in bold are those where selective lesions most commonly cause coma. These regions contain glutamatergic neurones of the reticular formation (open diamonds), noradrenergic and other catecholaminergic neurones (open circles) and cholinergic neurones (filled circles). Projections from the thalamus are not shown. AC = anterior commissure; CB = cerebellum; CC = corpus callosum; Hi = hippocampus; OB = olfactory bulb; OT = optic tract; S = sagittal; SC = spinal cord.

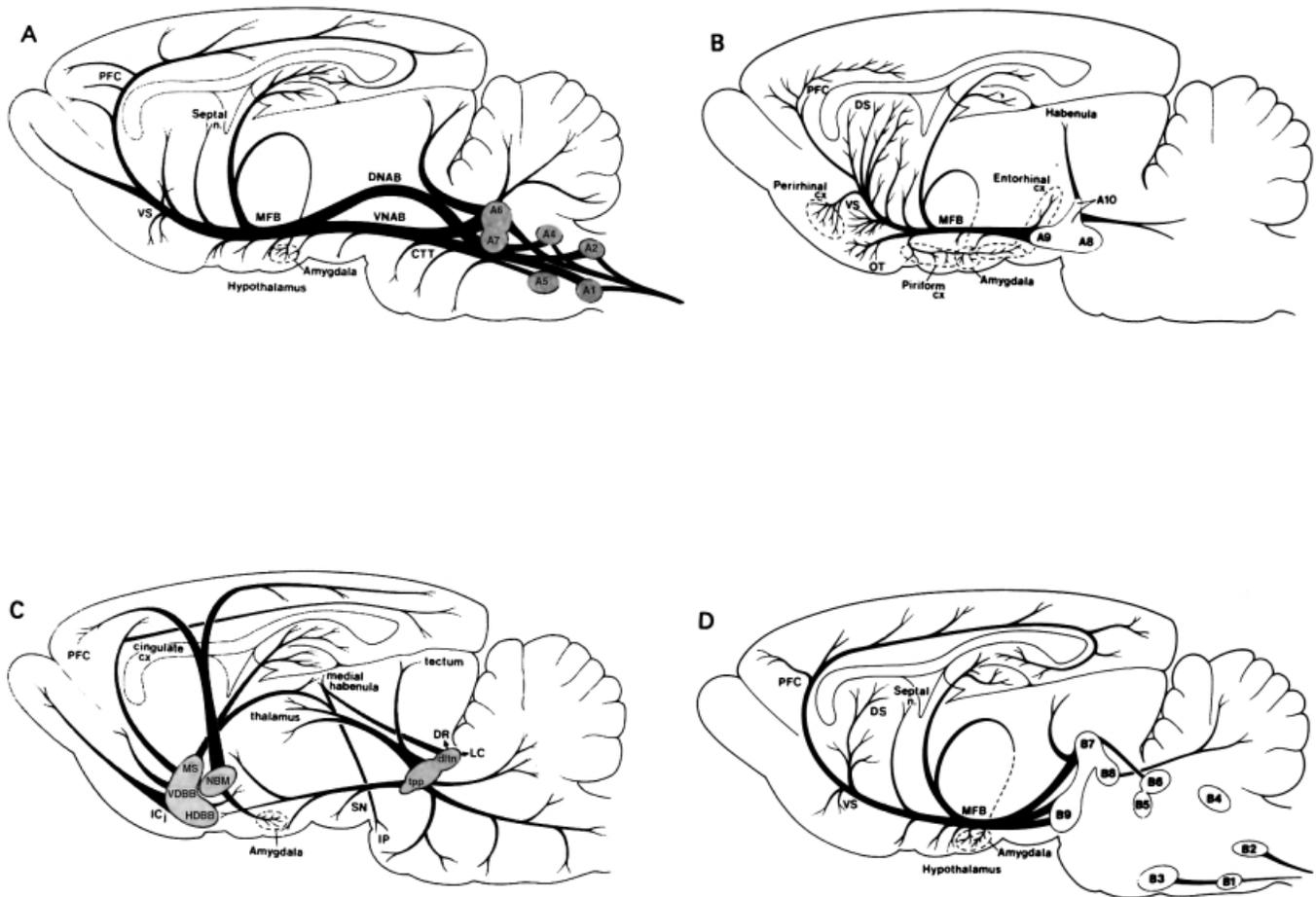
1993; McCarley, 1999; Steriade, 1999). This induces a ‘spike mode’ of response in thalamocortical cells, permitting faithful onward transmission of afferent signals to the thalamus. The reduction of the depolarizing input in SWS induces a contrasting ‘burst mode’ of response, dependent upon a low threshold calcium conductance, which predisposes these cells to patterns of repetitive discharge while hyperpolarized (Fig. 8). The simultaneous disinhibition of the reticular nucleus of the thalamus in early sleep, following reduction of inhibitory cholinergic input from the brainstem, allows it to exert a synchronized GABAergic inhibition of thalamocortical cells, which ultimately gives rise to the ‘spindles’ abounding in stage 2 sleep. Further hyperpolarization of thalamocortical cells, as sleep deepens, allows them to participate in slow wave oscillations to which the individual and network properties of thalamocortical cells, corticothalamic cells and neurones of the reticular nucleus all contribute. Reduction of direct non-specific excitatory inputs to the cortex, as well as effects occurring primarily at the level of the thalamus, is conducive to the generation of these rhythms. Thus the distinction at an electrophysiological level between spike and burst modes of response in thalamocortical neurones corresponds with the behavioural distinction between the responsiveness of the waking state

and the inaccessibility of sleep, and mirrors the global shift between the waking and the sleeping EEG.

### (iii) Pharmacology

As we have seen, the pharmacological dissection of the ‘reticular activating system’ has revealed the presence of several chemically distinct but interactive subsystems: cholinergic, noradrenergic, dopaminergic, serotonergic and histaminergic (Robbins and Everitt, 1993; Marrocco *et al.*, 1994). The actions of each of these transmitters are complex, depending upon the site of release and the nature of the receptor targeted. Nonetheless it is at least clear that the firing of cells in cholinergic and noradrenergic brainstem nuclei is ‘state dependent’ and correlates with wakefulness (McCarley, 1999).

Evidence that REM sleep is dependent upon activity in cholinergic nuclei, while noradrenergic and serotonergic nuclei are least active in this phase of sleep has given rise to a ‘reciprocal interaction’ model of sleep architecture. This proposes that the regular alternation of SWS and REM sleep over the course of the night is regulated by the waxing and waning of mutually inhibitory activity in these nuclei. The pharmacological basis of ‘sleep debt’, i.e. the increasing pressure to sleep as the period of wakefulness extends,



**Fig. 6** The pharmacologist's view of the activating system (from Robbins and Everitt, 1993; with permission—copyright 1993, MIT Press, USA): (A) shows the origin and distribution of the central noradrenergic pathways in the rat brain, (B) the dopaminergic pathways, (C) the cholinergic pathways, (D) the serotonergic pathways. CTT = central tegmental tract; dltn = dorsolateral tegmental nucleus; DNAB = dorsal noradrenergic ascending bundle; DR = dorsal raphe; DS = dorsal striatum; HDBB = horizontal limb nucleus of the diagonal band of Broca; Icj = islands of Calleja; IP = interpeduncular nucleus; LC = locus ceruleus; MFB = medial forebrain bundle; MS = medial septum; NBM = nucleus basalis magnocellularis (Meynert in primates); OT = olfactory tubercle; PFC = prefrontal cortex; SN = substantia nigra; tpp = tegmental pedunculopontine nucleus; VDBB = vertical limb nucleus of the diagonal band of Broca; VNAB = ventral noradrenergic ascending bundle; VS = ventral striatum.

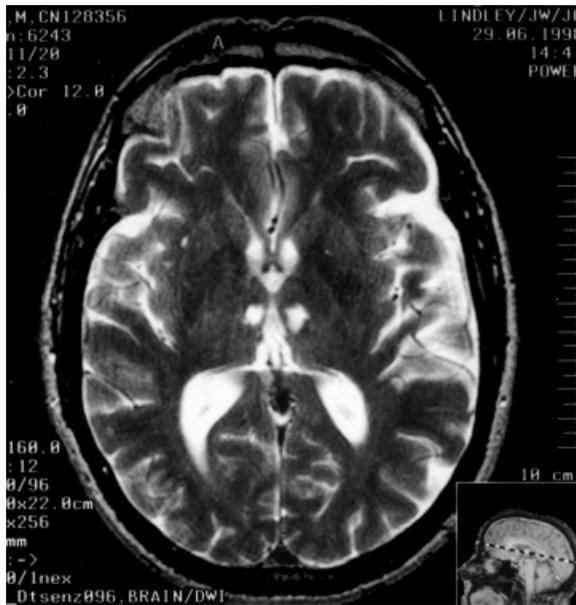
remains a confusing area. A number of potential 'hypnotoxins', i.e. sleep-promoting substances, usually accumulating during wakefulness, have been identified (Zolotki *et al.*, 1999). Recent evidence suggests that a particularly important role may be played by elevation of extracellular adenosine concentrations during wakefulness, with resulting inhibition of activating cholinergic nuclei in the upper brainstem and basal forebrain (McCarley, 1999).

Future work on the pharmacology of wakefulness is likely to demonstrate distinctive roles for the neurotransmitters of the 'activating system' in modulating different aspects of arousal. 'Wakefulness', after all, is shorthand for a set of associated neural, behavioural and psychological functions which are, to some extent, independently controlled. This is evident from a range of pathological states in which the usual associations between these states break down. Thus, for example, a sleepwalker is capable of coordinated movement, may be able to avoid obstacles and accede to

gentle urging to return to bed, but will later have no recall of the episode; motor control, perception and memory have lost their usual relationship (Mahowald and Schenck, 1992). In work exploring the idea that the neurotransmitters linked with arousal make distinctive contributions, Robbins and Everitt have found, using a consistent set of behavioural tests, that selective damage to the noradrenergic system impairs selective attention, damage to the cholinergic system impairs baseline accuracy, damage to the dopaminergic system lengthens response latency and probability of response, and damage to the serotonergic system leads to impulsive responding (Robbins and Everitt, 1995).

### (c) Pathologies of wakefulness

The key roles of the brainstem and thalamus in the maintenance of wakefulness help to explain the pathologies of arousal which follow structural brain damage.



**Fig. 7** Axial MRI scan in a patient who presented with depression of consciousness. Symmetrical high signal abnormalities are seen in the thalami. The aetiology was thought to be ischaemic. A common vascular supply to both sides of the thalamus from a single posterior cerebral artery is common and explains the occurrence of symmetrical bilateral abnormalities (Warlow *et al.*, 1996).

Coma is a state of continuous ‘eyes-closed’ unconsciousness, in the absence of a sleep–wake cycle (Plum and Posner, 1982; Plum, 1991; Schiff and Plum, 2000). It varies in degree from mild to profound unresponsiveness, and is associated with a comparably variable reduction in cerebral metabolism. It results from diffuse hemispheric or focal brainstem/thalamic injury and is usually a transitional state, *en route* to full recovery, brainstem death or a state of chronically impaired awareness with recovery of the sleep–wake cycle. The risk of confusing the ‘locked-in state’ with coma is now well recognized. In this syndrome, which follows brainstem lesions abolishing the descending control of voluntary movement, patients are only able to communicate using movements of the eyes or eyelids.

Brainstem death (Pallis and Harley, 1996) implies the irreversible loss of all brainstem functions. In the United Kingdom it renders legal the removal of organs for transplantation, provided that appropriate consent has been obtained. It is generally followed by cardiac death, within hours to weeks, although there are reported exceptions to this rule (Shewmon, 1998).

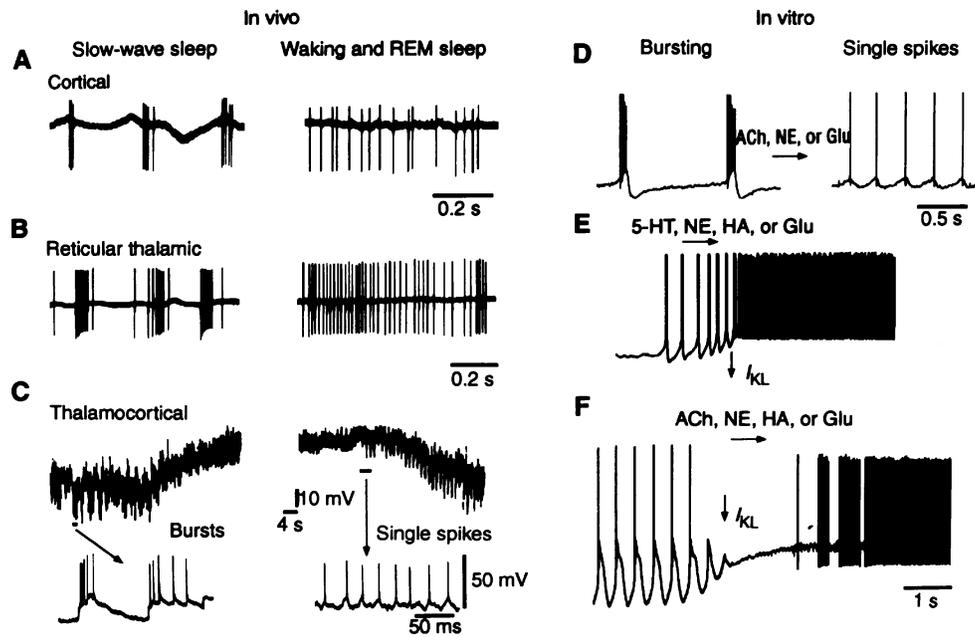
The vegetative state, described by Jennett and Plum in 1972 (Jennett and Plum, 1972; Multi-Society Task Force on PVS, 1994; Zeman *et al.*, 1997), is in a sense the converse of brainstem death: in this condition, characterized by ‘wakefulness without awareness’, patients regain their sleep–wake cycle, and may be aroused by painful or salient stimuli, but show no unambiguous signs of conscious perception or deliberate action, including communicative acts. Recovery

from a vegetative state often occurs: younger age and a traumatic, rather than hypoxic–ischaemic, causation improve the outlook. After 1 month the state is often termed ‘persistent’, and in patients in whom recovery appears highly unlikely it may be deemed ‘permanent’, although permanence cannot be predicted with certainty. The underlying pathology usually involves (i) diffuse cortical injury, typically cortical laminar necrosis; (ii) diffuse white matter injury, typically diffuse axonal injury or leucoencephalopathy; or (iii) thalamic necrosis. ‘Minimally responsive states’ are sometimes distinguished from the vegetative state; these involve some consistent or inconsistent evidence of intelligent awareness in the presence of profound continuing physical and cognitive impairment. Akinetic mutism is a related state of profound apathy with evidence of preserved awareness, characterized by attentive visual pursuit and an unfulfilled ‘promise of speech’. It is often associated with damage to the medial frontal lobes.

The distinctions between coma, brainstem death and the vegetative state are useful and moderately robust in clinical practice. But they are not immune to practical or theoretical criticism. At a practical level, there are apparent examples of long survival in ‘brain dead’ patients (Shewmon, 1998), and evidence that the vegetative state is often misdiagnosed in patients who are, in fact, aware of their surroundings (Childs *et al.*, 1993; Andrews *et al.*, 1996). At a theoretical level, it is conceivable that brainstem death may become a treatable disorder as neural prostheses are developed, and it is open to question whether patients in vegetative states are wholly unaware. These clinical concepts are still in evolution.

#### IV. The science of awareness

If the study of healthy and disordered states of consciousness has clarified the neural basis of arousal, the study of vision has provided the outstanding source of evidence on the mechanisms underlying the contents of consciousness. Research in animals and man is defining increasingly fine-grained correlations between cerebral activity and visual experience, including experiences which occur, like visual hallucinations, or change, like ambiguous figures, without any corresponding changes in external stimuli. A second, complementary, approach has been opened up by the distinction between ‘explicit’ neural processes which directly give rise to conscious awareness, e.g. conscious vision, and ‘implicit’ processes, e.g. blindsight, which allow visuomotor performance in the absence of awareness. Subtracting the processes required for implicit perception from those involved in explicit perception should, in principle, help to delineate the key neural substrates of awareness. Work on learning has developed a parallel distinction between declarative and procedural memory. A third approach, linked to the last, looks toward action: changes in cerebral activity, as consciously directed action becomes automatized or otherwise divorced from conscious control, may illuminate the neurology of consciousness. I shall outline these approaches in turn.



**Fig. 8** State-dependent activity in thalamic and cortical neurones (from Steriade *et al.*, 1993; with permission—copyright 1993, The American Association for the Advancement of Science, USA). Neurones from the cerebral cortex of chronically implanted, behaving cats, in the cerebral cortex (A), reticular thalamic nucleus (B) and thalamic relay nuclei (C) change their activity from rhythmic spike bursts during natural slow-wave sleep to firing of single spikes during waking and rapid eye movement sleep. Similar changes can be demonstrated *in vitro* in response to neurotransmitters involved in modulating sleep and wakefulness. (D) Cortical cell; (E) reticular thalamus cell; (F) thalamic relay cell. Depolarization results from the reduction of specialized conductances including  $I_{KL}$ , a potassium conductance. ACh = acetylcholine; Glu = glutamate; HA = histamine; 5-HT = serotonin; NE = noradrenaline.

### (a) Exquisite correlations

Although it is not usually framed in these terms, one of the main goals of visual physiology is to account for the contents of visual consciousness. A series of discoveries in the course of this century have revolutionized our picture of the cerebral events which underlie conscious vision. Key findings include the definition of the retinotopic map in striate cortex (Holmes and Lister, 1916); the discovery of its orientation-specific columns by Hubel and Wiesel (Hubel and Wiesel, 1977); the realization that 30–40 functionally and anatomically visual areas surround area V1 (Covey, 1994); the evidence that parallel, though interconnected, streams of visual information flow through these areas, subserving the perception of form, colour, depth and motion (Livingstone and Hubel, 1988); and the broad distinction between an occipitotemporal stream concerned with object identification and an occipitoparietal stream concerned with visually-guided action (Milner and Goodale, 1995).

A few examples will illustrate the grain of the correlations which have emerged between visual experience and regional activity in the visual system. Comparison of regional brain activation by coloured and by moving grey-tone visual stimuli indicates that the former strongly activate an area in the fusiform gyrus, possibly the human homologue of area V4, which has been shown in monkeys to contain a high proportion of colour selective neurones (Zeki, 1993). By

contrast, moving achromatic stimuli selectively excite a more lateral region, at the occipitotemporal junction, termed ‘human V5’ by Zeki and colleagues. The anatomical distinction between these regions helps to explain the existence of selective deficits of colour and movement perception in man occurring after focal cortical lesions, respectively cerebral achromatopsia and akinetopsia (Zeki, 1990, 1991). Zeki has extended these findings by showing that a figure which gives rise to a compelling illusion of movement also selectively activates human V5 (Zeki *et al.*, 1993). There is some evidence that activity in area V5, in the absence of striate cortex, may be sufficient to generate the conscious experience of visual movement (Barbur *et al.*, 1993), although this has recently been challenged (Covey and Walsh, 2000). As visual information travels further down the stream of visual areas, retinotopic mapping tends to become less precise and the stimulus requirements of the local neurones become more complex. These downstream areas, demonstrated by functional MRI, are illustrated by the ‘fusiform face area’ (Kanwisher *et al.*, 1997), an area in the human fusiform gyrus selectively responsive to faces, and the ‘parahippocampal place area’ (Epstein and Kanwisher, 1998), a region responsive to the visual appearance of locations.

Inferences about the generation of visual awareness, drawn from work of this kind, are open to the potential objection

that, although these areas are activated by appropriate stimuli, they may not mediate the conscious experience of vision. Correlation does not imply cause, and, after all, much of the work on cortical visual responses in animals has been performed under anaesthesia. It is open to question, for example, whether area 17, primary visual cortex, directly contributes to visual awareness (Crick and Koch, 1995; Rees *et al.*, 2000). This objection can be met, at least in part, by the use of paradigms in which visual awareness changes while external stimulation is held constant. Change in neuronal activity, detected under these circumstances, is more likely, although not guaranteed, to be linked specifically to changes in awareness. Work on visual imagery, hallucinations, attentional shifts and binocular rivalry has exploited this strategy.

We can summon up visual images ‘in the mind’s eye’ and interrogate them much as we do a real visual scene (Kosslyn and Shin, 1994). Shepard, Kosslyn and colleagues have shown that, in many respects, mental images are processed in similar ways to percepts of items in the real world (Shepard, 1978; Kosslyn and Shin, 1994). For example, decision tasks involving a minor mental rotation of complex shape are performed more quickly than tasks involving more extensive rotation, and the time required to search an imagined map is proportional to the distance travelled in the mind’s eye. Recent reports, using functional imaging, indicate that these similarities in processing extend to similarities in the brain regions activated by real and imagined objects: for example, imagining a small object activates the representation of central vision in striate cortex, while imagining a large object activates more peripheral parts of the cortical map (Kosslyn *et al.*, 1995); mental rotation tasks engage areas involved in tracking moving objects and encoding spatial relations (Cohen *et al.*, 1996); imagining faces and places respectively excite the fusiform face area and hippocampal place area mentioned earlier (Kanwisher, 2000). This line of research suggests that the neural correlates of mental imagery overlap substantially with the correlates of perception.

Like mental images, visual hallucinations are visual percepts which occur in the absence of a corresponding external stimulus, but, unlike images, hallucinations are perceived as if they originated in the external world. A wide range of pathologies is associated with complex visual hallucinations, including partial epilepsy, structural disorders of the central visual pathways, disorders of the brainstem causing peduncular hallucinosis, narcolepsy, drug-induced hallucinosis, psychotic disorders and peripheral visual disorders giving rise to the Charles Bonnet syndrome (Manford and Andermann, 1998). A recent functional imaging study of this last condition found that hallucinations of faces, colours, textures and objects were associated with activity in the ventral occipital lobe (ffytche *et al.*, 1998). Comparable results have been reported recently in the auditory domain (Griffiths, 2000). The blood flow changes detected by functional imaging, which usually follow the onset of the responsible stimulus by an appreciable delay, occurred in

these patients several seconds before the onset of the hallucinations, suggesting that the neural activity responsible for a hallucination must evolve to a certain state before it gives rise to a conscious percept.

States of imagery and hallucination are rather special cases of perception. The very best opportunities to isolate the neural correlates of ordinary visual perception may come instead from studies of attentional shifts and ‘multistable’ visual percepts.

Attention is the sentry at the gate of consciousness: ‘My experience is what I agree to attend to’ (James, 1890).<sup>3</sup> The essence of attention is ‘selection’: whether we are displaying ‘preparatory attention’ as we await an anticipated event, switching our attention between the senses or between the targets presented to a single sense, or sustaining our attention to a task, we are excluding a range of rival stimuli from the focus of our interest. Changes in the neural representation of items as they move in and out of the focus of attention should shed light on the neural accompaniments of consciousness. These changes have been termed the neural ‘expression’ of attention (LaBerge, 1995).

Single cell recording from monkeys trained to shift visual attention without altering their gaze indicates that firing rates are increased in cells responding to attended stimuli, and reduced in cells responding to unattended stimuli in extrastriate visual areas, for example, areas V4 and V5 (Moran and Desimone, 1985; Treue and Maunsell, 1996). Recent functional imaging studies suggest that the neural expression of attention in man also involves focal enhancement and inhibition of neural activity; for example, switches of attention between faces and places presented simultaneously are associated with detectable modulations of activity in the fusiform and parahippocampal regions mentioned above (Kanwisher, 2000).

Multistable or ambiguous visual stimuli, like the Necker cube, which appears to reverse in depth during protracted viewing, are open to alternative visual readings. If incompatible stimuli are presented to the two eyes, they also tend to generate a ‘multistable’ percept, as most viewers experience an alternation between the two images rather than a fusion. This paradigm has been applied both to animals and man to study the neural correlates of the alternating percept. Logothetis, working with monkeys, has reported that while many neurones in visual areas respond to the stimuli throughout their presentation, regardless of the current conscious percept, a subset of extrastriate neurones recorded in V4 and V5 raise or lower their firing rate markedly as the stimulus to which they respond gains or loses perceptual dominance (Logothetis and Schall, 1989; Leopold and Logothetis, 1996). Work by Engel and colleagues suggests that cells responding to the currently perceived member of a pair of rivalrous stimuli synchronize their discharges during

<sup>3</sup>Although attention may not be sufficient for consciousness. In the context of blindsight, Kenridge *et al.* (1999) have shown that spatial cues can enhance performance in detection tasks in which both cue and stimulus are unseen.

the period of perceptual dominance to a greater degree than during periods of suppression (Engel *et al.*, 2000).

Further down the processing stream, in experiments with human subjects, the modulation of neuronal activity in the fusiform face area and parahippocampal place area, as simultaneously presented faces and places alternate in awareness, is of similar size to the modulation seen when faces and places are alternately presented (Kanwisher, 1998). Thus, by this stage of processing in the human brain, activity correlates with the contents of awareness rather than with the raw features of the impinging stimuli. Using magnetoencephalography, Tononi and colleagues have reported that as conscious perception shifts between two gratings of different orientations, one horizontal, the other vertical, flickering at different frequencies, so the power of electromagnetic activity at the corresponding frequency waxes and wanes by 30–60% over wide regions of cortex (Tononi and Edelman, 1998*b*). Lumer and colleagues have found that right frontoparietal activation is associated with the transition between multistable percepts, suggesting that the neural control of these transitions, as opposed to the neural expression of the resulting percepts, may share common ground with the direction of spatial attention (Lumer *et al.*, 1998).

These experiments, investigating imagery, hallucinations, attentional shifts and ambiguous percepts, are beginning to capture the neural correlates of visual experience. The precise definition of the ‘neural correlate of consciousness’ in man (Koch, 1998) remains a goal for the future, and will probably require more sophisticated methods than those we currently have at our disposal, allowing the detailed measurement of disparate neuronal activity over short time scales in the human brain. Nevertheless the work we have just reviewed is helping to justify the neuroscientist’s long-held article of faith: that every distinction drawn in experience and behaviour will be reflected by distinctions between patterns of neuronal activity.

### **(b) Implicit perception**

The correlation of aspects of conscious vision with details of neural activity has been complemented by another approach, which aims to highlight the neurology of consciousness by contrasting it with the neural basis of unconscious processes. This offers an alternative solution to the problem mooted in the last section, that the correlation of neural events with conscious processes does not prove a causal relationship. ‘Subtraction’ of the neural processes associated with unconscious processing from those correlated with consciousness should help to define the fundamental neural substrate of awareness. This is, so to speak, an attempt to conquer consciousness by stealth.

Two lines of research can be distinguished. The first studies the neural effects of sensory stimulation which may have no discernible effect whatever on behaviour; the second, more provocative, project is to explore the neural consequences of

events which ‘might influence our experience, thought and action even though they themselves have not been consciously perceived’ (Kihlstrom *et al.*, 1992).

Some conceptual difficulties and terminological distinctions deserve a mention before reviewing examples of these two lines of research. The major conceptual problem for students of ‘unconscious’ or ‘implicit’ psychological processes is how to determine the presence or absence of consciousness. Much of the neuropsychological work in this area relies on verbal report (Barbur *et al.*, 1993) or the use of a ‘commentary key’ (Weiskrantz, 1997) to indicate the degree of awareness. But verbal reports and presses on commentary keys may not provide exhaustive measures of the information available to consciousness. Indeed, there are no conclusive reasons for thinking that consciousness should always be reportable, even in principle (Zeman, 2000). On the other hand, if any successful discrimination is taken to provide evidence of conscious perception, the possibility of unconscious perception is ruled out by definition (Kihlstrom *et al.*, 1992). The lack of any ‘exhaustive measure that exclusively indexes relevant conscious perceptual experiences’ is therefore a significant problem, though not necessarily an insuperable one (Merikle and Reingold, 1992). Psychologists have suggested a number of sophisticated solutions to the dilemma (Jacoby *et al.*, 1992; Merikle and Reingold, 1992).

The following terminological distinctions are helpful. Implicit—or unconscious or subliminal—perception generally refers to perception occurring in the absence of any conscious experience of the information perceived. In this respect it is distinct from ‘residual experience’, a term describing the rudimentary conscious experience which sometimes survives the substantial loss of a sense: for example, residual but clearly conscious experience of visual movement sometimes survives the substantial loss of sight caused by injuries to striate cortex (Weiskrantz, 1980; Barbur *et al.*, 1993). Whether or not some examples of vaguely sensory ‘knowledge’ surviving such damage represent residual experience has given rise to debate (Zeki and ffytche, 1998). A second useful distinction is between explicit and implicit processes and direct and indirect tasks (Weiskrantz, 1997). Explicit processes are those which yield overt knowledge of the material on which a given act or judgement is based; for example, if I can straightforwardly see the target at which I take my aim, the visual process is explicit and, if I can later recall the event, my process of recollection is as well. Direct tasks, by contrast, are those which involve instructions referring directly to the dimension of interest to the experimenter: a direct test of memory might ask for the contents of a word list; an indirect test might examine whether prior exposure to the words increased the ease with which they are identified on a very brief presentation. But a direct task may tap an implicit process—if, for example, we are asked to guess at the location of a visual stimulus which we have not consciously perceived—and an indirect task

may tap an explicit process, if I recognize the items from the word list on their brief presentation.

Turning to the first line of research on implicit perception, the idea that stimuli impinging on the nervous system may have neural effects in the absence of any discernible effect on 'experience, thought and action' is familiar. It is no surprise, for example, that some reflex responses, like the pupillary light reflex, can be elicited from the visual system in the absence of visual experience. Penetrating deeper into brain, cortical evoked potentials can be recorded from patients under anaesthesia (Jones, 1994; Schwender *et al.*, 1994), as they can from anaesthetized experimental animals; low concentration odours evoke activity in olfactory regions at levels which do not permit even accurate guesses about their presence (Sobel *et al.*, 1999); brief masked presentations of fearful faces activate the amygdala in the absence of acknowledged awareness (Morris *et al.*, 1998; Whalen *et al.*, 1998); some patients with prosopagnosia who are unable to classify faces as familiar or unfamiliar when asked to do so, nevertheless display galvanic skin responses to familiar ones (Bauer, 1984).

But the possibility that events which are not consciously perceived can 'affect our experience, thought and action' is more startling. Evidence in favour of this possibility flows from a number of sources, including the effects of 'unperceived' stimuli on judgements made by normal subjects in direct and indirect tasks, their effects in normal controls subjected to procedures like anaesthesia or hypnosis, and their effects in subjects with neuropsychological syndromes such as blindsight, neglect and, possibly, hysteria (Kihlstrom *et al.*, 1992).

A 19th century experiment by Pierce and Jastrow illustrates the effects of stimuli too subtle to allow confident verbal report on a judgement in a direct test. The experimenters attempted to judge, by guesswork if need be, which of two similar pressures was the greater. At the same time they indicated their degree of confidence in their judgement. Even when their confidence was rated at 0, their guesses proved correct significantly more often than chance would have allowed (Kihlstrom *et al.*, 1992). In two modern reworkings of the theme, Libet has found that after brief trains of threshold stimuli delivered directly to the cerebral cortex, which fail to excite any conscious sensation, subjects 'forced' to guess whether or not they had been delivered can do so with a success rate well above chance (Libet, 1996); a study of the functional imaging correlates of the perception of low concentration odours demonstrated above-chance detection of the odour in the absence of awareness, associated with brain activation in the anterior medial thalamus and inferior frontal gyrus (Sobel *et al.*, 1999). 'Mere exposure' effects illustrate the effects of unperceived stimuli on normal subjects in an indirect test. Abstract stimuli presented extremely briefly, for 1 ms, tend to be chosen in a subsequent task in which subjects are asked to state which of two stimuli they prefer, even though they are not recognized as familiar (Zajonc, 1980). The occurrence of implicit perception under

anaesthesia has been supported by a number of studies. Schwender and colleagues, for example, showed that, in the absence of any explicit recall of events during cardiac surgery, some patients associated key words with material which had been read aloud during the operation; these patients were distinguished by the relatively normal latency of their mid-latency auditory evoked potentials under anaesthesia (Schwender *et al.*, 1994).

In neuropsychology, the most widely quoted example of unconscious perception is undoubtedly blindsight (Stoerig and Cowey, 1997; Weiskrantz, 1997; Weiskrantz, 1998). This paradoxical term, coined in 1974 (Sanders *et al.*, 1974), describes a range of visually based abilities which can be demonstrated, in the absence of visual awareness, following damage to striate cortex, in some, but not all, subjects. The possibility that abilities of this kind might exist was suspected on the basis of the relatively good recovery of visuomotor functions in monkeys after experimental lesions of striate cortex. An experiment by Poppel and colleagues, in which war veterans with scotomata due to gun shot wounds were encouraged to look in the direction of 'unseen' flashes, suggested that similar abilities might be found in man (Poppel *et al.*, 1973). This was confirmed when D.B., a patient in whom the right calcarine cortex had been excised as part of the surgical treatment of an arteriovenous malformation, was 'urged to guess' the nature or location of stimuli in his blind field (Sanders *et al.*, 1974). His guesses, to his 'great surprise' and despite his insistence 'that he saw nothing except in his intact visual field', proved to be substantially correct. Subsequent work has shown that besides mediating neuroendocrine and reflex responses, blindsight can subserve accurate performance in a range of indirect and direct tasks (Stoerig and Cowey, 1997). Indeed, on some measures, blindsight allows levels of accuracy well beyond the performance of normal observers making judgements close to the threshold of awareness. Its capacities include localization of the 'unseen' target, by hand or eye, and simple judgements about orientation, shape and the presence or absence of motion. A range of sceptical interpretations of these results, in terms of covert eye movements, scatter of light into the intact visual field, the survival of islands of cortex and the persistence of 'degraded', but conscious, visual awareness have been substantially rebutted (Weiskrantz, 1997), although blindsight continues to provoke lively debate (Zeki and ffytche, 1998). This line of research, which had been inspired by observations made in monkeys, has recently come full circle with ingenious experimental evidence that destriated monkeys, like destriated man, may lack 'phenomenal vision' and rely on blindsight for their well preserved visuomotor skills (Cowey and Stoerig, 1995).

The study of neglect also illustrates the effects of unperceived stimuli on subsequent behaviour, although the puzzle in such cases, superficially at least, lies just as much in the subjects' initial failure to perceive the stimuli as in their subsequent effects (Robertson and Marshall, 1993). Following brain lesions, most commonly those affecting the

right inferior parietal lobe, subjects may fail to attend to stimuli in contralateral space. The failure affects imagined scenes as well as real ones (Bisiach and Luzzatti, 1978). The syndrome has been fractionated into several sub-types: it may be primarily perceptual or primarily motor (Tegner and Levander, 1991), spatially or object-based, and may affect near-space or far-space (Halligan and Marshall, 1991). Yet despite the apparent failure of awareness of stimuli in the affected field among subjects with neglect, there is clear evidence for implicit processing of information about these stimuli. Thus subjects with left hemi-neglect, invited to express a preference for one of two line drawings which differ only in the plume of smoke rising from a house fire on the far left, tend to choose the smoke-free home (Marshall and Halligan, 1988). In a similar vein, unidentified words presented on the neglected left hand side of space can influence the identification of words presented later on the attended side (Berti and Rizzolatti, 1992).

The findings described in this section support the view that unacknowledged and presumably unperceived stimuli can exert detectable effects on neural activity and subsequent behaviour (Covey, 1997; Frith *et al.*, 1999). What is the explanation for these subliminal processes, and why do they remain subliminal? Why, for example, is blindsight blind?

Possible explanations for these phenomena fall into two broad classes: sub-threshold stimulation of systems which sometimes subserve consciousness, and activation of systems which invariably operate in the absence of consciousness. Thus, stimuli which are too faint or brief for conscious perception may nonetheless be capable of eliciting neural activity and biasing responses via pathways in which stronger or more sustained stimuli would elicit conscious awareness; or, alternatively, after destruction of brain regions which are supposedly required for awareness, stimuli may excite other structures, perhaps subcortical ones, which can drive behaviour but always operate unconsciously.

Different explanations may apply to different instances. Functional imaging opens up the possibility of examining this question in man. This line of research is relatively new, but several results have been reported, offering tentative support for both kinds of explanation. Thus, Sahraie and colleagues compared brain activity generated by stimuli which give rise to awareness with activity generated by stimuli permitting similar levels of discrimination without awareness, in a single subject (G.Y.) with blindsight (Sahraie *et al.*, 1997). They found that the shift between 'aware' mode and 'unaware' modes was associated with a shift in the pattern of activity from cortical to subcortical, and from dorsolateral to medial prefrontal. Psychophysical data also suggest that blindsight is unlike normal, near-threshold vision, and involves a distinctive form of visual processing (Azzopardi and Covey, 1997). In contrast, Zeki and ffytche, in work performed with the same patient, report enhanced activation of area V5 in the aware mode, with additional activation of a region of the brainstem (Zeki and ffytche, 1998). In a similar vein, Dehaene and colleagues have

provided evidence that in a task requiring subjects to classify numbers as larger or smaller than five by pressing a button with the left or right hand, presentation of masked, unreported, numerical primes sets in train a stream of perceptual, cognitive and motor processes in precisely the areas which are also engaged by the perceived stimulus (Dehaene *et al.*, 1998). Morris and colleagues, comparing the neural responses to masked and unmasked presentations of an angry face previously paired with an aversive stimulus, found that masked unreported presentations activated the right amygdala, while unmasked reported presentations activated the left amygdala (Morris *et al.*, 1998). The exploration of the neural correlates of implicit processes, along these lines, will complement the pursuit of the neural correlates of consciousness over the coming decade.

### ***(c) Insights from the study of memory and action***

Perceptual awareness supplies our most vivid conscious experiences, but there are many other kinds of conscious content. These range from the quasi-perceptual, like 'flash-bulb memories', to the mainly propositional, like a philosophical belief. Two areas of study, in particular, promise further insights into consciousness. Work on memory has developed a distinction between 'declarative' and 'procedural' memory which parallels the contrast between explicit and implicit perception. Research on action has developed in two directions which are especially relevant to consciousness: the first concerns the neural basis of the distinction between the attentive 'conscious' acquisition of a new skill and its practised or automatic performance; the second concerns disorders which disturb our sense of the ownership of our intentions. I shall review each area briefly.

It is well established that the ability to acquire new episodic memories, i.e. conscious recollections of our own experiences, depends upon structures in the medial temporal lobes and diencephalon, linked in the circuit of Papez (Squire *et al.*, 1990). Damage to these structures is the cause of the classical 'amnesic syndrome' which gives rise to a profound anterograde amnesia. But the amnesic syndrome spares the ability to acquire memories of other kinds, i.e. procedural memories which are demonstrated by their effects on our actions. The spared varieties of learning include classical conditioning (the process by which a conditioned stimulus, like a bell, paired with an unconditioned stimulus, like a morsel of food, becomes able to elicit the response, salivation, normally elicited by the unconditioned stimulus), priming (the process by which prior exposure to an item can facilitate subsequent behaviour in relation to the item, even in the absence of conscious perception or recollection of it) and acquisition of motor skills. This has the surprising result that patients who are wholly amnesic about their prior exposure to the tests in question can improve their performance in parallel with healthy controls.

There is a growing body of evidence about the neural basis of procedural learning. Classical conditioning is intact in patients with the amnesic syndrome; research in animals suggests that both acquisition and retention of classically conditioned responses are dependent on the cerebellum (Clark and Squire, 1998). Functional imaging studies of priming reveal that the facilitation of response is associated with a stimulus specific reduction in local brain activity, for example, a reduction in the extrastriate cortex on repeated presentation of pictorial stimuli (Buckner and Koutstaal, 1998). Acquiring new motor skills engages prefrontal regions linked with attention to action, discussed further below, the basal ganglia and the cerebellum (Berns *et al.*, 1997; Passingham, 1997).

Like perception and memory, our actions can be more or less 'conscious'. This aspect of consciousness has been approached by studying the difference between effortful acquisition of skills and their accomplished performance. This is a less direct approach to the neurology of consciousness than the study of implicit and explicit perception: we can easily revive our awareness of a skill, like driving, but nevertheless the tendency for our actions to recede from awareness as they are automatized offers one experimental avenue for examining the neural basis of awareness. Studies of skill acquisition suggest two generalizations. First, as skills are acquired, global brain activation declines (Haier *et al.*, 1992); secondly, the ensemble of areas engaged tends to change as automatic task performance takes over from effortful, attentive acquisition. Activity in prefrontal regions, in particular, falls away as behaviour becomes automatic (Passingham, 1997; Petersen *et al.*, 1998; Raichle, 1998). Thus, for example, Passingham and colleagues found that prefrontal cortex was activated by learning new motor sequences, but not by reproducing a learned sequence: the learned sequence activated more posterior premotor and supplementary motor areas (Passingham, 1997). Similarly, Raichle found that a task requiring subjects to generate related verbs in response to nouns initially activated left frontal and anterior cingulate cortices and the right cerebellar hemisphere; as responses to given nouns became automatic, activity shifted to the sylvian-insular cortices also implicated in word reading (Raichle, 1998).

The essence of willed or voluntary acts is that they have aims of which we are conscious and are, usually, willing to acknowledge. This sense of consciousness therefore throws a bridge between perception and action, the events we experience and the ones we bring about (Hurley, 1998). Disorders which disrupt our sense of ownership of our actions are a potential source of insight into the neurological basis of volitional or conscious action (Spence and Frith, 1999): these include the 'alien limb syndrome' (Feinberg *et al.*, 1992; Della Sala *et al.*, 1994), the delusions of alien control sometimes seen in schizophrenia (Spence and Frith, 1999), and the varieties of 'involuntary' action and inaction which occur in motor hysteria and under hypnosis (Halligan and David, 1999). Where it is possible to correlate these

syndromes with underlying neuropathology, the evidence points, once again, towards the importance of prefrontal cortex in framing and 'owning' conscious intentions.

Despite the many differences between the domains of perception, memory and action, there are potentially informative parallels between the operation of conscious and unconscious processes in each of these three contexts. Circumstances in which perceptual experience changes without any change in external stimuli (with or without an effect on behaviour) offer a direct avenue to the study of consciousness, while circumstances in which changing stimuli impinge on the nervous system without altering awareness make it possible to study implicit perception, the 'dark side' of visual awareness. In the context of memory, studies of declarative memory will illuminate our understanding of conscious recollection, while procedural memory offers a parallel to implicit perception. In the context of action, attentive performance supplies an analogous contrast with automatic behaviour. These parallels have recently been emphasized by Frith and colleagues in a lucid discussion of the experimental framework of research on consciousness (Frith *et al.*, 1999). The results of work in these three domains furnish much of the raw material for the overarching theories discussed in the following section.<sup>4</sup>

## V. Theories of consciousness

The renaissance of empirical research on consciousness has stimulated several more general accounts of its mechanisms. Some have tried to specify the neurological mechanisms of consciousness, often with the aim of showing that these parallel and can explain the features of experience; others have focused on the computational tasks which conscious processes might perform, usually with reference to putative functions of consciousness in the control of behaviour; a third group of theories have addressed the possible social origins and roles of consciousness. These approaches are not mutually exclusive: conscious visual experience, for example, has subjective qualities, a neural basis, a behavioural role and a social context. In this section I shall review a selection of the more prominent proposals in each of these groups. I will reserve the question of whether any such theory is capable of giving a complete account of consciousness for Section VI.

### (a) Neurobiological theories

The majority of neurobiological theories of consciousness take for granted two broad principles which have emerged from the past century of research, reviewed in Sections III and IV: that structures in the upper brainstem core play a critical role in arousal, and that thalamic and cortical activity

<sup>4</sup>Cowey (1997) and Weiskrantz (1997) show that examples of preserved implicit functioning in the presence of damaged explicit processes can be found throughout the domain of neuropsychology: this paper focuses on the evidence for implicit perception and implicit memory.

supplies much of the content of consciousness. Most of these theories also work from the premise that the neural correlate of consciousness (NCC) will prove to be, to use the term championed by Donald Hebb, a 'neuronal cell-assembly' of some kind. [Hebb defined the cell-assembly as 'a diffuse structure comprising cells in the cortex and diencephalon . . . capable of acting briefly as a closed system, delivering facilitation to other such systems' (Hebb, 1949).]<sup>5</sup>

Agreement on the role of neuronal assemblies in the genesis of consciousness leaves scope for disagreement about many critical details: how large must an assembly be to give rise to consciousness? Need it incorporate particular neuronal types or specific cortical layers? Need it involve given cortical regions, or possess a particular range of connections with regions elsewhere? Must an assembly engage in any particular pattern or duration of activity to generate awareness? Is a certain degree of complexity of interactions within the assembly a prerequisite for consciousness?

Edelman and Tononi have proposed a model which abstracts from the subjective properties of experience, from biological features of the thalamocortical networks which plausibly supply its contents, and from detailed computer simulations of their activity (Edelman, 1992; Tononi and Edelman, 1998*a*). A recent formulation has three key tenets: that consciousness arises from 'the fast integration of a large amount of information within a dynamic core of strongly interacting elements'; that re-entry, via reciprocal interconnections between regions of the thalamocortical system, mediates this rapid integration; and that the emergence of 'primary consciousness', the construction of our multimodal perceptual world, depends upon the integration of current sensory processing with previously acquired affect-laden memories. Tononi argues that this model of a constantly shifting 'dynamic core' of neural elements subserving consciousness accounts for many of its properties: its continuity and changefulness, its selectivity, the existence of a focus of attention and a more diffuse surround, its coherence, its pace of change and the wide 'access' of its contents to other psychological operations. In terms of the variables mentioned in the last paragraph, this theory plays down the role of particular neuronal types and cortical regions, but stresses the importance of the complex integration of thalamocortical subsystems which are both functionally segregated and highly interactive.

Crick and Koch (Crick, 1994; Koch, 1998) have made a series of proposals along broadly similar lines to these, with some instructive differences. They argue that, in the case of visual awareness, the NCC must be an 'explicit, multi-level, symbolic interpretation of part of the visual scene'. 'Explicitness' implies that the NCC must somehow reference those features of the visual scene of which we are currently

aware, for example, by a synchronized elevation of the firing rate of the cells which represent those features; the NCC for vision is likely to be 'multi-level' in the sense that several levels of processing in the hierarchy of cortical visual areas are involved; it is 'symbolic' in the sense that there is a close correlation between features of the visual scene and the neural activity which represents them. Crick and Koch anticipate that the NCC at any given time will involve a sparse but spatially distributed network of neurones, and that its activity must stand out above the background of neuronal firing for at least 100–200 ms. They suggest that the neurones directly involved in the NCC may have 'some unique combination of molecular, pharmacologic, biophysical and anatomic properties': for example Crick has speculated that 'bursty' pyramidal cells in layer 5 of the cortical visual areas may play a critical role in the NCC. With the aim of honing the definition of the NCC for vision, Crick and Koch have recently made the controversial proposal that neurones within area V1, primary visual cortex, do not directly participate in the NCC for visual awareness, despite supplying much of the information which is processed in visual areas downstream (Crick and Koch, 1995). This idea has two sources: the empirical observation that several characteristics of our visual experience correlate more closely with the behaviour of neurones in higher visual areas, such as area V4, than in V1; and the theoretical view that only cortical regions which can directly influence action, via interconnections with the frontal lobes, can contribute to consciousness.

Three other suggestions offer variations on these themes, illustrating the wide divergence of views on the level of organization which is relevant to understanding consciousness. Zeki has drawn attention to a subtle temporal asynchrony in the perception of different visual characteristics, for example, colour and motion, which he attributes to discrepancies between the processing times in the associated cortical visual areas (Zeki and Bartels, 1998). On the basis of these data he has proposed the existence of 'multiple visual micro-consciousnesses'. This suggestion envisages the emergence of consciousness of simple kinds from relatively local cortical activity. Milner and Goodale have marshalled evidence that the 'dorsal' stream of visual processing is dedicated to the on-line control of visuomotor behaviour, in the absence of visual awareness, which depends instead on the ventral stream, dedicated to object perception, recognition and the creation of our visual world (Milner and Goodale, 1995). This suggestion envisages a special role in awareness, or the lack of such a role, for substantial swathes of cortex (Milner, 1995). Finally, in a theory inspired by the phenomenon of blindsight, Weiskrantz has underlined the crucial importance for visual consciousness of 'the ability to render a parallel acknowledged commentary' on our visuomotor behaviour (Weiskrantz, 1997). This ability, rather than the capacity to make visual discriminations *per se*, is the casualty in patients with blindsight.<sup>6</sup> Weiskrantz distinguishes

<sup>5</sup>Although this premise is open to question, for the purposes of this review, I will accept the widespread assumption that conventional neuronal activity and synaptic interactions in neural networks supply the substrate of consciousness. In this section I will therefore not discuss further those proposals which explain consciousness in terms of quantum processes or the putative consciousness of single cells.

<sup>6</sup>The ability to render a commentary is not the only casualty: the capacity to make visual discriminations, while intact, is markedly impaired.

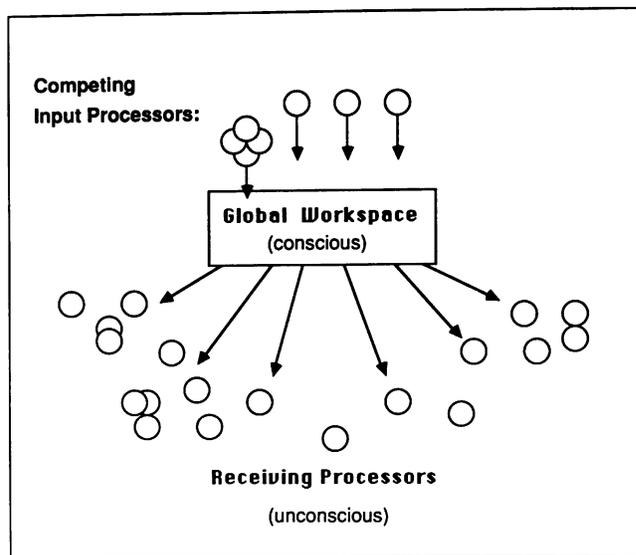
two views of this ‘commentary stage’: the view that it merely enables the acknowledgement of awareness, achieved by other means; and the view, which he favours, that the commentary ‘is actually endowing: it is what is meant by being aware and what gives rise to it’. This idea is reminiscent of Crick and Koch’s view that consciousness presupposes the possibility of action, and, like Crick and Koch, Weiskrantz speculates that the commentary stage depends on the ‘fronto-limbic complex’. This theory, therefore, envisages that visual consciousness requires the engagement of much of the brain, including both areas with primarily sensory functions and areas linked with action.

Several of these theories agree that a certain ‘kind’ or pattern of distributed neuronal activity may be crucial for consciousness; that, in other words, awareness has a physiology as well as an anatomy. The most popular current physiological candidate for a key role in consciousness is neural activity synchronized in the gamma frequency range, at ~35–45 Hz. This proposal originated in the theoretical suggestion that selective synchronization of neuronal firing might provide a basis for figure-ground segmentation in vision (von der Malsburg, 1995). More recently, evidence has accumulated that neurones responding to a single object synchronize their activity both within and between visual areas, and that synchrony increases as a stimulus moves into the focus of attention (Engel *et al.*, 2000). We have seen that widespread synchronized activity at these frequencies characterizes the magnetoencephalogram in the waking state and REM sleep (Llinas and Ribary, 1993); resonance of the auditory evoked potential in the same range has been used successfully as a marker of awareness in studies of the measurement of depth of anaesthesia (Munzlani *et al.*, 1993). Engel and Singer have proposed that synchrony may play a role in all the major processes subserving visual consciousness: arousal; sensory segmentation; selective attention; working memory; and also in the processes of ‘higher order consciousness’: motivation; action planning; and symbolic processing (Engel *et al.*, 2000). Proving the case for synchronization in the human brain will be technically demanding, but the evidence already available suggests that selective synchronization of neuronal activity may well be required for consciousness.

These theories reverberate between the twin poles of neural activity and experience, even if some of them, particularly Crick and Koch’s and Weiskrantz’s, look beyond experience, to action. But there is another broad theoretical approach to understanding consciousness, in terms of its functions. What difference does consciousness make to our lives? The most detailed attempts to answer this question have used the framework of information processing and cognitive psychology.

### (b) Information processing theories

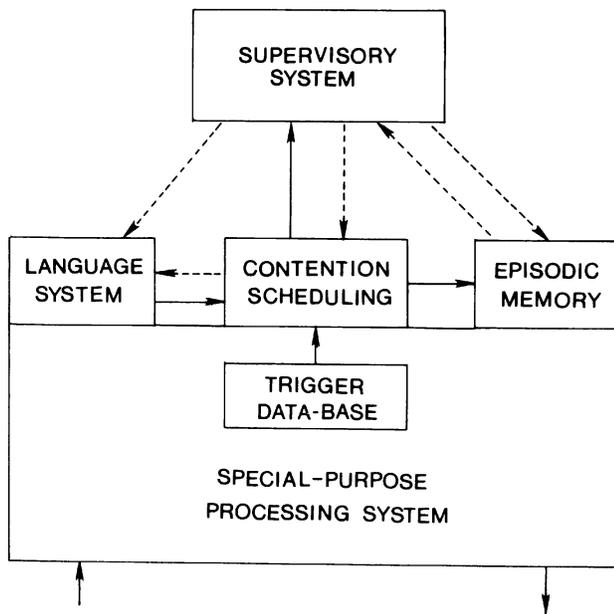
If consciousness has a biological function at all, it must ultimately be manifest in behaviour. Much of what we do



**Fig. 9.** Baars’ ‘global workspace’ model of consciousness (from Baars, 1998; with permission from Cambridge University Press, UK). The global workspace provides a ‘central information exchange’ which receives its input from the input processors which currently command attention, and broadcasts it widely through the brain.

can be done with little or no conscious attention: riding a bicycle and driving a car make few demands on consciousness under normal conditions once we have mastered the skills required. By contrast, circumstances which mobilize conscious attention tend to involve novel challenges or unpredictable events to which we need to devote a substantial part of our psychological resources. These uncontroversial observations are the cornerstone of most functional theories of consciousness.

These theories tend to follow the lead of William James in associating consciousness with selective attention and ‘primary’ or ‘working’ memory (Velmans, 2000). Thus, Baars has proposed a model which identifies the contents of consciousness with the contents of a ‘global workspace’ which can be broadcast widely through the nervous system, to recruit the operation of the numerous unconscious specialized subsystems to the task in hand (Baars, 1988; Baars and McGovern, 1996) (Fig. 9). Use of the global workspace trades automaticity, speed and high capacity parallel processing among unconscious modules of psychological function for flexibility of response under relatively slow, voluntary, limited capacity, serial control. One could hypothesize that the deficit in patients with a damaged ‘explicit’ but preserved ‘implicit’ function is precisely the loss of the ability to broadcast the output of the function through psychological space: as a result the residual capacity is unavailable for use in thinking or in imagery (Weiskrantz, 1997). Information processing theories of this kind harmonize with the widely held view that consciousness is linked to the emergence of flexible patterns of learned behaviour from more rigid instinctive patterns of response, as the ‘synaptic



**Fig. 10** Shallice's information-processing model of consciousness (from Shallice, 1988; with permission—copyright 1988, Oxford University Press, UK). The four subsystems extensively interconnected by arrows are those involved in awareness: we are conscious only of information which is successively represented in two or more of these. Solid arrows represent obligatory information transfer; dashed arrows optional transfer, under the control of the supervisory system.

bridge' that links sensation to action gradually lengthens in the course of cerebral evolution (Mesulam, 1998).

Shallice has proposed an alternative model which also emphasizes the role of consciousness in integrating the activities of psychological sub-systems (Shallice, 1988). He identifies a number of higher level systems which influence, rather than control, lower level processes. The higher level systems are concerned with language, episodic memory, the selection of actions given the organism's current overall goals ('contention scheduling'), and goal setting (in the 'supervisory system'). Shallice envisages that consciousness arises only when the same information is simultaneously represented in two or several of these higher level systems (Fig. 10). The contents of consciousness correspond to the information flowing between these systems. This model allows for the possibility that, in certain circumstances, it may be difficult or impossible to state unequivocally whether information is or is not conscious, when the integration of the controlling systems is impaired or incomplete.

Theories of this kind identify consciousness with a module, stage or aspect of information processing, on the assumption that consciousness plays an important, more or less continuous role in directing our waking behaviour. This assumption has been questioned, on conceptual and empirical grounds (Velmans, 2000). The conceptual doubt, which applies equally to all scientific theories of consciousness, is that no theory of this kind satisfactorily explains what phenomenal awareness contributes to information processing: we will return to this in

the following section. The empirical doubt arises from evidence that consciousness arises too late to perform the kinds of function which these theories envisage. Thus, for example, Velmans has argued that awareness of the meaning of a sentence arises too late for consciousness to inform the analysis of input on which comprehension depends (Velmans, 2000); similarly, Libet's work on the awareness of action suggests that the awareness of an intention to act is the result rather than the cause of preconscious processes which finally issue in action (Libet, 1996). One recurring response to this evidence draws on the idea that, rather than guiding action from moment to moment, consciousness serves its biological purpose on a longer, more reflective, time scale: on these theories the function of consciousness is to supply us with insight into the minds of others by illuminating the workings of our own (Humphrey, 1978; Frith and Frith, 1999). They suggest that understanding consciousness requires us to look beyond the brain and behaviour of isolated subjects to their social functioning.

### (c) Social theories

Several basic intuitions underlie social theories of consciousness. The first echoes its etymology: as we have seen (Section II) the Latin root of 'consciousness' refers, among other things to knowledge shared with another. Thus, our language itself hints at the social dimension of consciousness. A second, related, intuition is that the sharing of knowledge with oneself, in consciousness, and the sharing of knowledge with others, in social exchanges, are interdependent. This thought can be fleshed out in several ways: it can be argued, at a conceptual level, that the concept of one's own mind presupposes the concept of other minds (Strawson, 1974) and the thought receives empirical support from work in child development showing that awareness of self and awareness of others are acquired in parallel (Parker *et al.*, 1995). A third source of social theories of awareness is the idea that language is vital to human consciousness—and language is self-evidently a social creation. Proponents of these theories of consciousness sometimes claim that its social dimension accounts for the bafflement we tend to feel when we try to explain how the brain can generate experience: experience, on these views, is as much a social construction as a physiological or a psychological phenomenon (Rose, 1998; Singer, 1998).

Humphrey provided a lucid example of theories which propose a social function for awareness (Humphrey, 1978). He suggested that the purpose of consciousness is to allow social animals to model each other's behaviour on the basis of their insight into its psychological motivation. In other words, our knowledge of our own mental states supplies us with insight into the mental states underlying the actions of others; the ability to predict these is a major determinant of our biological success. More recently, such knowledge has been described in terms of the possession of a 'theory of mind' (Frith and Frith, 1999); some social theories broadly equate this with consciousness.

There is no doubt that a comprehensive theory of consciousness will need to take account of its social dimension. But most contemporary views agree that this is the wrong level of explanation for the simpler forms of consciousness. There is a logical worry about these social theories, if they are taken to provide a complete explanation of awareness: they seem to imply that introspection simultaneously provides access to and creates the mental states in question. The consensus view in the current literature is that these theories offer a helpful analysis of ‘self-consciousness’, or ‘higher-order awareness’, rather than addressing the more basic problem of perceptual awareness. But they do remind us that all our thinking about consciousness takes place in a social context, conditioned by our language and its history. The following section will examine how contemporary philosophers have tried to sharpen our understanding of the problems with which consciousness confronts us and of the terms in which we frame them.

## VI. The philosophy of consciousness

Philosophy aims to clarify the nature and logical relationships of our basic concepts. Its main areas of concern, and arguably its main unsolved dilemmas, have been constant throughout its history. The key philosophical question posed by consciousness concerns its relationship to the neural processes which correlate with it. How do the events which register in our experience relate to those occurring in our brains? This ‘problem of consciousness’ is the modern formulation of the ancient ‘mind–body problem’ (Churchland, 1984). Philosophical analysis takes a step back from the accounts discussed in the previous section, to ask what kind of theory has the potential, in principle, to offer a solution, or, indeed, to redraft the question (O’Regan and Noe, 2001).

There is no consensus. Rather than summarize the 15 or more radically differing philosophical approaches on offer (Chalmers, 1996), I shall map the positions of the main contenders in relation to three key intuitions which are regularly on view in philosophical discourse on consciousness.

### (a) *Three intuitions about consciousness*

The first intuition is that consciousness is a robust phenomenon which deserves to be explained rather than being explained away (Chalmers, 1996; Libet, 1996; Velmans, 2000). Sensory experiences like those of colour, sound or pain, the simplest and most vivid instances of consciousness, are phenomena which any full description of the world must reckon with. Indeed, experiences of this kind are arguably our point of departure in gaining knowledge of the world. Consciousness, in this sense, is the ‘sea in which we swim’ (Velmans, 2000).

The second intuition is that consciousness is bound up with our physical being. This thought is pre-scientific: everyone

knows that fatigue, alcohol, knocks on the head and countless other physical events can modify the state and contents of consciousness. But science has fleshed out the thought: Sections III and IV of this review summarize some of the observations which suggest that consciousness is rooted in the brain, and that the structure of consciousness is mirrored by the structure of a set of neural processes. It has become reasonable to suppose that every distinction drawn in experience will be reflected in distinctive patterns of neural activity.

The third intuition is that consciousness makes a difference (Fodor, 1989). It seems self-evident that much of our behaviour is explained by mental events; if we could not see or hear or touch, if we could not experience pain or pleasure, if we lacked conscious desires and intentions, we would not and could not behave as we do. If this is true it is natural to suggest that consciousness is a biological capacity which evolved in the service of action.

Much more could be said about each of these these intuitions and, of course, the fact that they are ‘natural’ does not guarantee that they are true. But they help to identify the main points of disagreement between the contending theories of consciousness in the philosophy of mind. I shall focus on three of these: the view that neural events are identical with the corresponding conscious events; the view that conscious events are best understood as functions of neural events; and the view that conscious and neural events are closely correlated but fundamentally distinct classes of phenomena.

### (b) *Identity theory*

The suggestion that conscious events are identical with the corresponding neural events offers a reductionist and materialist, or physicalist, solution to the mind–body problem. There is nothing particularly modern about such a view; it was taken by Lucretius in the ancient world and Thomas Hobbes in the 17th century. But a number of recent examples of successful scientific reduction have reinvigorated materialist theories of mind. Often cited examples include the ‘reduction’ of heat to the kinetic energy of atoms, the explanation of light in terms of electromagnetism and, perhaps of clearer relevance to consciousness, the analysis of ‘life’ as a set of physicochemical properties possessed by certain complex systems which are able to utilize energy from their surroundings to reproduce themselves. Why should consciousness be an exception to the stream of successful reductions of phenomena once considered to be beyond the reach of science?

There are some arguments suggesting that it might be an exception. Current physical theory teaches that light, in its physical existence, is identical to, is nothing more than, a certain type of electromagnetic radiation. To know everything about such radiation would be to know everything about light. But it is less clear that if we could know everything about the physicochemical properties and the behaviour of an organism we would know everything about its experience. For example, how far can scientific investigation take us

towards an appreciation of the subjective experience of an animal equipped with a sense we lack, like the echolocatory sense of bats and dolphins (Nagel, 1979)? Or, to come closer to home, could a blind student of the visual system ever gain the knowledge, which the sighted naturally possess, of 'what it is like to see' (Jackson, 1982)? This line of questioning suggests that conscious experience has subjective properties which are not fully specified by and can not be reduced to the neural structures and processes on which they depend. On this view, the analogy between the reduction of conscious processes to neural processes and the reduction of 'light' to electromagnetism fails, because in the latter case we legitimately set aside the properties of experience (e.g. what light looks like to us), while in the former the properties of experience are precisely the matter at issue.

In terms of our three intuitions, mind–brain identity theories, with their claim that conscious events 'simply are' brain events, do justice to the physical basis of experience and allow for its functional role. But they fail to satisfy the first intuition, that the properties of experience are robust phenomena in need of explanation. In John Searle's uncompromising words: '...the deeper objection [to physicalism] can be put quite simply: the theory has left out the mind' (Searle, 1992).

### (c) *Functionalism*

A second much canvassed view, particularly associated with Daniel Dennett, is that the essence of consciousness lies in the functions that it serves, i.e. in a certain subset of the transformations of input into output which our nervous systems achieve (Dennett, 1991). This theory of consciousness, functionalism, owes much to the developing science of artificial intelligence. Dennett develops an analogy between the activity of the brain and the implementation of a software package in a computer to create a 'virtual machine': 'human consciousness ... can best be understood as the operation of a ... virtual machine ... in the parallel architecture of a brain'. Taking vision as an example, functionalism suggests that visual experience consists in the countless acts of discrimination and classification which sight permits, and in their consequences for the rest of our mental life: functionalism reinterprets our experience in terms of a series of acts of judgement. O'Regan and Noe have developed a challenging account of vision along these lines, conceptualizing seeing as 'a way of acting' rather than as the creation of a visual representation (O'Regan and Noe, 2001).

This approach has many attractions. Like physicalism, it finds a place for consciousness in the natural world. It accounts for, indeed it originates with, our conviction that consciousness has effects. It escapes the superficiality of its intellectual predecessor, behaviourism, by taking seriously what goes on within our heads. It allows for the occurrence of consciousness in other organisms, or machines, which perform the same intellectual computations as we do.

But functionalism is arguably vulnerable to the same

fundamental objection as identity theory: that it fails to account for the qualitative properties of consciousness. Once again a third-person description, this time in terms of function, fails to specify the features of experience as we know them in the first person. We can ask of a 'virtual machine', just as we can of a neural assembly, why should it be conscious at all, and if it is conscious, why should its consciousness be like this (Chalmers, 1996)? Critics of functionalism argue that, like physicalism, it fails to bridge this explanatory gap, to answer the 'hard question' of consciousness. This is not to say that either physical or functional considerations are irrelevant to consciousness, merely that they fall short of giving a comprehensive philosophical account. Functionalists and physicalists deny this, claiming that once we fully understand the structure and function of the systems which support consciousness there will be nothing left to explain.

### (d) *Dualism*

A third view of the relationship between conscious and neural events is that they are closely correlated, but fundamentally distinct, classes of phenomena: theories of this kind are sophisticated intellectual descendants of Descartes' dualistic separation of material and spiritual substances (Descartes, 1641 in Descartes, 1968). Chalmers' 'naturalistic property dualism' is a much discussed recent example (Chalmers, 1996).

Chalmers argues that no third-person account, framed in terms of the structure and function of the nervous system, can bridge the explanatory gap which separates conscious from neural events. Third-person accounts can, however, explain the physical correlate of consciousness. Chalmers terms this physical state 'awareness', I will refer to this technical use of the word as awareness\*. He defines awareness\* as 'the state wherein information is accessible for verbal report and deliberate control of behaviour', a definition reminiscent of the information-processing theories of consciousness we encountered in Section V. This distinction between awareness and awareness\* has similarities with Bloch's distinction between phenomenal and access consciousness (Young and Bloch, 1996).

In Chalmers' terms a computer, which could report and act on information from its surroundings, would thereby be aware\*, but it might or might not be conscious. Explaining consciousness requires a further step: the generation of a set of psychophysical laws which describe the relationship between conscious events and neural ones, between awareness\* and consciousness. Chalmers subscribes to one interim principle which we have already touched on, which he terms 'the principle of structural coherence'—that any distinction in experience will be mirrored by a distinction in neural activity, and the pattern of experience will be matched by the pattern of awareness\*.

Why is this theory described as 'naturalistic property dualism'? Unlike Descartes, a 'supernatural substance dualist', Chalmers does not postulate that there is a mental substance, or that the mental domain is supernatural. Instead,

he regards the mental features of reality as properties of natural physical entities, animals, and he believes that their mental properties bear a lawful relation to their physical ones. In terms of our three intuitions, Chalmers' theory 'takes consciousness seriously', in accepting that conscious mental events are an irreducible feature of reality; he does justice to the second intuition by assuming that these events are causally dependent on their physical or functional base. But in his insistence that consciousness involves a separate class of non-physical properties, Chalmers falls foul of our third intuition, that the subjective qualities of experience make a difference to our lives: how could non-physical properties of conscious events make a difference to the physical trajectory of human behaviour? Chalmers accepts the unpalatable conclusion that they do not. This is the main stumbling block for theories which hold that conscious and neural events are correlated but distinct. In their laudable effort to do justice both to consciousness and to science, these views tend to be driven to the conclusion that experience is a beautiful but functionally irrelevant embellishment of physical processes.

A fourth, final, view of the nature of consciousness deserves a mention. It is inspired by the thought that some kinds of event might be inherently neural and mental, not linked by any contingent correlation, but of necessity, perhaps a necessity the human mind is not equipped to comprehend (Nagel, 1986; McGinn, 1991). If the crucial events were simultaneously physical and conscious, all three intuitions could be saved: that conscious experience is irreducible, intimately tied to the physical world and efficacious. But at present, as Nagel has commented, we have no clear understanding of what kind of property could render physical events intrinsically mental, and speculations of this kind are prone to head off in the direction of panpsychism, the view that all physical things have a conscious 'inner' aspect, which many find implausible.<sup>7</sup> Nevertheless, a number of commentators believe that some version of this 'neutral monism' (Velmans, 2000) or 'dual-aspect' theory holds out the greatest promise of an eventual solution to the philosophical conundrum of consciousness.

Against the background of this ancient debate, what are the chances that science will give us a comprehensive analysis of consciousness? Our estimate of these odds must depend on our view of the nature of consciousness. If we believe that conscious events can be redescribed without residue in terms of neural events or of their roles in governing behaviour, as physicalism and functionalism, respectively suggest, then science will give a full account of them. If not, the best we can expect is an ever richer science of correlation, filling out our understanding of the neural mechanisms of consciousness. Would this leave consciousness a 'mystery'? Perhaps no more and no less so than the fact of existence itself, and for the same reason: our awareness is the point of departure for our knowledge of the universe and of ourselves, 'the sea in which we swim'. Future work in philosophy and science

may change the ways in which we think about consciousness so radically that its reduction to physical process begins to look plausible. For the time being we have no alternative but to continue to use all three vocabularies of biology, behaviour and experience in our efforts to understand the mind.

## VII. Conclusion

Describing the process by which mechanistic models replaced animistic ones to become the standard approach to biological explanation, the historian of medicine, Charles Singer, wrote: 'The course of physiological advance may be described, briefly, as the expulsion of the mental element from process after process associated with vital activity' (Singer, 1928). As our mental lives are a crucial aspect of our biology, the process of expulsion eventually had to stop. The current fascination with consciousness reflects the mounting intellectual pressure to explain how 'vital activity' in the brain generates a 'mental element', with rich subjective content.

There has been impressive scientific progress towards this goal over the past century, transforming our biological understanding of both conscious states and their contents. Landmarks include the discovery of the electrical correlates of wakefulness, slow wave and paradoxical sleep; the exploration of the structures which control these states, in the brainstem and diencephalon; and the gradual definition of the processes of neural analysis and synthesis which underlie perceptual awareness. A number of new paradigms are beginning to allow the focused investigation of the neurology of experience. In particular, the neural correlates of changes in perceptual experience occurring in the absence of stimulus change should help to reveal the neural events which are most closely tied to awareness; the definition of the wide range of implicit neural processes, from blindsight to habit, will highlight the neurology of consciousness by contrast. The advent of functional imaging combined with sophisticated electroencephalographic techniques is transforming our ability to study these processes directly in man. While this work is on course to illuminate the biological basis of the forms of consciousness we share with other animals, parallel investigation of the 'social brain' (Brothers, 1990) will clarify the neurology of the more peculiarly human capacity for awareness of self, and of others.

Whether scientific approaches will provide a completely satisfying explanation of awareness is a vexed philosophical issue. But the determination of recent work in neuroscience, psychology and philosophy to do justice both to the rich texture of experience and to its intimate, revealable relationship to neuronal events holds out great promise for the study of consciousness in years to come.<sup>8</sup>

<sup>8</sup>This review should close with an apology. Reviews are nowadays quite rightly expected to be systematic surveys of the literature bearing on a subject. A systematic review of the theory and evidence bearing on consciousness would be more than a lifetime's work. I have not attempted it. This review is a selective, partial and undoubtedly biased reading of the subject. Its only justification is as an introduction, which may encourage some of its readers to investigate the details, where God dwells. I apologize to the scientists and philosophers whom I have failed to cite through lack of space or ignorance of their work.

<sup>7</sup>Many, but not all. See, for example, Chalmers (1996, pp. 293–301).

## Acknowledgements

I am very grateful to Rustam Al-Shahi, Rebecca Aylward, David Chalmers, Alan Cowey, Anthony Grayling, John Gruzelier, Tony Hope, Sally Laird, David Mitchell, Kevin O'Regan, Jon Stone, James Tickell, Larry Weiskrantz and Sophie Petit-Zeman for their comments. I thank Roger Cull and Sharon Hoy for help in preparing Fig. 2, and Stephanie Sharp and Deidre Davidson for their secretarial help.

## References

- Alkire MT. A unitary physiologic theory for the mechanism of anesthetic-induced loss of consciousness. In: *toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 179.
- Andrews K, Murphy L, Munday R, Littlewood C. Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *BMJ* 1996; 313: 13–16.
- Aserinsky E, Kleitman N. Two types of ocular motility occurring during sleep. *J Appl Physiol* 1955; 8: 1–10.
- Azzopardi P, Cowey A. Is blindsight like normal near-threshold vision? *Proc Natl Acad Sci USA* 1997; 94: 14190–4.
- Baars BJ. *A cognitive theory of consciousness*. Cambridge: Cambridge University Press; 1988.
- Baars BJ, McGovern K. Cognitive views of consciousness. In: *Velmans M, editor. The science of consciousness*. London: Routledge; 1996. p. 63–95.
- Barbur JL, Watson JD, Frackowiak RS, Zeki S. Conscious visual perception without V1. *Brain* 1993; 116: 1293–302.
- Baron-Cohen S. *Mindblindness*. Cambridge (MA): MIT Press; 1995.
- Bauer RM. Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia* 1984; 22: 457–69.
- Beninger RJ, Kendall SB, Vanderwolf CH. The ability of rats to discriminate their own behaviour. *Can J Psychol* 1974; 28: 79–91.
- Berger H. *Über das Elektroencephalogramm des Menschen*. *Arch Psychiat* 1929; 87: 527–70.
- Berns GS, Cohen JD, Mintun MA. Brain regions responsive to novelty in the absence of awareness. *Science* 1997; 276: 1272–5.
- Berti A, Rizzolatti G. Visual processing without awareness: evidence from unilateral neglect. *J Cogn Neurosci* 1992; 4: 347–51.
- Bisiach M, Luzzatti C. Unilateral neglect of representational space. *Cortex* 1978; 14: 129–33.
- Brazier MAB. *A history of the electrical activity of the brain: the first half-century*. London: Pitman; 1961.
- Bremer F. Cerveau 'isole' et physiologie du sommeil. *C R Seanc Soc Biol* 1929; 102: 1235–41.
- Brothers L. The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. *Concepts Neurosci* 1990; 1: 27–51.
- Buchsbaum MS, Gillin JC, Wu J, Hazlett E, Sicotte N, Dupont RM, et al. Regional cerebral glucose metabolic rate in human sleep assessed by positron emission tomography. *Life Sci*, 1989; 45: 1349–56.
- Buckner RL, Koutstaal W. Functional neuroimaging studies of encoding, priming, and explicit memory retrieval. [Review]. *Proc Natl Acad Sci USA* 1998; 95: 891–8.
- Chalmers DJ. *The conscious mind*. Oxford: Oxford University Press; 1996.
- Childs NL, Mercer WN, Childs HW. Accuracy of diagnosis of persistent vegetative state. *Neurology* 1993; 43: 1465–7.
- Chokroverty S. An overview of sleep. In: *Chokroverty S, editor. Sleep disorders medicine*. Boston: Butterworth Heinemann; 1999.
- Churchland PM. *Matter and consciousness*. Cambridge (MA): MIT Press; 1984.
- Clark RE, Squire LR. Classical conditioning and brain systems: the role of awareness. *Science* 1998; 280: 77–81.
- Cohen MS, Kosslyn SM, Breiter HC, DiGirolamo GJ, Thompson WL, Anderson AK, et al. Changes in cortical activity during mental rotation. *Brain* 1996; 119: 89–100.
- Cowey A. Cortical visual areas and the neurobiology of higher visual processes. In: *Farah MJ, Ratcliff G, editors. The neuropsychology of high-level vision*. Hillsdale (NJ): Lawrence Erlbaum; 1994. p. 3–31.
- Cowey A. Current awareness: spotlight on consciousness. [Review]. *Dev Med Child Neurol* 1997; 39: 54–62.
- Cowey A, Stoerig P. Blindsight in monkeys. *Nature* 1995; 373: 247–9.
- Cowey A, Walsh V. Magnetically induced phosphenes in sighted, blind and blindsighted observers. *Neuroreport* 2000; 11: 3269–73.
- Crick F. *The astonishing hypothesis*. London: Simon & Schuster; 1994.
- Crick F, Koch C. Are we aware of neural activity in primary visual cortex? *Nature* 1995; 375: 121–3.
- Dehaene S, Naccache L, Le Clec'H G, Koechlin E, Mueller M, Dehaene-Lambertz G, et al. Imaging unconscious semantic priming. *Nature* 1998; 395: 595–600.
- Della Sala S, Marchetti C, Spinnler H. The anarchic hand: a fronto-mesial sign. In: *Boller F, Grafman J, editors. Handbook of Neuropsychology*, Vol. 9. Amsterdam: Elsevier; 1994. p. 233–55.
- Dement W, Kleitman N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalogr Clin Neurophysiol* 1957; 9: 673–90.
- Dennett DC. Quining qualia. In: *Marcel AJ, Bisiach E, editors. Consciousness in contemporary science*. Oxford: Clarendon Press; 1988. p. 42–77.
- Dennett DC. *Consciousness explained*. Boston (MA): Little, Brown; 1991.
- Descartes R. *Discourse on method and the meditations*. Harmondsworth (UK); Penguin; 1968.
- Edelman GM. *Bright air, brilliant fire*. London: Penguin Books; 1992.

- Engel AK, Fries P, Roelfsema PR, Konig P, Singer W. Temporal binding, binocular rivalry, and consciousness. 2000. Available from: <http://www.phil.vt.edu/ASSC/engel/engel.html>
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998; 392: 598–601.
- Feinberg TE, Schindler RJ, Flanagan NG, Haber LD. Two alien hand syndromes. [Review]. *Neurology* 1992; 42: 19–24.
- ffytche DH, Howard RJ, Brammer MJ, David A, Woodruff P, Williams S. The anatomy of conscious vision: an fMRI study of visual hallucinations. *Nat Neurosci* 1998; 1: 738–42.
- Fiset P, Paus T, Daloze T, Plourde G, Meuret P, Bonhomme V, et al. Brain mechanisms of propofol-induced loss of consciousness in humans: a positron emission tomographic study. *J Neurosci* 1999; 19: 5506–13.
- Flanagan OJ. *Consciousness reconsidered*. Cambridge (MA): MIT Press; 1992.
- Fodor JA. Making mind matter more. *Philos Top* 1989; XVII: 59–79.
- Frith CD, Frith U. Interacting minds—a biological basis. [Review]. *Science* 1999; 286: 1692–5.
- Frith C, Perry R, Lumer E. The neural correlates of conscious experience: an experimental framework. *Trends Cogn Sci* 1999; 3: 105–14.
- Gallop GG Jr. Chimpanzees: self-recognition. *Science* 1970; 167: 86–7.
- Greenfield S. How might the brain generate consciousness? In: Rose S, editor. *From brains to consciousness*. London: Allen Lane; 1998. p. 210–27.
- Griffiths TD. Musical hallucinosis in acquired deafness. Phenomenology and brain substrate. *Brain* 2000; 123: 2065–76.
- Haier RJ, Siegel BV Jr, MacLachlan A, Soderling E, Lottenberg S, Buchsbaum MS. Regional glucose metabolic changes after learning a complex visuospatial/motor task: a positron emission tomographic study. *Brain Res* 1992; 570: 134–43.
- Halligan PW, David AS. *Conversion hysteria: towards a cognitive neuropsychological account*. Hove (UK): Psychology Press; 1999.
- Halligan PW, Marshall JC. Left neglect for near but not far space in man. *Nature* 1991; 350: 498–500.
- Hebb DO. *The organization of behavior*. New York: John Wiley; 1949.
- Heiss W-D, Pawlik G, Herholz K, Wagner R, Wienhard K. Regional cerebral glucose metabolism in man during wakefulness, sleep, and dreaming. *Brain Res* 1985; 327: 362–6.
- Hofle N, Paus T, Reutens D, Fiset P, Gotman J, Evans AC, et al. Regional cerebral blood flow changes as a function of delta and spindle activity during slow wave sleep in humans. *J Neurosci* 1997; 17: 4800–48.
- Holmes G, Lister WT. Disturbances of vision from cerebral lesions, with special reference to the cortical representation of the macula. *Brain* 1916; 39: 34–73.
- Hubel DH, Wiesel TN. The Ferrier Lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci* 1977; 198: 1–59.
- Humphrey N. Nature's psychologists. *New Scientist* 1978; 78: 900–3.
- Hurlburt RT. *Denizens of the phenom*. In: *Toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 345.
- Hurley SL. *Consciousness in action*. Cambridge (MA): Harvard University Press; 1998.
- Jackson F. Epiphenomenal qualia. *Philos Q* 1982; 32: 127–36.
- Jacoby LL, Toth JP, Lindsay DS, Debner JA. Lectures for a layperson: methods for revealing unconscious processes. In: Bornstein RF, Pittman TS, editors. *Perception without awareness*. New York: Guilford Press; 1992. p. 81–120.
- James W. *The principles of psychology*. New York: Henry Holt; 1890.
- Jasper HH, Descarries L, Castelucci VF, Rossignol S, editors. *Consciousness at the frontiers of neuroscience*. Philadelphia: Lippincott-Raven, 1998. p. 75–94.
- Jennett B, Plum F. Persistent vegetative state after brain damage. *Lancet*, 1972; I: 734–7.
- John ER. Changes with loss and return of consciousness common across anaesthetics. In: *Toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 174.
- Jones JG. Perception and memory during general anaesthesia. [Review]. *Br J Anaesth* 1994; 73: 31–7.
- Jones BE. The neural basis of consciousness across the sleep-waking cycle. In: Jasper HH, Descarries L, Castelucci VF, Rossignol S, editors. *Consciousness: at the frontiers of neuroscience*. *Advances in neurology*, Vol. 77. Philadelphia: Lippincott-Raven; 1998. p. 75–94.
- Kanwisher N. Neural correlates of changes in perceptual awareness in the absence of changes in the stimulus. In: *Toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 164.
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 1997; 17: 4302–11.
- Kentridge RW, Heywood CA, Weiskrantz L. Attention without awareness in blindsight. *Proc R Soc Lond B Biol Sci* 1999; 266: 1805–11.
- Kihlstrom JF, Barnhardt TM, Tartaryn DJ. Implicit perception. In: Bornstein RF, Pittman TS, editors. *Perception without awareness*. New York: Guilford Press; 1992. p. 17–54.
- Kilduff TS, Kushida CA. Circadian regulation of sleep. In: Chokroverty S, editor. *Sleep disorders medicine*. 2nd ed. Boston: Butterworth Heinemann; 1999. p. 135–47.
- Kinomura S, Larsson J, Gulyas B, Roland PE. Activation by attention of the human reticular formation and thalamic intralaminar nuclei. *Science* 1996; 271: 512–15.
- Koch C. The neuroanatomy of visual consciousness. In: Jasper HH, Descarries L, Castelucci VF, Rossignol S, editors. *Consciousness:*

- at the frontiers of neuroscience. *Advances in neurology*, Vol. 77. Philadelphia: Lippincott-Raven; 1998. p. 229–43.
- Kosslyn SM, Shin LM. Visual mental images in the brain: current issues. In: Farah MJ, Ratcliff G, editors. *The neuropsychology of high-level vision*. Hillsdale (NJ): Lawrence Erlbaum; 1994. p. 269–96.
- Kosslyn SM, Thompson WL, Kim IJ, Alpert NM. Topographical representations of mental images in primary visual cortex. *Nature* 1995; 378: 496–8.
- Kutas M, Dale A. Electrical and magnetic readings of mental functions. In: Rugg MD, editor. *Cognitive neuroscience*. Hove (UK): Psychology Press; 1997. p. 197–242.
- LaBerge D. *Attentional processing*. Cambridge (MA): Harvard University Press; 1995.
- Leopold DA, Logothetis NK. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 1996; 379: 549–53.
- Lewis CS. *Studies in words*. Cambridge: Cambridge University Press; 1960.
- Libet B. Neural processes in the production of conscious experience. In: Velmans M, editor. *The science of consciousness*. London: Routledge; 1996. p. 96–117.
- Livingstone M, Hubel D. Segregation of form, color, movement and depth: anatomy, physiology, and perception. [Review]. *Science* 1988; 240: 740–9.
- Llinas R, Ribary U. Coherent 40-Hz oscillation characterizes dream state in humans. *Proc Natl Acad Sci USA* 1993; 90: 2078–81.
- Logothetis NK, Schall JD. Neuronal correlates of subjective visual perception. *Science* 1989; 245: 761–3.
- Lumer ED, Friston KJ, Rees G. Neural correlates of perceptual rivalry in the human brain. *Science* 1998; 280: 1930–4.
- Macquet P, Peters JM, Aerts J, Delfiore G, Degueldre C, Luxen A, et al. Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature* 1996; 383: 163–6.
- Macquet P, Degueldre C, Delfiore G, Aerts J, Peters JM, Luxen A, et al. Functional neuroanatomy of human slow wave sleep. *J Neurosci* 1997; 17: 2807–12.
- Mahowald MW, Schenck CH. Dissociated states of wakefulness and sleep. *Neurology* 1992; 42 Suppl 6: 44–52.
- Manford M, Andermann F. Complex visual hallucinations: clinical and neurobiological insights. [Review]. *Brain* 1998; 121: 1819–40.
- Marrocco RT, Witte EA, Davidson MC. Arousal systems. [Review]. *Curr Opin Neurobiol* 1994; 4: 166–70.
- Marshall JC, Halligan PW. Blindsight and insight in visuo-spatial neglect. *Nature* 1988; 336: 766–7.
- McCarley RW. Sleep neurophysiology: basic mechanisms underlying control of wakefulness and sleep. In: Chokroverty S, editor. *Sleep disorders medicine*. Boston: Butterworth Heinemann; 1999. p. 21–50.
- McGinn C. *The problem of consciousness*. Oxford: Basil Blackwell; 1991.
- Merikle PM, Reingold EM. Measuring unconscious perceptual processes. In: Bornstein RF, Pittman TS, editors. *Perception without awareness*. New York: Guilford Press; 1992. p. 55–80.
- Mesulam MM. From sensation to cognition. [Review]. *Brain* 1998; 121: 1013–52.
- Metzinger T, editor. *Conscious experience*. Schoningh: Imprint Academic; 1995.
- Milner AD. Cerebral correlates of visual awareness. [Review]. *Neuropsychologia* 1995; 33: 1117–30.
- Milner AD, Goodale MA. *The visual brain in action*. Oxford: Oxford University Press; 1995.
- Milner AD, Rugg MD. *The neuropsychology of consciousness*. London: Academic Press; 1992.
- Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science* 1985; 229: 782–4.
- Morris JS, Ohman A, Dolan RJ. Conscious and unconscious emotional learning in the human amygdala. *Nature* 1998; 393: 467–70.
- Moruzzi G, Magoun HW. Brain stem reticular formation and the activation of the EEG. *Electroencephalogr Clin Neurophysiol* 1949; 1: 455–73.
- Multi-Society Task Force on PVS. Medical aspects of the persistent vegetative state (1). *N Engl J Med* 1994; 330: 1499–508.
- Multi-Society Task Force on PVS. Medical aspects of the persistent vegetative state (2). *N Engl J Med* 1994; 330: 1572–9.
- Munzlani R, Andrade J, Sapsford DJ, Baddeley A, Jones JG. A measure of consciousness and memory during isoflurane administration: the coherent frequency. *Br J Anaesth* 1993; 71: 633–41.
- Nagel T. What is it like to be a bat? In: Nagel T. *Mortal questions*. Cambridge: Cambridge University Press; 1979. p. 165–80.
- Nagel T. *The view from nowhere*. New York: Oxford University Press; 1986.
- Natsoulas T. Concepts of consciousness. *J Mind Behav* 1983; 4: 13–59.
- O'Regan JK. Is visual consciousness a grand illusion? In: *Toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 221.
- O'Regan JK, Noe A. A sensorimotor account of vision and visual consciousness. *Behav Brain Sci*. In press 2001.
- Pallis C, Harley DH. *ABC of brainstem death*. 2nd ed. London: BMJ Publishing Group; 1996.
- Parker ST, Mitchell RW, Boccia ML, editors. *Self-awareness in animals and humans*. Cambridge: Cambridge University Press; 1994.
- Passingham R. Functional organisation of the motor system. In: Frackowiak RSJ, Friston KJ, Frith CD, Dolan RJ, Mazziotta JC, editors. *Human brain function*. San Diego: Academic Press; 1997. p. 243–74.
- Paus T, Zatorre RJ, Hofle N, Caramanos Z, Gotman J, Petrides M, et al. Time-related changes in neural systems underlying attention

- and arousal during the performance of an auditory vigilance task. *J Cogn Neurosci* 1997; 9: 392–408.
- Penrose R. *Shadows of the mind*. Oxford: Oxford University Press; 1994.
- Petersen SE, van Mier H, Fiez JA, Raichle MA. The effects of practice on the functional anatomy of task performance. *Proc Natl Acad Sci USA* 1998; 95: 853–60.
- Plum F. Coma and related global disturbances of the human conscious state. In: Peters A, Jones EG, editors. *Cerebral cortex*. New York: Plenum Press; 1991. p. 359–425
- Plum F, Posner JB. *The diagnosis of stupor and coma*. 3rd ed. Philadelphia: F.A. Davis; 1982.
- Poppel E, Held R, Frost D. Residual visual function after brain wounds involving the central visual pathways in man. *Nature* 1973; 243: 295–6.
- Raichle ME. The neural correlates of consciousness: an analysis of cognitive skill learning. [Review]. *Phil Trans R Soc Lond B Biol Sci* 1998; 353: 1889–901.
- Rees G, Wojciulik E, Clarke K, Husain M, Frith C, Driver J. Unconscious activation of visual cortex in the damaged right hemisphere of a parietal patient with extinction. *Brain* 2000; 123: 1624–33.
- Robertson IH, Marshall JC, editors. *Unilateral neglect: clinical and experimental studies*. Hove (UK): Lawrence Erlbaum; 1993.
- Robbins TW, Everitt BJ. Arousal systems and attention. In: Gazzaniga MS, editor. *The cognitive neurosciences*. Cambridge (MA): MIT Press; 1995. p. 703–20.
- Rose S, editor. *From brains to consciousness?* Princeton (NJ): Princeton University Press; 1998.
- Rosenthal DM. Two concepts of consciousness. *Philos Studies* 1986; 49: 329–59.
- Sahraie A, Weiskrantz L, Barbur JL, Simmons A, Williams SC, Brammer MJ. Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proc Natl Acad Sci USA* 1997; 94: 9406–11.
- Sanders MD, Warrington EK, Marshall J, Weiskrantz L. 'Blindsight': vision in a field defect. *Lancet* 1974; 1: 707–8.
- Schiff ND, Plum F. The neurology of impaired consciousness: global disorders and implied models. 2000. Available from: <http://www.phil.vt.edu/assc/niko.html>
- Schwender D, Kaiser A, Klasing S, Peter K, Poppel E. Midlatency auditory evoked potentials and explicit and implicit memory in patients undergoing cardiac surgery. *Anesthesiology* 1994; 80: 493–501.
- Searle JR. *The rediscovery of the mind*. Cambridge (MA): MIT Press; 1992.
- Shallice T. Information-processing models of consciousness. In: Marcel AJ, Bisiach E, editors. *Consciousness in contemporary science*. Oxford: Clarendon Press; 1988. p. 305–33.
- Shepard RN. The mental image. *Am Psychol*; 1978; 33: 125–37.
- Sherin JE, Shiromani PJ, McCarley RW, Saper CB. Activation of ventrolateral preoptic neurons during sleep. *Science* 1996; 271: 216–19.
- Shewmon DA. Chronic 'brain death': meta-analysis and conceptual consequences. *Neurology* 1998; 51: 1538–45.
- Singer C. *A short history of medicine*. Oxford: Oxford University Press; 1928.
- Singer W. Consciousness from a neurobiological perspective. In: Rose S, editor. *From brains to consciousness?* London: Allen Lane; 1998. p. 228–45.
- Sobel N, Prabhakaran V, Hartley CA, Desmond JE, Glover GH, Sullivan EV, et al. Blind smell: brain activation induced by an undetected air-borne chemical. *Brain* 1999; 122: 209–17.
- Spence SA, Frith CD. Towards a functional anatomy of volition. *J Consciousness Studies* 1999; 6: 11–29.
- Squire LR, Zola-Morgan S, Cave CB, Haist F, Musen G, Suzuki WA. Memory: organization of brain systems and cognition. [Review]. *Cold Spring Harb Symp Quant Biol* 1990; 55: 1007–23.
- Steriade M. Neurophysiologic mechanisms of non-rapid eye movement (resting) sleep. In: Chokroverty S, editor. *Sleep disorders medicine*. 2nd ed. Boston: Butterworth Heinemann; 1999.
- Steriade M, Gloor P, Llinas RR, Lopes de Silva FH, Mesulam MM. Basic mechanisms of cerebral rhythmic activities. [Review]. *Electroencephalogr Clin Neurophysiol* 1990; 76: 481–508.
- Steriade M, McCormick DA, Sejnowski TJ. Thalamocortical oscillations in the sleeping and aroused brain. [Review]. *Science* 1993; 262: 679–85.
- Stoerig P, Cowey A. Blindsight in man and monkey. [Review]. *Brain* 1997; 120: 535–59.
- Strawson PF. Self, mind and body. In: Strawson PF. *Freedom and resentment*. London: Methuen; 1974. p. 169–77.
- Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. *Lancet* 1974; 2: 81–4.
- Tegner R, Levander M. Through a looking glass. A new technique to demonstrate directional hypokinesia in unilateral neglect. *Brain* 1991; 114: 1943–51.
- Thines G. Phenomenology. In: Gregory RL, editor. *The Oxford companion to the mind*. Oxford: Oxford University Press; 1987. p. 614–16.
- Tononi G, Edelman GM. Consciousness and the integration of information in the brain. In: Jasper HH, Descarries L, Castelucci VF, Rossignol S, editors. *Consciousness: at the frontiers of neuroscience*. *Avances in neurology*, Vol. 77. Philadelphia: Lippincott-Raven; 1998a. p. 245–80.
- Tononi G, Edelman GM. Consciousness and complexity. [Review]. *Science* 1998b; 282: 1846–51.
- Treue S, Maunsell JH. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 1996; 382: 539–41.
- Varela FJ, Shear J, editors. *The view from within*. Thorverton: Imprint Academic; 1999.

- Velmans M, editor. *The science of consciousness*. London: Routledge; 1996.
- Velmans M. *Understanding consciousness*. London: Routledge; 2000.
- von der Malsburg C. Binding in models of perception and brain function. [Review]. *Curr Opin Neurobiol* 1995; 5: 520–6.
- Von Economo C. *Encephalitis lethargica: its sequelae and treatment*. London: Oxford University Press; 1931.
- Warlow CP, Dennis MS, Van Gijn J, Hawkey GJ, Sandercock, PAG, Bamford JM, Wardlow J. *Stroke: a practical guide to management*. Oxford. Blackwell Scientific; 1996.
- Weiskrantz L. Varieties of residual experience. *Q J Exp Psychol* 1980; 32: 365–86.
- Weiskrantz L. Disconnected awareness for detecting, processing and remembering in neurological patients. The Hughlings Jackson Lecture. *J Roy Soc Med* 1991; 84: 466–70.
- Weiskrantz L. *Consciousness lost and found*. Oxford: Oxford University Press; 1997.
- Weiskrantz I. *Blindsight: a case study and implications*. Oxford: Clarendon Press; 1998.
- Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike MA. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci* 1998; 18: 411–18.
- Young AW, Block N. Consciousness. In: Bruce V, editor. *Unsolved mysteries of the mind*. Hove (UK): Erlbaum; 1996. p. 149–79.
- Zajonc RB. Feeling and thinking: preferences need no inferences. *Am Psychol*; 1980; 35: 151–75.
- Zeki S. A century of cerebral achromatopsia. [Review]. *Brain* 1990; 113: 1721–77.
- Zeki S. Cerebral akinetopsia (visual motion blindness). A review. [Review]. *Brain* 1991; 114: 811–24.
- Zeki S. *A vision of the brain*. Oxford: Blackwell Scientific; 1993.
- Zeki S, Bartels A. The asynchrony of consciousness. *Proc R Soc Lond B Biol Sci* 1998; 265: 1583–5.
- Zeki S, ffytche DH. The Riddoch syndrome: insights into the neurobiology of conscious vision. [Review]. *Brain* 1998; 121: 25–45.
- Zeki S, Watson JD, Frackowiak RS. Going beyond the information given: the relation of illusory visual motion to brain activity. *Proc R Soc Lond B Biol Sci* 1993; 252: 215–22.
- Zeman A. Persistent vegetative state. [Review]. *Lancet* 1997; 350: 795–9.
- Zeman A. The problem of unreportable consciousness. In: *Toward a science of consciousness*. Thorverton: Imprint Academic; 2000. Abstr. No. 96.
- Zeman AZ, Grayling AC, Cowey A. Contemporary theories of consciousness [editorial]. *J Neurol Neurosurg Psychiatry* 1997; 62: 549–52.
- Zoltoski RK, Cabeza RJ, Gillin JC. Biochemical pharmacology of sleep. In: Chokroverty S, editor. *Sleep disorders medicine*. 2nd ed. Boston: Butterworth Heinemann; 1999. p. 63–94.

*Received December 7, 2000. Revised February 23, 2001.*

*Accepted March 14, 2001*