

# Book Reviews

---

**Christof Koch***The Quest for Consciousness:  
A Neurobiological Approach*

Englewood, CO: Roberts & Company, 2004, 448 pp.,  
\$45, ISBN 0974707708 (hbk).

---

Ten years ago, Francis Crick dedicated a popular book, *The Astonishing Hypothesis*, to Christof Koch. Now Koch has returned the favour in *The Quest for Consciousness*, which sums up a collaboration going back to the 1980s; hence in this review 'Koch' can be taken as shorthand for 'Crick and Koch' (as the author himself suggests). Their collective enterprise marries a sense of wonder — a 'quest' for the 'astonishing' — with a resolute focus on the microanatomy of the brain. The resulting blend might be called 'reductionism in Wonderland'.

It was Lewis Carroll's Alice who inspired the now-notorious quip addressed to the Crick's readers in 1994: 'You're nothing but a pack of neurons.' Though this is usually read (and often quoted) as the acme of brazen reductionism, those who persisted to the end of the book found a more cautious Crick: 'the words nothing but in our hypothesis can be misleading if understood in too naïve a way' (p. 261). Astonishing or not, the hypothesis was essentially a theoretical framework for an embryonic research program which he hoped might eventually contribute to solving the old mind-body problem. Nevertheless, Crick's final chapter declared eloquently that the more we understand the mechanics of the brain/mind, the more marvelous it becomes. 'To say that our behavior is based on a vast,

interacting assembly of neurons should not diminish our view of ourselves but enlarge it tremendously' (p. 260).

In his new book, Koch's expressions of wonder are a little more muted, and the hypothesis somewhat modified from the 1994 version, but the main outlines of the research program remain the same. The fruits it has borne are beautifully presented here in rich detail, lavishly illustrated and impeccably edited. Specialists will appreciate the wealth of empirical detail and documentation, but Koch's style — relatively informal, concise and straightforward, though perhaps less colorful than Crick's — should appeal to more general readers.

Like its predecessor, Koch's book focusses mainly on the visual system in the brain and its relation to the experience of seeing, on the assumption that all aspects and modalities of consciousness 'employ one or perhaps a few common mechanisms' (p. 15). The visual system is better known than others (so far) because it is more accessible to empirical research — mainly because the visual system in macaque monkeys is virtually identical to its counterpart in humans, and we can use laboratory techniques on monkey brains that can rarely be used on humans. The clarity with which Koch conveys this circumstance points to one significant virtue of his book: he provides a vivid sense of what neurobiologists actually do in their laboratories, and why.

Koch establishes his own motivation with an anecdote that happens to be non-visual. It all started in 1988 with a toothache: 'I pondered why it hurt,' he writes; though he already knew

quite a lot about the neuronal mechanisms generating the pain, ‘none of this explained why it felt like anything!’ (xv). At this point readers will naturally expect Koch’s *Quest for Consciousness* to deliver some kind of answer to the question of ‘qualia’, the ‘Hard Problem’ of Chalmers. Yet after 300 pages devoted mostly to describing neuronal mechanisms in exquisite detail, Koch admits that ‘why qualia feel the way they do remains an enigma’ (p. 310). Why? Not because the Problem is misconceived, but because we still need more data. This is the reductionist faith in a nutshell, and though Koch avoids dogmatic insistence on it, his ‘quest’ never deviates from it. He sums up the nature of this quest as follows:

I seek the physical basis of phenomenal states within brain cells, their arrangements and activities. My goal is to identify the specific nature of this activity, the neuronal correlates of consciousness, and to determine to what extent the NCC differ from activity that influences behavior without engaging consciousness (p. 19).

Koch’s framework thus fits within the global workspace family, along with other models that associate consciousness with a temporary ‘coalition’ of neuronal populations or circuits whose more or less synchronized activity persists long enough to command attention and guide action, until the next coalition takes over. His glossary (which is excellent and extensive) defines the NCC as the ‘minimal set of neuronal mechanisms or events jointly sufficient for a specific conscious percept or experience’. The distinctive features of Koch’s theory are most clearly indicated by the words *minimal*, *sufficient* and *specific*.

Koch’s minimalism is remarkable in that he considers it a form of emergence theory. ‘The working hypothesis of this book is that consciousness emerges from neuronal features of the brain’ (p. 10); and in the footnote which defines his terms, Koch says ‘A system has emergent properties if these are not possessed by its parts.’ Yet he does not consider it scientifically viable to treat consciousness as an emergent property of the whole brain, because its functioning includes all sorts of ‘activity that influences behavior without engaging consciousness.’ Therefore consciousness must

emerge from some part of the brain. Clearly this part (the NCC) does not have a simple fixed location, and membership in the coalition must be constantly changing from moment to moment, but the more narrowly we can specify the components of this shifting subsystem, the closer we are to understanding how consciousness emerges. What we need, then, is a ‘strategy for circumscribing the NCC’.

Accordingly, Koch provides a very detailed overview of cortical architecture to serve as the background against which the NCC can be defined. He also devotes considerable space to nonconscious brain functions, including a whole chapter on ‘the zombie within’. This refers not to the philosopher’s zombie (a notion which Koch finds ‘sterile’) but to ‘zombie agents, that carry out routine missions’ (p. 206). These ‘cortical reflexes’ must be eliminated in order to pin down the NCC. Koch also eliminates the ‘enabling factors’, such as biological mechanisms which maintain a general state of arousal. More controversially, he eliminates the primary visual cortex (V1): it is certainly necessary for the normal functioning of the visual system, but V1 activity is also observed in the absence of any conscious percept, so V1 is not sufficient for conscious seeing. The same goes for the retina; these parts of the visual system are ‘pre-NCC’. Further up in the visual hierarchy, we can also eliminate the dorsal stream, which enables ‘blindsight’ in patients whose ventral stream has been damaged: they do not consciously see anything but can manipulate objects as if they did.

The higher we go in the visual hierarchy, the more specific are the percepts or objects with which a neuron’s firing is correlated. Koch calls this specificity ‘logical depth of computation’ because it takes several layers of computation, including feedback loops, to derive it from retinal input. Taking this pattern to its linear conclusion would posit ‘grandmother neurons’ whose activity directly represents a consciously recognized object — an idea that often elicits ridicule from researchers. Koch, though, comes very close to embracing it when he reports the discovery of ‘Bill Clinton neurons’ (by Itzhak Fried, who recorded the firing of single neurons in an epileptic patient while viewing a series of 50 images). However, Koch explains that no single neuron could do the job

alone; it would take a coalition to bring Clinton's presence in a scene to consciousness.

Koch also flirts with embracing another oft-derided concept, the homunculus in the brain. It is essential to his theory that the neuronal coalition which supplies specific content to consciousness must project to the 'planning areas' in the frontal lobes. (This is another reason for eliminating V1 from the NCC, since it does not send axons to that region in the monkey brain.) So we are justified in using the metaphor of a 'homunculus residing in the front of the forebrain, receiving information from the sensory regions in the back, and relaying its output to the motor systems' (p. 302). The twist is that this is a nonconscious homunculus. Following Ray Jackendoff, Koch claims that 'neither the process of thought nor its content is knowable by consciousness' (p. 297). Thus he eliminates not only the lower levels of the 'computational mind' but the higher levels as well, leaving only the 'intermediate' to furnish the NCC. Also eliminated is the post-NCC, the 'penumbra', which Koch invokes (Chapter 14) to account for the meaning of qualia: the NCC of the moment will have outputs to other cells representing a myriad associated concepts, habits and so forth, which are lit up to some extent by the heightened activity in the NCC; some of these may participate in the next 'winning coalition', whereupon the associations become conscious.

Koch's brand of empiricism is not only minimalist but microscopic: he argues that the key to consciousness may lie in the differences between types of neurons (hundreds have been identified already) and/or in the fine structure of the cortex, its layers and columns and interconnections. His approach is also temporally microscopic, focussing on the short time span within which a coalition can arise and maintain itself. Thus Koch has little to say about diachronic matters such as the development of the organism or its brain. This would explain, for instance, why he rejects the 'enactive' approach which emphasizes the primacy of movement in generating consciousness. Koch dismisses this idea mainly because paralyzed people can still be conscious. The fact that the paralytic's NCC are functional components of an organism with an ontogenic history of movement (and an evolutionary heritage of motility) lies beyond the bounds of his theory, presum-

ably because this kind of fact is not observable in the laboratory.

Koch does recognize the need to place consciousness in a larger functional context (since he does not regard it as epiphenomenal), and deals with this in Chapter 14, but he is careful to mark these ideas as 'speculative'. He argues that consciousness provides the organism with an 'executive summary' of situations which demand attention because the 'zombie agents' cannot cope with them automatically. Koch does mention some other views on the nature and function of consciousness, and (near the end) some other neurobiological theories, but in each case he tells us just enough about the alternatives to explain why he has no use for them. For example, he acknowledges a kinship with Edelman and Tononi's 'dynamic core' hypothesis, only to comment that he finds it a bit too 'holistic'. This book offers 'only cartoon-like pocket sketches' (p. 4) of competing approaches to consciousness, but it certainly presents a full-color and high-resolution picture of brain structure at the cellular level. Readers oriented to higher levels may find Koch's definition of the NCC (and of consciousness) too narrow, yet still find much that is rewarding — even astonishing — in *The Quest for Consciousness*.

*Gary Fuhrman*

---

**Vincent Walsh & Alvarado Pascual-Leone**

*Transcranial Magnetic Stimulation: A Neurochronometrics of Mind*

Cambridge, MA: MIT Press, 2003, 297 pp., \$70. ISBN 0262232286 (hbk).

---

Written for research neurologists, this handsomely printed book introduces the subject of transcranial magnetic stimulation (TMS), which joins electroencephalography (EEG), event related potentials (ERPs), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET) as a new tool for studying the dynamics of the human brain. In simplest terms, TMS floods a restricted region of the neocortex with a large magnetic field (about 2 tesla or 20,000 gauss) for a fraction of a millisecond. In order to influence brain dynamics, it is now understood, a magnetic

field must be rather large (the earth's magnetic field is about 0.5 gauss), and it must be rapidly established. The broad effect of such a sudden intrusion of magnetic field energy is to introduce computational noise into the neocortical dynamics, interfering with motor activity and causing the perception of spots of light (phosphenes), in addition to more subtly influencing the brain's behavior in a variety of ways.

The basic idea of TMS is simple — a steady voltage source (power supply) charges a storage capacitor to some 2 kilojoules of energy, which is suddenly discharged as magnetic field energy through a magnetic stimulating coil by closing a solid-state switch. In an interesting early chapter, the authors of *Transcranial Magnetic Stimulation* trace the checkered history of related ideas from the discovery of magnetic induction by Michael Faraday in 1831 to the practical realization of TMS by Anthony Barker and his colleagues at Sheffield, England in the mid- 1980s. Why 'checkered'? Our brains are relatively insensitive to magnetic fields of ordinary strength (try waving a horseshoe magnet near your temple) so much of the nineteenth-century speculation on mental effects of magnetic fields was pseudoscientific or worse.

The brain is insensitive to magnetic fields because it is not an electromagnetic medium. In other words, cortical dynamics are dominated by interacting electric fields and ionic currents, in which magnetic fields play only minor roles. Normal cortical currents generate minute magnetic fields which are observed in MEG measurements, and the very large external magnetic fields that are suddenly turned on under TMS generate small transmembrane voltages that may influence the course of these currents, but this picture is very different from a true electromagnetic medium — such as a propagating radio wave or the output light beam from a laser — in which the total energy remains essentially conserved as it oscillates rapidly back and forth between electric and magnetic energy. In the brain, to the contrary, electric field energy is continually being generated by hydrolysis of adenosine triphosphate, stored in transmembrane potentials, and dissipated through myriad nonlinear electric field interactions with transmembrane ionic currents. Whatever small magnetic fields that are present do not influence normal functioning of the brain.

In neuroscience research, TMS has the advantage of acting as a noninvasive 'virtual lesion' which can be rapidly induced over a region that penetrates an inch or so into the neocortex and is transversely localized to roughly a square inch. Because cortical fibers are somewhat randomly oriented, such a localized and rapidly rising TMS field generates a random spectrum of transmembrane voltages that are either excitatory or inhibitory depending on the local orientation of a particular fiber, thereby introducing the functional equivalent of computational noise. About 20 ms after the end of a TMS pulse, the major effects disappear, and an experiment can be repeated; thus it is feasible to measure changes in timing delays for various motor responses as the stimulating coil is moved over the scalp, a research activity of considerable current interest. In addition to such timing experiments, which the authors of this book describe in some detail, there is also the possibility of using TMS in a "repetitive" mode called rTMS. In this mode, a periodic series of magnetic pulses are generated, and the experimenter has the opportunity of introducing steady noise into a restricted volume of the cortical dynamics for a well defined interval of time. This leads to the possibility of observing subsequent effects on a variety of subtle cognitive activities, many of which are reviewed and described in this book.

TMS safety is evidently a key concern; the authors mention that no one with a history of epilepsy should be used as a subject. Clearly this is a matter that neuropsychological researchers should consider carefully as they put their subjects through an increasing variety of subjective experiences, including visual suppression and extinction, search interference, geometrical perception, perception of temporal sequence, variations in attention, perceptual learning, and memory inhibition. Considering all of these possible applications, it seems safe to predict that both TMS and rTMS will make important contributions to research in neuropsychology over the next few years. As an introduction, *Transcranial Magnetic Stimulation* is highly recommended for all who would take up this exciting activity.

Alwyn Scott  
University of Arizona

---

**Pierre Jacob & Marc Jeannerod**
*Ways of Seeing: The Scope and Limits of Visual Cognition*

Oxford University Press, 2003, ISBN 0198509200 (hbk), 0198509219 (pbk)

---

In another fine book from the Oxford Cognitive Science Series, philosopher Pierre Jacob and well-known cognitive neuroscientist Marc Jeannerod collaborate to bring together many key findings in the visual ventral ('what') and dorsal ('where' and 'how') systems. They then go on to discuss the 'mirror neuron' system for perceiving and responding to object-oriented actions, a social perception system, and finally taxis-and-praxis systems for interacting with the environment.

Their major contribution is to clarify and integrate the mishmash of visual 'ventral' and 'dorsal' stream concepts, which the authors term 'dualistic models' of the visual processing of objects. Mishkin and Ungerleider in the early 1980s differentiated between two visual systems. The *ventral visuo-feature* ('what is it?') system is crucial for noticing and remembering an object's illumination, color, orientation and basic shape. The *dorsal visuo-spatial* ('where is it?') system discriminates and remembers an object's position in space in relation to other objects – translating visual information into 'allocentric' location coordinates.

In the early 1990s, however, Goodale and Milner contrasted the ventral, 'what' system with a somewhat different dorsal system, namely a *visuo-motor* ('how to respond to it') system. This 'how' system translates visual information into egocentric location coordinates – the position of each object in relation to ones body. This 'how', perception for action, system is much faster, evolved first and is more essential for basic survival, than the 'what' system. Goodale and Milner maintain that the 'what' system involves *conscious perception* but that their 'how' system operates *automatically*, outside of conscious awareness. Two striking illustrations of this are found in *visuo-spatial neglect patients* who still have intact 'how' systems: (a) most visuo-spatial neglect patients will consciously perceive *nothing* in their left visual space but can accurately reach for and grasp the objects they cannot consciously perceive; while (b) a few visuo-spatial neglect patients *misperceive* objects in their left

visual field, so that the objects appear smaller — despite this optical illusion, their grasping hand assumes the right size of the object rather than the consciously-perceived size.

Jacob's and Jeannerod's major contribution to this discussion is to say 'both/and' instead of 'either/or' to these two differing concepts of the dorsal stream by including both the Ungerleider 'where' and the Goodale 'how' systems into the stream, since both are centered in the parietal lobe. *Allocentric* coding seems to be centered in the *right hemisphere inferior parietal lobe*, while *egocentric* coding is *bilaterally* centered in still more *dorsal parietal areas*, in the intra-parietal sulcus. The authors also incorporate *attentional* systems into the dorsal – but without anatomical clues. While I will suggest some further complexities, Jacob and Jeannerod are very helpful in clarifying these visual systems.

Our authors make a second major contribution by pulling together various levels of motor responses. While subcortical motor systems activate many motor reflexes and unintended behaviors (such as absentmindedly rubbing ones eyes and automatic eye movements), both the 'taxis' and 'praxis' systems are intentional systems that produce 'actions', in contrast to unintended 'behavior'. The visuo-motor (how) system is an '*optic-taxis*' system. Patients with optic *ataxia* fail to reach or grasp objects. Frontal lobe *praxis* systems activate more skilled actions. Patients with *apraxia* can neither use tools properly nor pantomime their use.

J & J's third major contribution is pulling together the recently discovered 'mirror cell' evidence and various 'social system' findings in ways that show the strong wired-in linkage between perception-for-action and action. The perception of non-social intentional action by others involves, among other things, so-called 'mirror' cells in the ventral-premotor area of monkeys, discovered in the early 1990s. Mirror cells are sensori-motor cells that fire both when observing *others* make what seems to be an intentional action toward an object, and when one makes the *same* action toward that object. These pre-motor mirror cells have heavy two-way connections with the optic-taxis 'how' system. This circuit seems to transform object-specific information into motor plans for responding to those objects. The mirror-cell premotor area then sends infor-

mation to the primary motor area to activate the actual action, though this may sometimes be aborted by tonic inhibition from other sources.

Parallel to the non-social 'mirror system', a 'social perception' system processes emotionally-rich facial expressions, bodily posture and actions of others, and triggers appropriate emotional responses in oneself. Thus it detects other persons' assumed *social* motor intentions, triggers automatic *social/emotional* response plans in oneself, and so facilitates automatic intentional responses. The output from both the mirror and social perception systems serve as visual input to the mind-reading/Theory of Mind system, which our authors only touch on.

Jacob and Jeannerod clearly posit their position as a version of the 'dualistic approach' to human vision. The dualistic approaches they consider agree on the role of the ventral stream for object features, but then differ as to the second branch: either 'allocentric object location' circuits or 'egocentric object location' circuits. Our authors affirm all this and talk of the 'dual role of the parietal lobe' (p. xviii) for both allocentric and egocentric location. They also write vaguely of parietal *attentional* functions. I think this stretches the concept of 'dualistic approach'. I wonder why they do not join with some others in talking about three basic streams, or even talk of multiple streams. Recent anatomical studies of brain lesions in 'visuo-spatial neglect' (not detecting objects in ones left visual field), 'extinction' (not detecting them only if there is an object in ones right visual field) and ataxia (deficit in grasping or manipulating objects) suggest that visuo-spatial neglect implicates inferior-parietal allocentric coding areas, extinction implicates somewhat more dorsal parietal attentional areas, and ataxia implicates still more dorsal parietal egocentric coding areas. This suggests far more than a mere 'dualistic approach'. Complicating this further is the possibility that areas where damage can result in pure 'visuo-spatial neglect' may be crucial ones in which 'object features' and 'allocentric object location' (Ungerleider & Mishkin's ventral and dorsal streams) merge. If so, both of the *Ungerleider* streams are necessary for conscious perception, *but not* both of the *Goodale* streams!

Beyond the scope of this book is consideration of an additional syndrome called 'somatic

(body) neglect', where one will say that ones left arm, for example, 'does not belong to me', even if ones attention is drawn to it. This second type of 'neglect' involves destruction of the so-called 'body map', centered in right parietal *somatosensory* processing areas, rather than in the *visual* areas we have been talking about. Somehow the optic-taxis and praxis mechanisms draw upon the body map. The resultant dis-owning of one's arm (one form of anosognosia = lack of knowledge of one's disorder) clearly implicates these *somatic* areas in consciousness, even if the *visuo-motor* optic taxic circuits are not.

Despite these reservations, this is an excellent book!

Bill Faw

Brewton-Parker College

### **The Dalai Lama**

#### *Stages of Meditation*

New York: Snow Lion Publications, 2003, 210 pp., \$16.95, ISBN 1559391979 (pbk).

### **Alan Wallace**

#### *Choosing Reality: A Buddhist View of Physics and the Mind*

New York: Snow Lion Publications, 2003, 232 pp., \$16.95, ISBN 1559391995 (pbk).

### **Michele Martin**

#### *Music in the Sky: The Life, Art and Teachings of the 17<sup>th</sup> Karmapa*

New York: Snow Lion Publications, 2003, 352 pp., \$18.95, £12.95, ISBN 1559391952 (pbk).

'Every two years in Tucson we have another conference called "*Toward a Science of Consciousness*",' exclaimed Alan Wallace, physicist and Buddhist, in tones of kindly mockery, at a recent conference on meditation. '*Toward, toward ... any day now!*' These three books illustrate Wallace's proposition that a fully developed science of consciousness already exists, elaborated hundreds of years ago in the East, which the West continues largely to ignore.

Even those who would not accept that proposition are generally impressed and often moved by the loving wisdom of the Dalai Lama. His book is an elegant commentary given as a living teaching in 1989 at Manali in northern India on the intermediate part of the *Stages of Medi-*

*tation* by Kamalashila, a great scholar saint of the ninth century. It contains the complete Tibetan text of this classic, a glossary and a bibliography. The team of translators, to judge by the easy readability of the English text, have done excellent work.

The first chapter is entitled, 'What Is the Mind?' A form of answer is given, here and subsequently, though it emerges more readily from the Dalai Lama's commentary than in the passages from the root text that punctuate every chapter. The answer is difficult to summarise, as Kant's *Critiques* are difficult to summarise, with which indeed the exposition shares similarities. Two types of phenomena exist: the permanent and the impermanent. The impermanent are subject to, and limited by, causes and conditions. The permanent are not. The mind, like all the objects that it perceives, is impermanent. Its chief limitation derives from the erroneous belief that both the objects of perception and itself are permanent. When it recognises the error, it becomes less limited, though still impermanent.

And so the exposition proceeds. The following chapter describes techniques for ridding the mind of its cardinal error through training. Such training helps to eliminate unhappiness. The Dalai Lama continues:

Having considered the importance of training the mind, we might wonder what the mind is. If you ask them, most people respond by rubbing their heads and then pointing to their brains. This is partly correct. . . . The human mind does not have any existence independent of the human body. . . . The human mind, or consciousness, . . . actually consists of a vast number of minds, some subtle and some coarse. . . . The main cause of any mind is the preceding moment of consciousness, whose nature is clarity and awareness (p. 38).

From these beginnings the Dalai Lama goes on to weave together into a rich and satisfying pattern the themes of compassion, equanimity, suffering, wisdom, and 'the practice of calm abiding'. He emphasises the 'middle way' that utilises both feeling and intellect. Almost as an aside he introduces a set of criteria for deciding whether a scripture – or indeed any thesis – possesses authority. The overall impression is of

rigor exercised with discernment in the service of wisdom and love.

Nonetheless, almost all neuroscientists in professional mode and most Western philosophers, while endorsing some of these propositions and sentiments, would have little sympathy with the overall thrust and conclusions of this presentation, albeit the argument is lucid, rational and logical. The clash of views lies more in the premises and presuppositions of the parties. A key problem is that of limitation. Conventional scientists are seldom prepared to concede, at least publicly, that their worldview, like any other, is inevitably limited and as such subject to error. A re-examination and realignment of the underlying assumptions of our science, not least those concerning consciousness and its subordination or otherwise to matter, might indeed lead, as the Dalai Lama suggests, to an enlargement of knowledge and wisdom and to an increase in happiness.

Anyone who has sympathy with these ideas will enjoy, while those who do not might benefit from, Alan Wallace's own development of them in *Choosing Reality*. His writing is of course very much attuned to the Western mind, though the core ideas in it are essentially those given by the Dalai Lama in his commentary. It comprises a careful analysis of just those presuppositions and limitations of Western science mentioned above, together with the Buddhist experientially based response. As a physicist, he is penetrating concerning what he calls the 'assumptions of scientific realism', and knowledgeable about quantum reality and mathematical realism. As a Buddhist and experienced meditator he is able through personal experience to put into words, insofar as this is ever possible, that way of apprehending reality. Far from decrying the achievements and power of Western science, he, like the Dalai Lama, advocates a middle way, what he calls the 'centrist position'. This acknowledges the value of both paths, and sees them not as contradictory but as powerfully complementary. 'The great contemplative traditions . . . focus on the essential concerns of human existence, which are not addressed by physical science. The meaningfulness of scientific and contemplative knowledge is therefore complementary. In the absence of either, the world is impoverished.' Once again the beauty of a

well-structured and penetrating mind ruled by compassion shines through.

*Music in the Sky* is a remarkable volume containing biography, teachings, poetry and splendid colour photographs, not to mention extensive notes and a glossary. It concerns the life and times of the contemporary Buddhist sage and mystic Karmapa. He is regarded as a *tulka*, reincarnation, of a 13th century Tibetan monk; the seventeenth of such. At the age of 14, on 28 December 1999, Karmapa 'leapt from the roof of his monastery in Tibet and evading his Chinese guards undertook the gruelling and dangerous journey to India'. So he is now no more than 18 years old. The extensive material that Michelle Martin has assembled, almost overwhelming except perhaps for a disciple, is full of striking incidents and happenings, much of it not easy for a Western mind to accept as it stands. However, Karmapa's views on what it is to be human and on the nature of consciousness closely reflect those of Wallace and the Dalai Lama.

Beyond the value of the particular, contrasted yet similarly focussed, content of each of these three volumes lies a wider question. Why, at least for us in the West, is the 'hard problem' of consciousness so hard? Pragmatically, how should we tackle a problem that is so hard that many cannot see any way of making progress in solving it, other than denying or evading its existence? Deriving as they do, explicitly or implicitly, wholly or in part, from a different set of presuppositions to those of conventional Enlightenment science, these writings should provoke us to recognise that the only fruitful approach to the impasse is indeed to do that difficult and disagreeable thing: seek out, examine and question our own presuppositions, assumptions and axioms, something many good scientists are already doing. When we take that step, not only will our science break free from its current often unacknowledged limitations, but scientists themselves (and perhaps that means all of us) will become open to an enlargement of consciousness, and thereby as the Dalai Lama asserts, to greater happiness. Only then might a Tucson conference legitimately be entitled 'The Science of Consciousness'.

Julian Candy

---

**Paula Droege**

*Caging the Beast: A Theory of Sensory Consciousness*

Amsterdam/Philadelphia: John Benjamins, 2003, 183 pp., €55, ISBN 9027251827, \$65.95 ISBN 1588113914 (pbk), €107, ISBN 9027251819, \$89 ISBN 1588113906 (hbk).

---

Sensory states can be strange. *Caging the Beast* (henceforth CtB) shows that they can be even stranger than you may have supposed. The key to CtB is the notion that mentality is more fundamental than consciousness (p. 9). Mentality cuts across the distinction between a conscious and unconscious state because it senses *change* without any need for consciousness. When it does so, some representation or perception may arise that can be either conscious or unconscious.

What makes a perception conscious is the 'principle question' (p. 11) CtB investigates. It asks: 'what distinguishes conscious experiences from mental states (or, indeed, any states) that aren't conscious experiences?' Droege calls this the *question of state consciousness*. She focuses on conscious *sensory* states while putting the *question of qualitative character* aside. For, according to her, it is not its 'raw feel' that characterizes a conscious sensory state but the experience of 'now'. Answering her question involves identifying a conscious sensory state as the experience of being in the present moment, while excluding other areas of experience.

Higher-order theories are rejected because, in them, 'a mental state becomes conscious without undergoing any fundamentally internal change.' (p. 46) Yet first-order theorist Fred Dretske offers a crucial insight: either consciousness is being conscious of a mental *object* or it is the mental *act*, by virtue of which one is conscious of something in the world. The mental act obviously fits in with Droege's notion of mentality. Unfortunately, Dretske does not make a distinction between conscious and unconscious states, for in his view all perceptions are conscious, allowing higher-order theorists to object that he conflates state and transitive (perceptual) consciousness (p. 56). Droege softens this criticism by interpreting Dretske's thoughts about blindsighters as a case of perception without state consciousness: i.e 'seeing' does not necessarily mean 'consciously seeing' (pp. 57–8).

She goes on to identify three types of sensory state that a theory of sensory consciousness must explain (p. 65). The second (Type 2) is her own proposal:

- (1) being state conscious of x and noticing/responding to x (Type 1)
- (2) being state conscious of x without noticing x (Type 2)
- (3) not being state conscious of x but nonetheless able to respond to x (Type 3)

Then she recapitulates her argument that higher-order theories conflate Type 1 with Type 2, because state consciousness implies representational consciousness, though not the other way around. So only Types 1 and 3 are possible according to these theorists. Dretske's first-order theory, on the other hand, conflates Type 2 with Type 3, because he assumes that representational consciousness implies state consciousness. He recognizes Type 1 and, according to Droege's reinterpretation of his thoughts, Type 3. She concludes that both have trouble explaining Type 2, which is neither fully conscious nor unconscious; therefore she proposes a more satisfactory theory, able to account for all three cases, encapsulated in the notion of a *second sense*.

This *second sense* selects and coordinates sensory representations of the external world at the present moment, thus establishing what is the aim of the exercise — a conscious sensory state. The *degree of coordination* needed for a particular task at hand determines the perceived intensity of consciousness, obviously greater in Type 1 than in Type 2. The difference between Type 1 and Type 3 can only be explained by the representation of *the present moment*. Type 3 is selected and coordinated in some other way (p. 96) so cannot be the outcome of a second sense. Even though it occurs *in* the present moment, it is not *of* the present moment. Puzzlingly, the coordination hypothesis indicates that Type 3 is more conscious than Type 2. Probably due to the *after-effect* on the subject: 'not only does attentional control [second sense] determine the content of conscious sensory states, but the reverse is true as well' (p. 155), whereas pure state conscious is more like the awareness of background information (p. 118).

CtB thus gives a representationalist theory of sensory consciousness that is richly psychological. It sees consciousness on a continuous scale from weak to intense and taps for inspiration into the dynamic world of dreams, sleepwalkers and repressed memories. Unfortunately, though, the pieces of the puzzle that CtB describes do not fit together as well as one might like.

The first difficulty is to do with the ontological status of Type 2. The difference between Types 1 and 2 is one of degree, so why set them apart? Droege appeals to 'the intuitive sense... that our sensory states...are more richly detailed than we are able to process' (p. 97). So Type 2 is a construct introducing *state consciousness without perception*. Likewise, Type 2 seems to dissolve into Type 3. For example a 'Type 3' entity might possess a survival kit full of unconscious intentions such as nipple sucking, face recognition or 'just being'. But Droege claims that 'now' is a 'part of the content' of this entity that causes a 'unified sense of how things are at the moment' (p. 77), represented by 'some kind of temporal ordering principle' (p. 76). Therefore she argues that it is really a Type 2 entity. However, could something abstract as 'now' make a difference in the absence of any subjective 'uptake' or unconscious intention? Well, it looks as though one has to choose between Type 1 (state conscious of 'now' and responding to 'now') or Type 3 (the perception of 'now' based on some unconscious a priori sensitivity). The concept of Type 2 seems redundant because *consciously* there is no 'between-ness' as Droege calls it, only *mental* in between-ness. To call a state conscious that has not been consciously experienced is surely to take the concept a step too far.

Second, one may wonder whether it is consciousness or the states themselves that are continuous: e.g. wakeful states, sleep states, good memories or weak ones, which, by virtue of a discontinuous mental act that senses change, lead to representations and in some cases to consciousness.

I agree with Droege that there must be a *mental act* to make consciousness possible. The merit of CtB is that it refocuses consideration of questions of consciousness onto foundational issues — the senses. A further implication is that the secret of consciousness

may not lie where one expects to find it, in the rich repository of mature conscious states, but in how the mind and senses parse the world into separate ‘chunks’ of experience.

*Justus de Swart*

---

**James H. Fetzer, ed.**

*Consciousness Evolving*

John Benjamins, 2002, 251pp., \$83.95 € 70.00,  
ISBN 9027251541

---

For some of us, multi-disciplinary forums are vital places to visit: as full of prospect as a high school disco to a rummaging teenager. Perhaps something concrete and tangible will emerge from the undefined; new relationships may be formed, or something wholly unexpected happen. They are also risky. Threads of logic are lost or misinterpreted in the fudge between disciplines; wool is pulled over naïve eyes. None of us is safe. We are all, to a lesser or greater extent, naïve. No one knows everything about every discipline. Is it worth taking the risk of visiting? Chances are that the successful academic, as a teenager, played it safe. On the other hand, those inclined to be seduced by the dance ended with a headache and too few points to get into university.

To my mind this book, volume 34 in the ‘Advances in Consciousness Research’ from John Benjamins, is an event worth turning up for. Philosophers, computer scientists and cognitive scientists contribute to an absorbing series of essays in order to discuss the past, present and future of consciousness. The principal topic of debate is whether consciousness has an independent adaptive function of its own and, if so, what this implies in terms of where it evolved from and what it could evolve into.

On one level, within the earlier essays, there appears to be an emerging consensus about the value of reverse engineering, the nature of adaptive advantage, and the types of conscious state that the contributors assume may relate to a perceived advantage.

The reader’s confidence begins to unravel, however, as the contributors grapple with the status and evolutionary significance of non-physical qualia. Initially and with a lucidity that makes the whole business, at first, seem child’s play, David Cole takes the reader through a series of thought experiments involving inverted and missing qualia, exploring, among

other puzzles, the inverted spectrum possibility. In searching for an argument for the adaptive advantage, he concludes that we cannot prove that qualia don’t have a functional role.

In fact, this conclusion of no proof, this agreement on reasonable doubt, becomes a predictable pulse throughout the collection – a measured beat over which the more sophisticated patterns of philosophical debate are laid. None of these patterns are ground breaking in nature. For the most part they are the revisited notions of dualism, supervenience and materialism, the original authors of which are, in every case, clearly and fairly accredited.

However, unless we get a grip on the fundamental issue at stake here – whether there is mind beyond matter (or matter beyond mind) – then it is hopeless to have a multidisciplinary, or even a multi-person, debate of any significance.

Stephen Clark grasps this particular bull by the horn and argues that we, in our reasoning, have chosen to create the ‘solid framework’ of physics; that there is nothing without mind. If one is to believe Clark, which incidentally I do, then it is impossible to take human consciousness out of the equation. Yet, in an apparent attempt to do just that, the compilers of this set of studies move headlong into the realm of computers and robots.

While most of the contributions are well written and understandable, two essays in the ‘Artificial Consciousness’ section require a degree of patience on the part of the reader. Philosophers can wrestle broad concepts using anecdotes and colourful illustrations; the computer scientist, however, has to deal with detail — small, modular programs and experiments — accounts of which are not always very stimulating. Nevertheless, if the reader perseveres through these two essays to Inman Harvey’s excellent study entitled ‘Evolving Robot Consciousness’ and matches it to an almost sentimental portrait of two ‘creative’ robot experiments given by Selmer Bringsjord, Ron Noel and David Ferrucci, then the significance of the journey will become apparent. With the experiments to develop programmes (aimed at achieving robot ‘consciousness’) which allow robots to apply selection criteria, the evolution of computer programmes is slipping like sand from the grasp of the computer programmer

into the hands of computers themselves. Could it ever, one wonders, in appearance at least, slip from there into the grip of the blind watchmaker (natural selection)?

Are we treading in dangerous territory, are we losing control? In the epilogue Neil Tennant seems to imply that we are. Drifting somewhat tongue-in-cheek in the direction of science fiction with the introduction of clones and Ripvan-Winkle characters he recommends that we give more thought to the issues — offering the reader a fanciful end to a spirited encounter, a romantic conclusion to that harmless high school disco. Or, without the resolution of the mind/matter problem, is it after all just the prospect of a quick fumble in a dimly lit minefield?

*Clare McNiven*

---

**John N. Findlay & Iain D. Gilchrist**  
*Active Vision: The Psychology of Looking and Seeing*

Oxford: Oxford University Press, 2003, xiii + 220 pp., £24.95, ISBN 019852479X (pbk).

---

Set the kids loose with the camcorder and the result is always the same. A mad dancing picture as if the scenes had been shot from a particularly hairy fair ground ride. Why can't they let the camera linger? As any psychologist will know, the lurching lens is just imitating the child's saccading eye. Kids naturally expect to be able to shift their attention in eyeball-swivelling leaps three or four times a second and still see a perfectly stable world. So no mystery. Yet in fact the business of saccades is emerging as one of the key questions in consciousness research.

In the 1990s, most would agree, the big issue was the binding problem. If the cortex is a hierarchy of processing modules, then how do the results of all the computations get glued together to make a seamless conscious whole? This question has been reasonably well answered with the discovery that coherent firing rhythms help knit widely scattered neurons into fleeting global ensembles (though of course there remain many rival views about precisely how such coherence is effected). Now mind scientists are asking the same kind of question but concerning the unity of consciousness over seconds rather than milliseconds. The naive view of the brain is that it merely reflects what the sense organs report. Yet if the eyes are

dancing from fixation to fixation, then the resulting cortex mappings must swing about like the pictures from the kids' camcorder. How can a steady stream of experience be constructed from such violently wrenching foundations?

Findlay and Gilchrist are experimental psychologists from Durham and Bristol University respectively. In this concise volume they aim to tackle the question of what they call trans-saccadic integration. Unfortunately much of the book is framed as an attack on old-fashioned passive vision research and so risks savaging the ankles of a strawman. Findlay and Gilchrist are certainly right that much vision work has done its best to factor out the need for eye movements. There has also been a near obsession with the phenomenon of covert attention — exploring the cognitive limits of peripheral vision. But this reductionist instinct to pursue the remote extremes of visual processing does not mean that the same researchers are not also aware that the pieces will eventually have to be put back together as a theory of normal 'active vision' where constant eye movements to exploit the acuity of central vision are the rule.

Nevertheless, the book is a timely one as there really is no good model of how consciousness is smoothed over in the face of lurching saccades, attention shifts, blanking eyeblinks and other such apparent disruptions to the rolling Jamesian stream of experience.

As Findlay and Gilchrist recount, the evolutionary need for saccade-based vision is clear enough. The cone-dense fovea covers only 1 degree of the visual field. It is reckoned that to have foveal level acuity over the whole retina, the brain would need to be hundreds of thousands times bigger and weigh ten tons!

Saccading itself is a complex affair. Even when the eyes are fixated, there are constant tremors and microsaccades that keep the retina 'refreshed'. Saccades then come at two rates depending on how habitual and predictable the situation happens to be. With highly predictable events, our eyes can flick to the spot in as little as a tenth of a second in an express saccade. For more exploratory looking, it takes a fifth of a second to find a new fixation point, and the movement often includes an overshoot and quick unconscious correction. During the saccade, retinal output is suppressed to cut

down the blur of motion — though recent experiments suggest it is more that the brain immediately forgets any visual input captured while the eye is in transit. Either way, add in the 20,000 eyeblinks we make each day and we are effectively blind about a third of the time!

Actual fixations last about a third of a second. But a recent surprise is that the time the eye lingers has more to do with the needs of the brain. A visual image can be ‘snatched’ in a tenth of second. It seems the brain then holds the eye in place longer largely to give itself grace to complete its processing. Fixation serves to prevent further input before the brain is ready!

A further surprise has come from change blindness experiments in which background visual features are sneakily altered while a subject’s eyes are in mid-saccade. Big changes in position, or the switching of object characteristics — a dog for a cow — will often go unnoticed. Consciousness accepts radical sensory change as part of its seamless flow. But on the other hand, if these visual features happen to have been actively attended during prior fixations, then any later changes become highly detectable. This is a puzzle as it used to be thought the brain stabilises vision by precise anticipation. Planned eye movements generate an efference copy — a forward motor projection — that exactly cancels every jump in the visual scene. But the change blindness studies show that sometimes there is this kind of anticipatory compensation, yet at other times the visual changes are simply glossed over.

To cut a long story short, the stability of conscious awareness seems the result of a clever mix of rapid forgetting and unconscious prediction. Raw sensory information appears to flow through to the cortex more freely than some older theories suggest. But it is just as freely ignored if it does not fit into a general, expectancy-based, running model of the world. The lurching shifts in visual input are not gated or cancelled out but downplayed, rendered vague, as the brain tracks its own gist-based view of what is happening. As with the child holding the camcorder, it is the inner narrative which IS the foundation for the mental experience and the thrice-a-second wrenching changes in eyeball imagery get cortically represented only to the degree that they play into the inner tale being told.

Findlay and Gilchrist’s *Active Vision* only really offers an introduction to the problem of trans-saccadic integration. In review article style, they skim a wealth of research, leaving themselves just a modest page or two to sketch their own conclusions. Yet still, so central is this issue to a proper understanding of the cognitive mechanisms underpinning conscious awareness, that many will consider it essential reading.

*John McCrone*

---

### **Benny Shanon**

*The Antipodes of the Mind: Charting the Phenomenology of the Ayahuasca Experience*

Oxford: Oxford University Press, 2002, 475 pp., £22.50 (UK pbk) / \$35 (US pbk), ISBN 0199252920 (hbk), 0199252939 (pbk).

---

Benny Shanon is a respected cognitive psychologist who has published many peer-reviewed papers, some of them here in the *JCS*. In the early 90s, he went to the Amazon basin to investigate the phenomenology of ayahuasca intoxication. This psychoactive brew was alleged to cause vivid visions with surprisingly similar content in those taking it. He gathered extensive information about its effects from interviewees as well as by taking the drug himself, which allowed him to describe the experiences of intoxication both anecdotally and quantitatively. These experiences are truly remarkable.

He obtained *objective* data from structured interviews with members of religions that regularly use the plant brew, from other experienced users, both urban and indigenous people, and also from a review of writings on the topic. He estimates that his data ‘are based on ... 2,500 Ayahuasca sessions’ (p. 410). His *subjective* data were obtained from his own ayahuasca ‘journeys’, undertaken in a variety of settings. The material in the book is mostly derived from his first 67 sessions, each summarized in notes at its conclusion. His very first major ayahuasca experience, amongst a Santo Daimo (which roughly translates as Saint Ayahuasca) group deep in the Amazon jungle, appears to have been truly profound. After *seeing* the atrocities of history, then the beauty humanity has produced, he felt born anew: ‘It seemed this was the first day of creation’ (p. 6).

Beyond his professional effort to establish categories of comparative data, Shanon also found compelling personal reasons to enter what he came to regard as the 'School of Ayahuasca'. Woven within his detailed analyses is a narrative of metamorphosis.

The book gives little space to the neuropharmacology of intoxication, perhaps because Shanon quickly came to '... totally reject the possibility that biological accounts — detailed as they may be — can offer viable psychological explanations' (p. 35). Yet within his work, the analytical approach of this stalwart cognitive psychologist seldom wavers. This must have required great self-discipline, particularly when the objects of study were his own overwhelming visions, emotional storms, or indeed the distractions of the gastro-intestinal 'purging' so often encountered (which Shanon learned to control by 'bringing up' songs instead of stomach contents!).

One of the most striking findings was that particular entities or objects played central parts in a startling array of visions, especially snakes of all sorts and jaguars. Strangely, such jungle denizens appeared in the visions of Europeans as regularly as in those of Amerindians. Often reported as well were sinuous dancers, unknown fabulous cities, and historical panoramas. Detailed tables of quantitative analyses are given in an appendix. Structural similarities in the progression through a single ayahuasca journey were also noted: the often terrifying visions of the first period — 'Quite commonly,' he states matter-of-factly, 'people feel that they are about to die' (p. 57) — to the serenity and communion sometimes experienced in later stages.

But Shanon emphasizes that undergoing ayahuasca 'education' is a long term project. Sitting in on a class or two just won't do. There are stages of advancement into these mysteries: the novice begins passively watching visions unfold as if on a screen but, with more experience and courage, learns to enter the vision and participate in its reality. Then comes a stage where some degree of control over the process is possible, though such control is always felt as exerted in partnership with some creative source outside the self. Mastery seems to involve gaining ability to engage with many *realities* simultaneously, thus allegedly conferring power *to act* in the everyday world in new

ways that a person might never previously have attained or attempted, through the expressive arts, for example, or guidance and healing.

Shanon emerges as something of a shamanic *ayahuasquero* himself. The most expressive symbol of ultimate truth is found, he concludes, in songs of praise for all creation; the *Hallelujah* of his Judaic ancestors. As to the ontological question of what exactly is being so praised, Shanon avers that it is not *anything* at all, but rather a universal process. *Creation* is what the name implies, the ongoing creative unfolding of all things, including ourselves.

Obviously, such 'knowledge' cannot be attained through analytic or phenomenological reduction alone. It is ever present beyond the edge of the 'known world', that is, beyond the confines of the conscious mind. This is the point at which Shanon the scientist admits that full communion exceeds communication: 'Yet, there were occasions that it was clear to me that I had to make a choice — if I really wished to undergo the experience presenting itself to me, I would have to forgo my future recollection of it and give up any thought of ever talking about it' (p. 355). This *unspeakable* awakening is not found with ordered signposts and structured roads but intuitive discovery, well beyond categorical reasoning:

[V]ery poignantly, I realized how limited the scientific approach is. ... I further comprehended that there are levels of knowledge that demand one to let go and relinquish all critical, distanced analysis. ... In this respect, despite all its limitations in terms of sociological power and cultural permanence, the indigenous stance has the upper hand (p. 356).

This is an important message from a contemporary scholar, but a puzzling one from many points of view. I had particular difficulty myself with Shanon's brief discussion of the ontological status of good and evil: 'The ultimate reality is beyond good and evil', he says, but 'there are visions in which one feels one is encountering the Supreme Good' (p. 174). Does the Supreme Good cast no shadow then? I'm likely missing the mystical paradox here, or at least the Trickster at play. Ayahuasca has a cosmic sense of humor (not always benign), according to Shanon, and hides as much as it reveals.

The whole question of applying orderly statistics to *other realities* seems perplexing. To what do the lists, tables, categories, and structures actually refer? Does structural analysis reveal the universal latticework of creation, or at least of the Ayahuasca experience? This is uncertain. Shanon is well aware of how boundaries in the realms of visionary experience seem to shift or even, with a wink and smile, disappear altogether. Such carefully delineated categories may really exist within psychic projections, but 'there is no clear-cut differentiation between interpretation and creation. ... In essence, all is interpretive, all is creative' (p. 351). It is not perfectly clear how categories of experience can be both evanescent creations yet simultaneously stable interpretations. But perhaps that's why Shanon uses music making as his primary metaphor: the player creates by interpreting and vice versa.

Many scientists may be unpersuaded, at the very least, by Shanon's claim that ayahuasca visions and revelations 'might be psychological and creative *and* real' (p. 401). New Age folk may be disappointed that he remains unconvinced of reincarnation, ESP, or 'the paranormal actuality of the astral' (p. 363). Instead of sharpening distinctions, he found them disappearing — 'between creation and discovery, between psychology and ontology, between mind and world, between internal consciousness and the *anima mundi*, between the human and the Divine' (p. 401). This is a slippery metaphysics. Shanon himself allows that the many ontological questions arising from his experiences need much closer examination than they are given in this book, which primarily derives from phenomenology. Nevertheless his songs of praise for a transcendent creativity active here and now, within and around us, point to something profound. The apparently firm outlines of our 'given' world may, it appears, be but one reality among an infinitude of (equally valid?) experiential possibilities.

Greg Nixon

---

**Bill Fulford *et al.*, ed.**

*Nature and Narrative: An Introduction to the New Philosophy of Psychiatry*  
Oxford University Press, 2003, 286 pp. ISBN  
0198526113 (pbk).

---

There seems to be increasing interest in the areas of overlap between mental health and philosophy. The editors of this book introduce it as the first of a series designed to address this interest, and it lays out some of the current preoccupations in the philosophy of mental health. Many of the topics and the contributors will be familiar to readers of the journal 'Philosophy, Psychiatry & Psychology'.

Just how 'new' is this 'new philosophy of psychiatry'? In their introduction (which serves as a useful overview of the book as a whole) the editors put forward several pieces of evidence for their claim that there is something distinctively different in current thinking from what has gone before. The foundation of the above-mentioned journal, the thriving academic interest in the subject (including the appointment of a chair in 'Philosophy & Mental Health' at Warwick University), the increasing emphasis on the humanities in medical and psychiatric training; all suggest that some sort of 'sea change' is occurring. The actual topics are not really new at all: such questions as the status of mental illness, the relationship between mind and body, the nature of consciousness and the self are fairly constant human preoccupations. The approaches to these problems, most prominently post-Wittgensteinian analytical philosophy, phenomenology and hermeneutics, again have continuity with previous work in the area. What could be seen as innovative over recent years is the contribution of ideas from more empirical disciplines, such as cognitive psychology, linguistics, information technology and functional brain imaging, to what were previously viewed as purely theoretical problems. There is some evidence of this opening out of the field in the present volume, although most of the contributions in fact take a purely conceptual approach to the problems they examine.

Of these themes, the most fundamental to this volume is the tension between causal and meaningful explanations of mental disorders. It is explored in different ways in several of the

contributions. Although contributions range widely over a number of other concerns, from an analysis of Wittgenstein's philosophical method to a discussion of the comprehensibility of delusions, this theme gives unity to the volume as a whole.

How practically relevant is this book to people like me who have a predominantly clinical interest in mental disorder? Individual chapters vary considerably in this respect. A few of them assume quite a sophisticated level of background philosophical knowledge. Derek Bolton's chapter on meanings and causes, for instance, develops arguments set out elsewhere and so might be frustrating for someone looking for an initial framework for thinking about this important topic. One chapter, that by Natalie Depraz, is frankly incomprehensible to anyone not thoroughly immersed in the intricacies of French cultural studies.

Despite these caveats, the majority of the chapters are useful and absorbing from both an intellectual and a practical point of view. To pick out two among many, I particularly enjoyed those by Eric Matthews and Alfredo Kraus. Matthews provides a model of mental illness, based on Merleau-Ponty's phenomenology, which successfully reconciles conflicting theories, whether dualist or reductionist. Its consideration by psychiatrists might encourage a more holistic view of patients and their problems. Similarly, Kraus's excellent contribution develops a 'phenomenological-anthropological' approach — roughly speaking, this involves trying to get closer to the patient's own experience — in contrast to the prevailing 'symptomatological-criteriological' approach, which equates with most current psychiatric practice. After reading this chapter I felt not only that I had learned something new but also that it contained potentially valuable implications for psychiatric training.

So, despite being something of a mixed bag, overall this is a highly successful first venture for the new series. Several of the issues raised are fundamental to the humane practice of psychiatry and should be part of any training curriculum. Despite the difficulty of some (by no means all) of the contributions, I think that this book would repay reading by anyone in the mental health field who has a curiosity about conceptual issues and who is prepared to ques-

tion ideas and assumptions which are so often taken for granted in our work.

*Peter Howorth*

---

**David Herman, ed.**

*Narrative Theory and the Cognitive Sciences*

Stanford, CA: CSLI Publications, 2003, 363 pp., \$27.50, £19.50 ISBN 1575864681 (pbk).

---

On the one hand you could say that humans have a passion for storytelling, so much so that we spend huge amounts of our free time in reading books or watching television programmes. Even when we are apparently engaged in employment, our methods of communication often come down to telling each other stories. On the face of it then, storytelling must have adaptive evolutionary advantages. It is very easy to form naïve theories as to how this could work: Telling stories is an easy way to teach children; it is a way of practicing something difficult we are planning to do; it is a way of experiencing vicariously; by telling stories to ourselves we can reinforce our own identities. Furthermore, if storytelling is so central to the human experience, it can surely tell us something about human cognition and intelligence.

On the other hand you could, for example, speculate that 'utterances should be viewed not as manifestations of a code that pre-exists all communicative acts, but rather as an interactional achievement, negotiated by participants using an inherently variable linguistic code in situated contexts' (p. 7). This could further be located in the space, known as cognitive narratology, where cognitive and social psychology meet linguistics and literary theory.

These two approaches, for me, typify why storytelling is such a fascinating human attribute but also show how we love to take a simple idea and obfuscate it. There is undoubtedly academic mileage to be had from the idea of systematically studying the construction and use of narratives in human cognition, but there is also a narrow line between such an enterprise and the further reaches of deconstructionism that some literary theory encourages.

A further complication, in this particular book, is that the editor is seeking to expose the

links between narrative theory and cognitive science. So there is a need to ensure we are asking the right questions and making linkages that can be tested experimentally. My thoughts on first reading this volume, bearing in mind that my own specialism is artificial intelligence, is that the authors seem mostly to be wedded to a view of cognitive science that is still rooted in a rule-based, symbolic view of human cognition. This is a rather outdated approach and one that the AI community lost faith in some years ago. A good example of relying on the rule-based approach is the editor's discussion (pp. 10–11) of Schank and Abelson's 'Scripts, Plans and Goals', which seeks to codify common human endeavours by extracting stereotypical actions. David Herman does mention the limits of this approach but nevertheless tells us that work in artificial intelligence 'continues to underscore the necessity of viewing narrative itself as a relational construct' (p. 11). Surely this strategy is like that of the drunk who insists on looking for his lost key under a streetlamp, not in the dark alley where he most likely dropped it. Most AI workers in computational linguistics have now abandoned the idea of trying to represent natural language as sets of relational constructs and moved on to statistical approaches. In other words the prevailing paradigm in artificial intelligence is a non-symbolic one, which makes it rather difficult to fit a largely rule-based cognitive theory of narratology into the currently popular non-symbolic view of human cognition.

This book is a useful collection of papers which will serve to introduce the general reader to narrative theory and the way that literary theorists in particular view its relationship to cognitive science. Like most collections of papers, the standard of writing and discussion is variable. Some topics are presented in a very clear fashion whilst others are obscure. The highlights for me include the introduction by the editor, David Herman, which gives an excellent overview of the emergence of narratology as an area of study and which largely succeeds in convincing that it is a worthwhile enterprise. Someone new to the topic would find this an excellent overview and come away with clear suggestions for further reading. I also commend Herman's own chapter 'Stories As a Tool for Thinking' which gives a very unambiguous overview of just what it says on the tin (not

always the case in academic papers!). He is also modest enough to conclude that this chapter 'merely outlines a program for future research.' (p. 185)

At the end of his main chapter Herman expresses the common sense view that I outlined in my first paragraph - 'stories can be studied as a primary resource for building and updating models for understanding the world.' (p. 185). Unfortunately many of his contributors deviate from this approach and seek to build models or theories which rely on unnecessarily complex ideas and in particular try to build on what we would now disparagingly call Good Old-Fashioned Artificial Intelligence (GOF AI). The book is certainly worth looking at if you want to see the view from narrative theory, but I would argue that the view from cognitive science presented here is somewhat outdated. The enterprise of situating narrative theory in cognitive science is a worthwhile one but most of the papers in this book do not make a believable case for how it might be done.

*David W Salt*  
University of Derby

---

**Eric B. Baum**  
*What Is Thought?*

Cambridge, MA: MIT Press, 2004, 495 pp., \$40,  
ISBN 0262025485 (hbk).

---

In the introduction to this handsomely bound book, the author suggests that it is an appropriate time for an explanation of how the dynamics of a human brain can be accounted for by computer science. His title is motivated by Erwin Schrödinger's enormously influential *What Is Life?*, which launched the field of evolutionary biology by inducing both Francis Crick and James Watson to successfully seek the molecular basis of biological evolution (Moore, 1989), but the analogy is strained for several reasons.

Schrödinger's book is less than 100 pages in a current edition (1992), while Baum's is about five times as long. In the context of Schrödinger's lifelong interest in biological problems and based on a series of three public lectures that he presented to the Irish intelligentsia in 1943 (as one of his statutory duties as the founding director of the Dublin Institute of Advanced Studies), *What is Life?* is a classic example of his exceptional expository skill —

in a second language, no less — whereas Baum's book would have profited from another round of copy-editing. But the most striking difference between these two titles lies in the cogency of their respective contents.

Although Max Delbrück and his colleagues (1935) had used measurements of mutation rates of fruit flies under X-radiation to show that their genes were necessarily of molecular dimensions in the mid-1930s, the implications of these data were unnoticed by the literate world of the mid- 1940s. Thus Schrödinger's public lectures were newsworthy, being favourably noted by *Time* magazine in the spring of 1943, and his subsequent book — after some difficulties with an Irish publisher and the Roman Catholic Church over the religious implications of his ideas — went on to sell over 100,000 copies for Cambridge University Press, with translations into seven languages. Is there a similar communications gap in our current understanding of the nature of thought?

In an internet review of the second international consciousness conference which was held at the University of Arizona in 1996, the *Times Higher Education Supplement* (THES) noted — with tongue firmly in cheek — no less than eleven current approaches to the understanding of human thought: traditional neuroscientists ('white-coat brigade'), philosophers ('hard-problemites'), artificial-intelligence researchers ('brain-on-a-disc brigade'), cognitive scientists ('it's the thought that counts'), carbon chauvinists ('get a life'), emergent hierarchists ('wait, there's more'), mysticians ('why bother?'), quantum physicists ('consciousness plugs into the universe'), contemplators ('turn off your mind, relax, and start taking notes'), folk psychologists ('yup, I'm conscious'), and finally the parapsychologists ('what am I thinking?'). With the present book, evidently, Baum is digging into some rather thoroughly tilled ground.

Noting his background in computer science, the THES would archly classify Eric Baum among those who believe that 'our souls are software', but this is not quite fair. Although he states that 'the obvious inability of present-day computer science to account for [the brain's behaviour] is no reason at all for doubting that they can be accounted for by computer science', the intellectual perspectives of *What Is Thought?* are broader than this assertion seems

to suggest. The book begins with several interesting chapters on the nature of computation (I particularly liked the presentation of the traveling-salesman problem), which include discussions of the importance of making decisions at the level of semantics, the Turing test, properties of neural nets, hill climbing in a fitness landscape, among several other relevant topics. These discussions lead into the author's central thesis that the mind, like all efficient computer programs, is necessarily modular. In other words, each aspect of the brain's dynamics comprises several subroutines, which presumably can be further broken down into hierarchical structures of nested activities, and he discusses several permutations of this important concept.

Curiously, Baum's otherwise comprehensive list of references does not include Donald Hebb's seminal and classic work, in which the notion of 'cell assemblies' (which are dynamically self-sufficient modules of neurons) was first suggested over a half-century ago (Hebb, 1949). As a psychologist, Hebb aimed to 'bridge the long gap between the facts of psychology and those of neurology,' and coming at about the same time as the development of the digital computer, his formulation has provided the basis for many numerical studies starting in the 1950s and continuing to the present day which are in accord with a growing body of electrophysiological data (Edelman, 2004). Setting this quibble aside, Baum offers compelling psychological evidence for the modular structure of mind and provides his readers with an interesting and informative account of how the structure of our thinking may have developed over the course of biological evolution, with particular attention paid to computational constraints on the development of learning mechanisms. Importantly, his perspectives are broader than those of many of his colleagues, as he asserts that the 'whole program' of a brain's dynamics includes the 'complex society' in which it is embedded. Indeed, the author's evident humility in the face of awesome intricacy of mental activity is, to me, one of the more appealing aspects of *What Is Thought?*

The often suggested possibilities for quantum computation are discussed in some detail, along with an analysis of the widely noted example of 'Schrödinger's cat' which was

originally proposed to emphasize the difficulties of applying ideas developed for atomic dynamics to complex macroscopic systems. Considering that a quantum computer — if it is at all possible to construct one — must be carefully isolated from structural irregularities and operated near absolute zero of temperature, Baum joins the majority of physical scientists in concluding that it is ‘highly unlikely that quantum computation is relevant to the mind’.

Eric Baum has a dog, and — like most of us dog owners — he is convinced that his pet is conscious, but he goes on to assert that ‘we do not need to posit new qualitative modes of thinking to explain human advance over animals. To my mind, the difference between human intelligence and animal intelligence is straightforwardly explainable by cumulative progress once there is the ability to communicate programs.’ Here, again, Baum could profit from reading Hebb’s (1949) book, which contains but a single mathematical expression, namely

A/S.

This parameter represents the ratio of the associative area (A) of a mammalian neocortex to its sensory area (S), and it becomes greater as one progresses from rats through dogs to humans. A related physiological parameter — with profound significance for the ease and rate at which modules (or cell assemblies) can switch on and off — is the percentage of inhibitory inter-cortical neurons, varying as follows: rabbit (31%), cat (35%), monkey (45%), human (75%) (Rakic, 1975). Of course, these relative differences may be examples of the ‘cumulative progress’ to which Baum refers.

In a penultimate section, Baum discusses the question of free will, noting that ‘our decisions look, from any reasonable perspective short of knowing the exact state of our brains and simulating them in detail, like they are introducing genuinely new information’. In reaching this conclusion, he may be confused by the continuing tendency of many scientists to overlook a phenomenon called ‘sensitive dependence on initial conditions’ first studied by the eminent French mathematician Henri Poincaré (2001/1903) and widely observed nowadays by those who study nonlinear dynamic phenomena (chaos theory). As Poincaré famously put it over a century ago:

If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment, but even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation approximately. If that enabled us to predict the succeeding situation with the same approximation, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon.

For an author who bases many of his conclusions on close mathematical reasoning and offers a theory that purports to be ‘capable of explaining everything,’ the implications of these ‘fortuitous phenomena’ should be carefully digested.

Alwyn Scott  
University of Arizona

### References

- Delbrück, M., Timofeef-Ressovsky, N.W. and Zimmer, K.G. (1935), *Nachr. Biol. Wiss. Göttingen*, **1**, pp. 189–245.
- Edelman, G.M. (2004), *Wider than the Sky: The Phenomenal Gift of Consciousness* (New Haven: Yale University Press).
- Moore, W. (1989), *Schrödinger: Life and Thought* (Cambridge: CUP).
- Hebb, D.O. (1949), *The Organization of Behavior* (New York: John Wiley & Sons).
- Poincaré, H. (2001), *Science and Method* (Chicago, IL: St. Augustine’s Press; first published in 1903).
- Rakic, P. (1975), *Neurosci. Res. Prog. Bull.* **13**.
- Schrödinger, E. (1992/1956), *What is Life? with Mind and Matter* and autobiographical sketches (Cambridge: CUP).