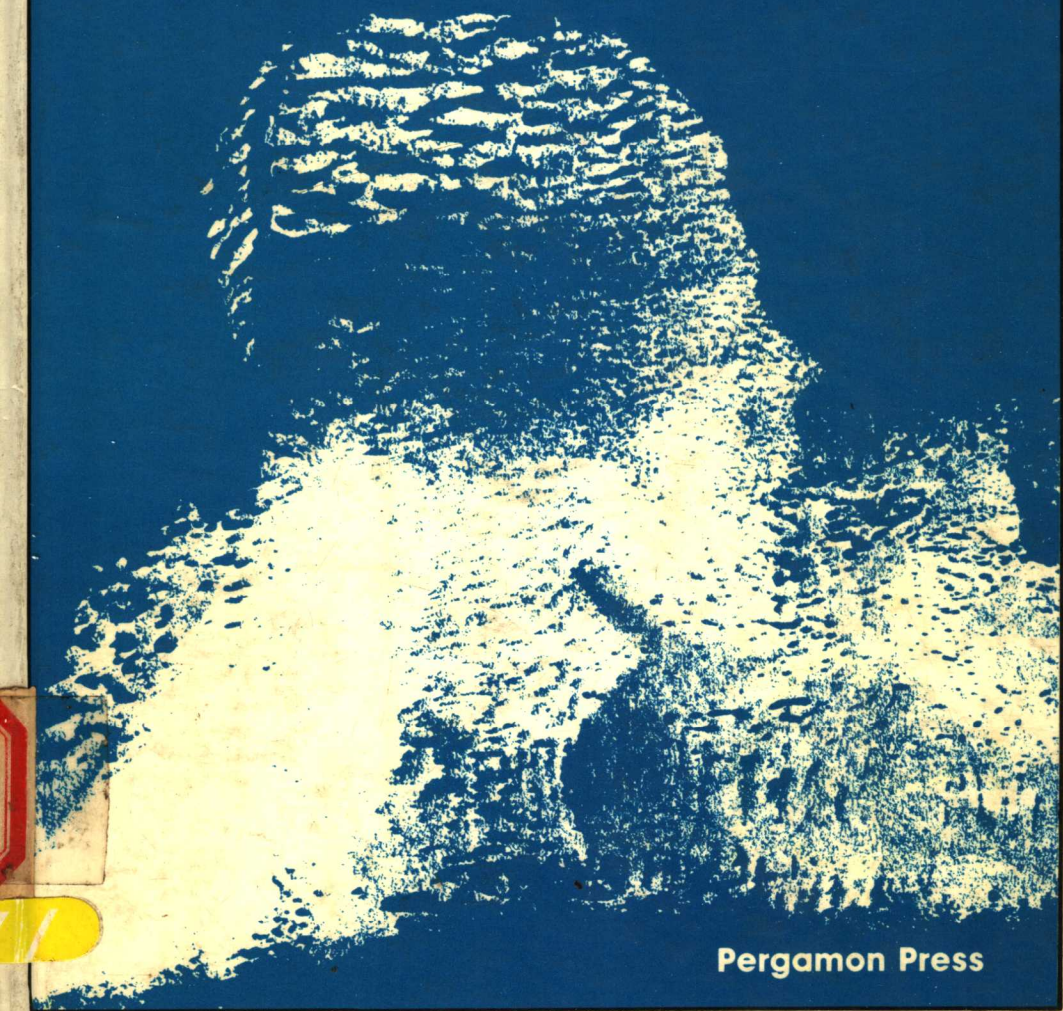


Consciousness and the Physical World

Edited by **B D Josephson** and **V S Ramachandran**



Pergamon Press

CONSCIOUSNESS AND THE PHYSICAL WORLD

EDITED PROCEEDINGS OF AN INTERDISCIPLINARY SYMPOSIUM ON
CONSCIOUSNESS HELD AT THE UNIVERSITY OF CAMBRIDGE IN JANUARY 1978

EDITED BY

B. D. JOSEPHSON

Cavendish Laboratory, Cambridge

and

V. S. RAMACHANDRAN

California Institute of Technology, U.S.A.

with a Foreword by

F. J. DYSON

Institute for Advanced Study, Princeton, New Jersey



PERGAMON PRESS

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

Foreword

F. J. DYSON

This book stands in opposition to the scientific orthodoxy of our day. The orthodox dogma is stated by the biologist Jacques Monod in his book *Chance and Necessity* with characteristically French sharpness: "The cornerstone of the scientific method is the postulate that nature is objective. In other words, the systematic denial that true knowledge can be got at by interpreting phenomena in terms of final causes—that is to say, of purpose." Monod labels those who disagree with him "animists". The arch-animist is Teilhard de Chardin, for whom Monod reserves his deepest scorn: "The biological philosophy of Teilhard de Chardin would not merit attention but for the startling success it has encountered even in scientific circles. . . . There is no inert matter, and therefore no essential distinction between matter and life. . . . For my part I am most of all struck by the intellectual spinelessness of this philosophy. In it I see more than anything else a systematic truckling, a willingness to conciliate at any price, to come to any compromise. Perhaps, after all, Teilhard was not for nothing a member of that order which, three centuries earlier, Pascal assailed for its theological laxness."

The authors of this book are not followers of de Chardin. They represent a variety of scientific disciplines and a variety of philosophical viewpoints. But they are all, according to Monod's definition, animists. That is to say, they are not willing to exclude *a priori* the possibility that mind and consciousness may have an equal status with matter and energy in the design of the universe. They are trying to extend the boundaries of scientific discourse so that the subjective concepts of personal identity and purpose may come within its scope. They are all to some extent exposing themselves to the charges of ideological laxity with which Monod lambasted de Chardin. They are accepting a certain risk that their orthodox colleagues will consider them a little soft-headed.

I am delighted to see that the contributors to this book include more biologists than physicists. In recent years biologists have usually been more inhibited than physicists in stepping outside the accepted norms of scientific respectability. Monod was, after all, a biologist. In dealing with the problems of consciousness, physicists have had courage but no competence, biologists have had competence but no courage. In this book we see some examples of competence combined with courage.

Why have the biologists during the last century been so inhibited? I believe they are still suffering from the after-effects of the great nineteenth-century battle between the evolutionists led by Darwin and Huxley, and the churchmen led by Bishop Wilberforce. The high point of the battle was the great debate in Oxford in 1860 during which Bishop Wilberforce asked Huxley whether he was descended from a monkey on his grandfather's or on his grandmother's side. Huxley won the debate, but the biologists are still fighting the ghost of Bishop Wilberforce. In the bitterness of their victory over the forces of religious orthodoxy, they have made the meaninglessness of the universe into a new dogma. "Any mingling of knowledge with values is unlawful, forbidden", says Monod.

The authors of this book have defied Monod's anathema. They have wandered freely over the borderland between science and philosophy, where knowledge and values are inextricably mixed. I believe they have brought back some insights which will be illuminating not only to scientists but also to anybody with a philosophical turn of mind who enjoys pondering over the mysteries of mind and consciousness.

Contents

<i>Foreword</i>	vii
F. J. DYSON	
Introduction	1
V. S. RAMACHANDRAN	
Part I. General	17
1. What Defines Privacy?	19
G. VESEY	
2. Regarding Consciousness	31
R. L. GREGORY	
3. Is Consciousness a Phenomenon?	49
H. C. LONGUET-HIGGINS	
Part II. Consciousness and Behaviour	55
4. Nature's Psychologists	57
N. K. HUMPHREY	
5. Nature's Joke: A Conjecture on the Biological Role of Consciousness	81
H. B. BARLOW	
6. Conscious Agency with Unsplit and Split Brains	95
D. M. MACKAY	
7. Some Hypotheses Concerning the Role of Consciousness in Nature	115
B. D. JOSEPHSON	

Part III. Subjective Experience	121
8. Consciousness and Psychopathology M. ROTH	123
9. Twins, Split Brains and Personal Identity ' V. S. RAMACHANDRAN	139
10. Mind-Matter Interaction in the Psychokinetic Experience S. PADFIELD	165
11. Phenomenal Space M. J. MORGAN	177
Afterword to the Conference: <i>The Prospects for Consciousness Research</i> B. D. JOSEPHSON	193
<i>List of Participants</i>	197
<i>Name Index</i>	199
<i>Subject Index</i>	201

Introduction

V. S. RAMACHANDRAN

Trinity College, Cambridge

This book is about consciousness and is based on a symposium on that subject held at the University of Cambridge on 9–10 January 1978. Usually, books on scientific or philosophical subjects are edited by experts on the subject matter of the book itself. I make no apology for the fact that this particular book is an exception to that rule – since there can really be no such thing as an “expert” on a subject as nebulous as consciousness.

Although scientists often have their own private views on consciousness they are usually reluctant to talk about these views. There are two reasons for this. Firstly, scientists are generally unwilling to venture into realms outside their legitimate scope or to speculate on questions for which there can be no precise empirically demonstrable answers. Secondly, there is a widely prevalent superstition among them that interest in such “fringe areas” is a sign of woolly thinking and declining intellectual vigour. Perhaps this explains their curious silence and their unwillingness to publish philosophical speculations.

The purpose of the Cambridge conference was to encourage distinguished scientists to express their views on the relationship of conscious experience to the physical world.* To add a sense of proportion we also invited a professional philosopher (G. Vesey) and a person claiming psychokinetic powers (Suzanne Padfield). By doing this we have tried to represent as wide a spectrum of views on consciousness as possible.

And as the reader will notice, the spectrum is very wide indeed – ranging from Barlow’s materialistic account (that consciousness is nature’s “trick”

*We are grateful to Research Corporation of New York for a grant out of which this conference was supported.

to chain us to our herd) to Josephson's view that minds may even have certain attributes of their own (e.g. "creativity" or "intelligence") to help channel the activity of physical brains towards specific goals. Yet in spite of these wide-ranging views, some of them flatly contradicting each other, a surprising degree of communication was achieved between the various speakers. What emerged was this book, whose contents I shall attempt to summarize in this Introduction. *

The publication of an interesting book by Popper and Eccles,¹ *The Self and its Brain*, coincided with the conference, and since the ideas in that book are rather similar in spirit to some of those which were discussed at the conference, it may be relevant to begin our survey with some of Popper's ideas. Popper calls himself a "dualist" and "interactionist", and believes in what he calls World 1 (the material universe, including physical brains), World 2 (individual human minds) and World 3 (language, culture, science and other products of World 2). He suggests that although World 3 originally emerged as a product of World 2, it seems to have acquired a life of its own and is no longer chained to individual minds. He speaks of World 3 "objects" like numbers, ideas, numerical concepts, etc., which are in some respects analogous to the *physical* objects of World 1. Calling ideas and numbers objects may sound like an elaborate joke to some readers, but in defence of his thesis Popper points out that:

- (a) World 3 has a quasi-independent status and would exist even if individual men died.
- (b) Many World 3 attributes are *unplanned* consequences of collective culture (e.g. Goldbach's conjecture and other hidden properties of number systems that are discovered by mathematicians just as an archaeologist discovers a World 1 object).
- (c) World 3 properties are often novel and "emergent", i.e. irreducible to the properties of individual minds – just as brains may have properties which are irreducible to single neurons.
- (d) Finally, one can imagine chains of causation in World 3 that are logically independent of (though necessarily accompanied by) physical causation in World 1. For instance, two computers that are grossly

*The speakers were encouraged to correspond with each other after the conference and this additional discussion is also included in the book.

different physically can nevertheless operate according to the same "standards of logic" (which are World 3 entities).

Popper also emphasizes that Worlds 2 and 3 are symbiotic since culture can "feed back" to enrich and expand individual minds. "Matter", he argues, "can thus transcend itself by producing mind, purpose and a world of the products of the human mind. One of the first of these products is language. In fact I conjecture that it was the very first of these products, and that the human brain and the human mind evolved in interaction with language." Elsewhere: ". . . As selves, as human beings, we are all products of World 3 which, in its turn, is a product of countless human minds."

It is important not to evade the chicken-or-egg aspects of this theory. Fortunately both authors (Eccles and Popper) give some thought to the apparently insuperable problem of how a closed system like the physical universe can "interact" with minds. Eccles begins by making the deliberately outrageous suggestion that the physical world is in fact *not* a closed system and that World 2 can directly influence the activity of brains.* The self-conscious mind, according to him, may act on certain "open" elements in the nervous system (such as synaptic clefts), which are so minute that even Heisenbergian uncertainty can influence their behaviour. The activity of these structures could then become magnified to account for brain events corresponding to human "choice" or "creativity".

Not everyone would find this view very satisfactory. If a combination of sub-atomic uncertainty (World 1) and the constraints of rational thought (World 3) can account for human freedom and creative enterprise, then what *need* is there for World 2? There is, after all, nothing logically impossible about World 1 objects (brains) creating World 3 *without* the intervention of World 2; so Eccles's own argument seems to suggest that minds are redundant by-products of evolution!

In spite of these difficulties *The Self and its Brain* contains some bold and powerful arguments for dualism and is sure to provide a valuable stimulus

* The authors seem to rely largely on introspection for arriving at some of these conclusions. For instance, the fact that people can reverse Necker cubes or engage in adventurous mountain climbing (Popper, p. 146) is cited as evidence for the view that the conscious self has "taken over" the activities of brains!

to new enquiry. If the Cambridge Symposium (embodied in this book) provides a similar stimulus, it will have achieved its purpose. It begins, appropriately, with a scholarly chapter by G. Vesey which contrasts sharply with some of the more light-hearted chapters in the book. The other contributors include three psychologists (R. L. Gregory, N. K. Humphrey and M. J. Morgan), three physical scientists (B. D. Josephson, H. C. Longuet-Higgins and D. M. MacKay), two physiologists (H. B. Barlow and myself) and a psychiatrist (M. Roth).

THE SOCIAL DIMENSIONS OF CONSCIOUSNESS

Chapters 4 and 5 form the core of the book and deal with speculations on the possible evolutionary significance of consciousness. Barlow's suggestion (Chapter 5) is novel and surprisingly simple. He begins by rejecting "parallelism" (i.e. the view that consciousness simply parallels any complex neural event such as the activity of MacKay's "supervisory" system, described in Chapter 6) on the grounds that if consciousness merely parallels complex neural events there is no reason why only a tiny fraction of such events should emerge into awareness. He suggests, instead, that consciousness may have emerged as an evolutionary novelty among social animals to permit gregariousness and communication. Thus consciousness, according to him, is "interaction and not a property". We feel pain only in order to communicate it, and if the need to communicate it had not arisen (e.g. in non-social animals like frogs or lizards) there would only be reflex withdrawal unaccompanied by the subjective sensation of pain. Perhaps the fact that people generally shout when jabbed with a needle supports Barlow's argument, but then why is the pain often felt *after* the shout?

Barlow also suggests that archetypes of other people are modelled into our brains by natural selection, and that consciousness consist either of real conversations with other individuals or of imaginary conversations with those archetypes (psychologists would call this "internal rehearsal"). Consciousness in his view is synonymous with *communication*. It would be biologically useless to communicate certain brain events (like the pupillary light reflex and reflex arcs regulating visceral functions, etc.) and therefore these events never emerge into consciousness.

Note that Barlow is not merely saying that communication adds an extra dimension to consciousness (a point that is already implicit in Popper's ideas), but that communication *is* consciousness. What he claims to have found is a correlation between certain kinds of neural events and consciousness — namely those neural events which are involved exclusively in communicating with other brains. Of course Robinson Crusoe was also conscious, but that is because his brain was engaged in imaginary dialogues with archetypes of other people.

Humphrey (Chapter 4) also emphasizes social aspects of consciousness but in a sense of his argument is the exact converse of Barlow's. He points out that a person who has never felt (say) pain cannot meaningfully understand or interpret the behaviour of another person being exposed to painful stimuli and would consequently be unable to communicate* effectively with him. From this example, he argues that the biological function of the *sensation* of pain lies in its usefulness for social interaction. Thus we feel pain in order better to understand the pain felt by others. He argues further that such subjective sensations evolved primarily to permit an animal to attribute reasons for its own behaviour and consequently to make sense out of the behaviour of other members of the social group.

Although at first sight Humphrey's argument seems flatly to contradict Barlow's, there is really no fundamental inconsistency, since both authors emphasize the importance of social factors and suggest that consciousness may have an evolutionary function. Thus, while Humphrey suggests that introspection is necessary for modelling archetypes of other people, Barlow regards conversations with archetypes as almost synonymous with introspection. Barlow speaks of communication with people "enriching" our conscious experience whereas Humphrey points to people who seek out new subjective experiences in order to enrich communication with others! A biologically inclined philosopher might support Barlow, but Humphrey's more introspective account seems closer to common sense.

To use Popperian terminology, Barlow is suggesting that World 2 (mind) is compulsorily parasitic on World 3 (which includes languages and culture). This is a bold departure from Popper's own interactionist view that Worlds 1, 2 and 3 exist independently while interacting to enrich each

*Here, and elsewhere, I use the word communication in its widest sense (and interchangeably with social interaction). The word should not be taken to mean verbal communication alone.

other. Humphrey, on the other hand, sticks to the Popperian tradition, and his view would be consistent with the suggestion that World 3 (as well as simple communication with others which is a necessary antecedent of World 3) would not have arisen if World 2 had not crept into physical brains at some stage in evolution, i.e. in his account World 2 would necessarily precede World 3. However, it is not clear whether either author would want to argue that the survival value of World 3 *actually exerted selection pressure for the emergence of World 2*. This, it seems to me, is the crux of the whole debate.

These considerations must lead us to a synthetic view of the evolution of mind. Perhaps at some stage in phylogeny, consciousness emerged as an incidental by-product of certain complex neural events. This new property was unplanned for, but once it emerged it made communication possible, since animals could begin to "introspect" and (by analogy) make sense out of one another's behaviour. Since communication has survival value, natural selection seized upon these neural events which were associated with consciousness and this in turn led to a mutually reinforcing interaction between collective culture and individual minds.

Such an account would be wholly consistent with Humphrey and Popper but would also leave several questions unanswered. Implicit in all the views presented so far is the assumption that consciousness is *causally* important for communication. For if it were not causally important then natural selection could not have favoured its emergence and its absence would have made no difference to the course of evolution. On the other hand, if its presence *does* make a difference we would have to assume that minds can actually exert an influence (however indirect) on the course of events in the physical world – particularly on a small portion of the physical world consisting of communicating brains. The implication of this would be that (a) the physical world is not a "closed system" and that (b) minds *cause* communication and do not merely accompany it.

This gets us into logical difficulties. Can consciousness really cause neural events? Stimulating the cortex can lead to mental events (e.g. phosphenes), but the converse would be hard to demonstrate empirically. Josephson accepts "mind acting on brain" as being almost axiomatic but is there any evidence to justify such a view? Unfortunately we are not even sure of what cause-and-effect *means* when talking about brain events and mentation. Gregory (Chapter 2) points out that our common-sense notions

about causation are hopelessly muddled, and he illustrates this with the example of night following day. Obviously day does not cause night; nor are they both caused by some third agent. Instead we see the night-day sequence as part of our *conceptual model* of the solar system. Similarly, the nature of brain-mind causation may become clearer when we start seeing it as part of a larger (hitherto undiscovered) conceptual scheme. Perhaps the causal links between brain events and mentation belong to a logical category that is quite distinct from, and are of a much more subtle nature than, the causation we talk about in the context of objects and forces. (Though, heaven knows, these words beg enough questions themselves!)

And nowhere is the problem of disentangling cause and effect more difficult than in the World 2 \Rightarrow World 3 interactionism proposed by Popper and Eccles. Could World 3 have arisen at all in the absence of at least a rudimentary World 2, and if so could the survival value of World 3 have exerted any selection pressure for the emergence of World 2? Did the dim introspective abilities of *Proconsul* necessarily antedate his ability to communicate, and if so did the culture which emerged from such communication propel him onwards to become *Homo erectus*? The theories of Barlow and Humphrey (as well as Popper's interactionism) may well contain partial answers to these important questions.

WHAT IS CONSCIOUSNESS?

While engaging in philosophical discussions of this kind there is always the tendency to forget that problems of consciousness are not merely of academic interest. To a patient in a hospital, experiencing intense pain or anguish, what we have said so far in this chapter, and any talk about consciousness being a "ghost in the machine", would seem curiously irrelevant or even perverse (Gregory, p. 31). Fortunately, this deficiency is remedied by Roth (Chapter 8), who surveys the phenomenology of consciousness from a clinical point of view, and by Longuet-Higgins (Chapter 3), who examines the validity of common-sense criteria which people generally use for deciding whether someone is conscious or not. Relying largely on common sense, Longuet-Higgins argues that the encodability of events into memory seems to be an invariant correlate of conscious experience — i.e. if a person remembers something, he must have

been conscious of it in the first instance. This seems to be generally true, but it is not difficult to think of possible exceptions. For instance, we often remember dreams vividly and *attribute* consciousness to dreams while recalling them later, but does it necessarily follow that we were conscious during the dream?

Josephson's approach to consciousness (Chapter 7) differs radically from those of the other contributors. Most scientists start with the brain and ask themselves why certain brain events seem to be associated with consciousness. Josephson's point of departure, on the other hand, is in consciousness itself, which he suggests can be empirically studied by *introspection*.

He begins with consciousness as a "given thing" and points out that our minds seem to have certain obvious attributes like creativity or intelligence. He regards these attributes as being almost axiomatic since we *know* them to be there from our own personal experience. Might we then not start with these almost axiomatic observations on consciousness and then try to arrive at more general "laws" of behaviour? Josephson points out that there already exists an extensive introspective-phenomenological account of consciousness to be found in the Eastern philosophical literature.² He uses ideas from this literature and tries to construct a theory of consciousness based on concepts borrowed from systems engineering.

Professional psychologists frown on introspection largely because other professional psychologists would frown on them if they did not. There is, after all, no *a priori* reason for starting with brains and working up towards consciousness instead of vice versa. In fact, to a person untrammelled by conventional scientific training, Josephson's approach might seem much more simple and straightforward. Galileo and Newton began with observations about the *physical* world and went on to construct laws (such as the laws of motion) of steadily increasing explanatory power. Why sneer on the same approach being used for studying our own conscious experience?

Until now we have considered the evolutionary origins of consciousness and tried to answer the question "What is consciousness?" We must now turn to more ancient philosophical issues—like free will and personal identity. In my own contribution (Chapter 9) I have tried to point out that there are really two kinds of personal identity which I have dubbed "empirical identity" and "ontological identity". The empirical identity

question is philosophically trivial and has the form "What criteria do people generally use when trying to *identify* an agent A' as being the same as an agent A whom they have seen in the past?". The ontological identity question (i.e. what criteria *should* be used when trying to decide whether A' is existentially the same as A who lived in the past) is much more important and can be stated in the form of a series of "thought experiments". I have argued that nothing more can be said about personal identity than what is contained in these thought experiments.

FREE WILL – AN EVOLUTIONARY APPROACH

Any theory of consciousness must eventually contend with the problem of free will and determinism. If every event in the universe (including brain events) is the inevitable outcome of preceding events, then in what sense are our actions really free? Of course, if a person were *completely* free his behaviour would be chaotic. Freedom of behaviour (and consequently the will) is necessarily limited by environmental constraints, and hence the question of freedom arises only at what might be called "choice-points", where an agent is called upon to choose between alternate courses of action.

The situation is analogous to a donkey located exactly between two haystacks. Obviously the donkey would not starve to death. He would eventually move towards one haystack or the other, and one would be tempted to describe his choice as being random. A human being in a similar situation might claim that he was exercising the privilege of free will.

If there were no special reason for favouring one haystack, the donkey's choice would either (a) depend on a hidden, thermodynamic bias in the immediately preceding state of the animal's nervous system, or (b) be truly random. Such randomness could arise from a magnification of Heisenbergian uncertainty (see Eccles).³ The sequence of events would be identical in a man but an illusion of free will would accompany the events. Two questions arise. Firstly, why are events at choice-points accompanied by the subjective feeling of free will? And secondly, is there any sense in which an agent's behaviour may be said to be truly free?

Why human behaviour at choice-points is accompanied by the subjective sensation of "willing" is difficult to answer. I do not get this feeling (even at choice-points) if my behaviour is triggered off by (say) an epileptic fit. So

the presence of intervening variables, as opposed to a straightforward S-R sequence, and knowledge (or belief) that I could have acted otherwise are both necessary conditions for claiming to have chosen freely. Further I must be aware of the outcome of my action and must *intend* that outcome. (For there can be irrelevant consequences of my action which I am aware of but do not intend – see Kenny.⁴)

The criteria specified above are mainly self-testimonial. Further, they would be possible only in a nervous system that was capable of projecting itself into the future to anticipate consequences of different kinds of simulated behaviour. (Hence our donkey could not have acted freely.) The system could then use feedback from such anticipations to make what one could call a decision – based on certain goal criteria. If the anticipated consequences are the same for either of two kinds of behaviour then an element of randomness may be deliberately introduced to break the deadlock.

Thus free will seems logically possible only in situations where the outcomes of two anticipated courses of action are *equally* desirable (e.g. choosing between two identical peanuts – where all of Kenny's criteria would be satisfied). Yet, oddly enough, it is precisely in situations like this that a person often declines having chosen freely and says: "My choice was not based on any particular reason – it was random. . . ." One is almost tempted to conclude that free will exists only among philosophers!

We experience willing even in situations where one choice is clearly preferable to the other. The fact that rational considerations lead to one choice and not the other does not seem to be incompatible with feeling free (i.e. feeling that we could have acted otherwise). The sense of choosing freely seems to parallel closely the activity of the system in the brain that is involved in assessing priorities of action in the light of certain goal criteria. Actions uncoupled from this system (e.g. automatisms) are not "willed". Why the activity of this system should be accompanied by a feeling of conscious choice is a mystery, but we can speculate on its biological origins.

Perhaps belief in free will provides the drive or incentive to explore various strategies of action by turning and tossing over ideas in one's mind (just as hunger provides the drive for exploring one's *physical* environment). A drive of this kind would discourage passive acceptance of environmental constraints – and would therefore have obvious survival value. What demarcates Jean-Paul Sartre from *Homo habilis* may be free will rather than language or consciousness!

This analogy between hunger and free will may not be as superficial as it sounds. Consider a hypothetical organism living in an environment where food is always available in plenty. Such an organism would eat and excrete in a continuous and uninterrupted cycle and would never *need* to feel hungry. Hunger must have evolved as part of a control system to regulate the state of nutrition of the animal, when food supply became scarce and intermittent. A fall in blood sugar generates hunger and this in turn goads the animal on to look for food. Consistent with this argument is the fact that carnivores probably experience more intense hunger than herbivores, and plants and trees do not feel hungry at all.

Now in my view, just as the conscious sensation of hunger leads us to explore the environment around us, the inner feeling of freedom goads us on to explore strategies of action in an imaginary world which we construct in our minds. We then see ourselves as active agents striving to *do* things in this imaginary world; and this is possible only because we *feel* free.

Consider a fatalist who feels a sense of inevitability about his own future. To him all actions would seem futile and pointless. In extreme cases, such individuals are often profoundly depressed since they feel they have “lost control” over themselves. Conscious beings need to feel free in order to justify planning for the future and even to justify their very existence. As Sartre would put it, we need to believe in the permanent possibility of consciousness “. . . effecting a rupture with its own past, of wrenching itself away from its past . . .”. So, if consciousness is “nature’s joke” to chain us to each other (Barlow, Chapter 5), free will may be nature’s joke to permit human beings to plan their own future without feeling like puppets in a Laplacian world.

An animal will work only for a tangible reward that lies well within his reach. What characterizes all human actions, on the other hand, seems to be the willingness to participate in what Bronowski⁵ has called “unbounded plans”. Instead of going through a specific sequence of steps leading to a reward, we often adopt global strategies of action directed towards more general aims which we call values or ideals. This ceaseless striving towards abstract and sometimes even unattainable goals (such as “truth” or “perfection”) may also depend crucially on our belief in our freedom. Free will may therefore turn out to be a biologically useful delusion that has been built into our brains by natural selection, i.e. those who believed in their ability to will survived and those who did not died out.