

In *Consciousness and the Physical World*, ed. B. Josephson and V. Ramachandran, pp. 57-75, Pergamon, Oxford, 1980

CHAPTER 4

Nature's Psychologists*

N. K. HUMPHREY

University of Cambridge

On the temple at Delphi was written the stern message "Know thyself". Did the oracle realize she was uttering an evolutionary imperative? I shall argue presently that self-knowledge, and through it the possibility of "intuitive" knowledge of others, has made an essential contribution to the biological fitness of man and certain other social animals. The *means* to self-knowledge have consequently been promoted and perfected by selection. Within this argument lies a theory of the evolution of consciousness; within it, too, lie some humbler ideas about the evolution of overt behaviour.

In *The Nature of Explanation*¹ Kenneth Craik outlined an "Hypothesis on the nature of thought", proposing that "the nervous system is ... a calculating machine capable of modelling or paralleling external events. ... If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the future, and in every way to react in a much fuller, safer and more competent manner to the emergencies which face it." The notion of a "mental model of reality" has become in the years since so widely accepted that it has grown to be almost a cliché of experimental psychology. And like other clichés its meaning is no longer called in question. From the outset Craik's

¹This paper is based on the Lister Lecture delivered at the B.A.A.S. meeting, September 1977.

"hypothesis" begged some fundamental questions: A model of *reality*? What reality? Whose reality?

My dog and I live in the same house. Do we share the same "reality"? Certainly we share the same physical environment, and most aspects of that physical environment are probably as real for one of us as for the other. Maybe our realities differ only in the trivial sense that we each know a few things about the house that the other does not - the dog (having a better nose than I) knows better the smell of the carpet, I (having a better pair of eyes) know better the colour of the curtain. Now, suppose my dog chews up the gas bill which is lying on the mat by the door. Is the reality of that event the same for him as me? Something real enough has happened for us both, and the same piece of paper is involved. The dog hangs his head in contrition. Is he contrite because he has chewed up the gas bill? What does a dog know about *gas bills*? Gas bills are an important part of my external reality, but they are surely none of his.

If mine and the dog's realities differ in this and other more important ways they do so because we have learned to conceptualize the world on different lines. To the dog paper is paper, to me it is newspaper or lavatory paper or greaseproof paper or a letter from my friend. These ways of looking at paper are essentially human ways, conditioned of course by culture, but a culture which is a product of a specifically human nature. I and the dog are involved with different aspects of reality because, at bottom, we are *biologically adapted* to lead different kinds of lives.

To all biological intents and purposes the portion of reality which matters to any particular animal is that portion of which it must have a working knowledge in the interests of its own survival. Because animals differ in their life-styles they face different kinds of "emergencies" and they must therefore have different kinds of knowledge if they are to react in the full, safe, competent manner which Craik - and natural selection - recommends.

But different kinds of knowledge entail different ways of knowing. In so far as animals are biologically adapted to deal specifically with their own portions of reality, so must their nervous "calculating machines" be adapted to construct very different kinds of models. This is not to say merely that the calculating machines may be required to do different kinds of sums, but rather that they may have to work according to quite different heuristic principles. Depending on the job for which Nature has designed

them the nervous systems will differ in the kind of *concepts* they employ, the *logical calculus* they use, the *laws of causation* they assume, and so on. They will differ in what may properly be called their "ideologies". Ideology, in the sense I use the term, means simply a framework of ideas. Ideologies provide, if you like, the "conceptual language" in terms of which questions are asked, calculations made and answers given.

Let us call these nervous calculating machines "minds". It is the thesis of this paper that a revolutionary advance in the evolution of mind occurred when, for certain social animals, a new set of heuristic principles was devised to cope with the pressing need to model a special section of reality - the reality comprised by the behaviour of other kindred animals. The trick which Nature came up with was *introspection*; it proved possible for an individual to develop a model of the behaviour of others by reasoning by analogy from his own case, the facts of his own case being revealed to him through "examination of the contents of consciousness".

For man and other animals which live in complex social groups reality is in larger measure a "social reality". No other class of environmental objects approaches in biological significance those living bodies which constitute for a social animal its companions, playmates, rivals, teachers, foes. It depends on the bodies of other conspecific animals not merely for its immediate sustenance in infancy and its sexual fulfilment as an adult, but in one way or another for the success (or failure) of almost every enterprise it undertakes. In these circumstances the ability to model the behaviour of others in the social group has paramount survival value.

I have argued in more detail before now that the modelling of other animals' behaviour is not only the most important but also the most difficult task to which social animals must turn their minds.² In retrospect I do not think I took my own case seriously enough. The task of modelling behaviour does indeed demand formidable intellectual skill - social animals have evolved for that reason to be the most intelligent of animals - but intelligence alone is not enough. If a social animal is to become - as it must become - one of "Nature's psychologists" it must somehow come up with the appropriate ideology for doing psychology; it must develop a fitting set of concepts and a fitting logic for dealing with a unique and uniquely elusive portion of reality.

The difficulties that arise from working with an *inappropriate* ideology are well enough illustrated by the history of the science of experimental

psychology. For upwards of a hundred years academic psychologists have been attempting, by the "objective" methods of the physical sciences, to acquire precisely the kind of knowledge of behaviour which every social animal must have in order to survive. In so far as these psychologists have been strict "behaviourists" they have gone about their task as if they were studying the behaviour of billiard balls, basing their theoretical models entirely on concepts to which they could easily give public definition. And in so far as they have been strict behaviourists they have made slow progress. They have been held up again and again by their failure to develop a sufficiently rich or relevant framework of ideas. Concepts such as "habit strength", "drive", or "reinforcement", for all their objectivity are hopelessly inadequate to the task of modelling the subtleties of real behaviour. Indeed, I venture to suggest that if a rat's knowledge of the behaviour of other rats were to be limited to everything which behaviourists have discovered about rats to date, the rat would show so little understanding of its fellows that it would bungle disastrously every social interaction it engaged in; the prospects for a man similarly constrained would be still more dismal. And yet, as professional scientists, behaviourists have always had enormous advantages over an individual animal, being able to do controlled experiments, to subject their data to sophisticated statistical analysis, and above all to share the knowledge recorded in the scientific literature. By contrast, an animal in nature has only its own experience to go on, its own memory to record it and its own brief lifetime to acquire it. "Behaviourism" as a philosophy for the *natural* science of psychology could not, and presumably does not, fit the bill.

Chomsky in his famous review of Skinner's *Verbal Behavior*³ argued on parallel lines that it would be impossible for a child to acquire an understanding of human spoken language if all the child had at its disposal was a clever brain with which to make an unprejudiced analysis of public utterances. Chomsky's way round the problem was to propose that the child's brain is not in fact unprejudiced: the child is born with an innate knowledge of transformational grammar, and this knowledge of the grammar provides it, in my terms, with the ideology for modelling human language. Though there are snags about Chomsky's thesis, it would not, I suppose, be wholly unreasonable to suggest something similar with regard to the acquisition of a model of behaviour: the essential rules

and concepts for understanding behaviour might simply be innately given to a social animal. There is, however, an alternative, and to my mind more attractive, possibility. This is to suggest that the animal has access not to "innate knowledge" but to "inside evidence" about behaviour. Nature's psychologists succeed where academic psychologists have failed because the former make free use of introspection.

Let us consider how introspection works. I shall write these paragraphs from the position of a reflective conscious human being, on the assumption that other human beings will understand me. First let me distinguish two separate meanings of what may be called "self-observation", a weak one and a strong one. In the weak sense self-observation means simply observing my own body as opposed to someone else's. It is bound to be true that my body is the example of a human body which is far the most familiar to me. Thus even if I could only observe my behaviour through "objective" eyes it is likely that I would draw on self-observation for most of my evidence about how a human being behaves (in the same way that a physicist who carried a billiard ball about in his pocket might well use that "personal" billiard ball as the paradigm of billiard balls in general).

But the importance of self-observation does not stop there. In the strong sense of the term self-observation means a special sort of observation to which I and I alone am privileged. When I reflect on my own behaviour I become aware not only of the external facts about my actions but of a conscious presence, "I", which "wills" those actions. This "I" has *reasons* for the things it wills. The reasons are various kinds of "feeling" - "sensations", "emotions", "memories", "desires". "I want to eat because I am hungry", "I intend to go to bed because I am tired", "I refuse to move because I am in pain". Moreover, experience tells me that the feelings themselves are caused by certain things which happen to my body in the outside world. "I am hungry because my body has been without food", "I am in pain because my foot has trodden on a thorn". It so happens (as I soon discover) that several sorts of happening may cause a particular feeling and that a particular feeling may be responsible for my willing several sorts of action. The role of a feeling in the model I develop of my own behaviour becomes, therefore, that of what psychologists have called an "intervening variable", bridging the causal gap between a set of antecedent circumstances and a

set of subsequent actions-between what happens to "me" and what "I" do.

Now, when I come to the task of modelling the behaviour of another man, I naturally assume that he operates on the same principles that I do. I assume that within him too there is a conscious "I" and that his "I" has feelings which are the reasons for "his" willing certain actions. In other words I expect the relation between what happens to his body and what he does to have the same causal structure - a structure premised on the same intervening variables - as I have discovered for myself. It is my familiarity with this causal structure and these variables which provides me with the all-important ideological framework for doing natural psychology.

Without introspection to guide me, the task of deciphering the behaviour of fellow men would be quite beyond my powers. I should be like a poor cryptographer attempting to decipher a text which was written in a totally unfamiliar language. Michael Ventris could crack the code of Linear B because he guessed in advance that the language of the text was Greek; although the alphabet was strange to him he reckoned - correctly - that he knew the syntax and vocabulary of the underlying message. Linear A remains to this day a mystery because no one knows what language it is written in. In so far as we are conscious human beings we all guess in advance the "language" of other men's behaviour.

But it may be objected that I have not really made out a case for there being any unique advantage in using introspection since non-introspective psychological scientists do in fact also allow themselves to postulate certain intervening variables such as "hunger" and "fear". And so they do. But think of how they derive them. To establish what variables are likely to prove useful to their models they must (assuming they do not cheat) make a vast and impartial survey of all the circumstances and all the actions of an animal and then subject their data to statistical factor analysis. In practice, of course, they usually do cheat by restricting their data to a few "relevant" parameters - relevance being decided on the basis of an *intuitive* guess. But even so their task is not an easy one. Before postulating even such an "obvious" variable as hunger the experimental psychologist must go through a formidable exercise in data collection and statistical cross-correlation (cf. Hinde).⁴ An ordinary introspective human being has, however, no such problem in devising a "psychological" model of

his own and other men's behaviour: he knows from his own internal feelings what intervening variables to go for. Indeed he knows of subtle feelings which no amount of objective data crunching is likely to reveal as useful postulates. Speaking again for myself, I know of feelings of awe, of guilt, of jealousy, of irritation, of hope, of being in love, all of which have a place in my model of how other men behave.

Before I can attribute such feelings to others I must, it seems, myself have had them - a proviso which the academic psychologist is spared. But it is generally the case, for reasons I shall come to in a moment, that in the course of their lives most people do have most of them, and often indeed it takes only a single seminal experience to add a new dimension to one's behavioural model. Let a celibate monk just once make love to a woman and he would be surprised how much better he would understand the Song of Solomon; but let him, like an academic psychologist, observe twenty couples in the park and he would not be that much wiser:

A garden inclosed is my sister, my spouse; a spring shut up, a fountain sealed. Thy plants are an orchard of pomegranates, with pleasant fruits.... Let my beloved come into his garden, and eat his pleasant fruits. ... I sleep, but my heart waketh; it is the voice of my beloved that knocketh, saying Open to me my sister, my love, my dove. ... My beloved put in his hand by the hole of the door, and my bowels were moved for him.

The translators of the King James Bible, who summarized these lines of the Song as: "Christ setteth forth the graces of the church; the church prayeth to be made fit for his presence" were themselves perhaps somewhat restricted in their ideological perspective.

People are I think well aware of the value of novel experiences in "broadening" their minds. I admit, *pace* my last example, that mind-broadening is not the usual motive which lies behind people's first experiments in making love; carnal knowledge, so called, has intrinsic attractions over and above the insight it may give into what the psalmist meant by an orchard of pomegranates. But there are times when people do apparently seek new experiences for no other reason than to help themselves "make sense", through introspection, of the behaviour of other people. The clearest cases are those where someone deliberately undergoes an unpleasant experience in order to gain insight into the associated state of mind. My mother once discovered that my young sister had swallowed twenty plumstones, whereupon she herself swallowed

thirty plumstones in order, she said, that she should be able to understand my sister's symptoms. My father, in the days when he was politically active, deprived himself of food for a week in order that he should know what it feels like to be a starving peasant. A colleague of mine, studying a tribe of Amazonian Indians, joined the Indians in drinking a strongly emetic and hallucinogenic drug in order that, having experienced the sickness and the visions, he should be better placed to interpret the Indian's behaviour. I could multiply examples, and so I am sure could you.

These acts of calculated self-instruction have, however, a rather artificial ring to them. They are the acts of "intellectuals", hardly to be expected of ordinary people, let alone of ordinary infra-human social animals. Yet every one of Nature's psychologists, if they are to make good use of the possibilities of introspection, must somehow or another acquire a broad base of inner experience to which they can refer. Had they but time, they might perhaps hope to pick up the requisite ideas simply by waiting passively for relevant experiences to come their way. Sooner or later, without seeking it, most animals will no doubt find that they have, say, run short of food or been beaten in a fight or had a narrow escape from danger; they may even - if they are lucky (or unlucky, depending on how you look at it) - find that they have accidentally swallowed twenty plumstones. But what if the experience comes later rather than sooner? The costs of naivety are likely to be heavy in terms of psychological misunderstanding.

The matter is so serious that it would be surprising if it had been neglected by natural selection in the course of evolution. I believe that biological mechanisms have in fact evolved for ensuring that young animals, like it or not, rapidly receive the ideological instruction required to turn them into competent psychologists. They fall into three categories: (i) play, (ii) parental manipulation, (iii) dreaming.

The role of play in extending inner experience is so obvious as to need little elaboration. For all animals, and not just man, play involves adventures for the mind as well as for the body. If we could ask a young animal, as we can ask a child, why it is doing whatever it is doing in play, it would probably reply that it is simply "having fun": but in the course of having fun the animal is unwittingly educating itself. It is throwing itself into new kinds of interaction with the physical and social world and thereby introducing its mind to a whole new range of feelings - new sensations, new

emotions, new desires. Look at a child playing hide-and-seek, or look at a young monkey playing king of the castle: feelings of anxiety, of excitement, of satisfaction, of disappointment, of competitiveness, even perhaps of compassion; these and many other rarer and often unnamable ideas are being planted and tended in the youngsters' minds. One day, when the games are for real, the child or the monkey will use its introspective knowledge of such feelings to interpret and predict the behaviour of another member of its social group.

There are, however, limits to the range of feelings which animals are likely to learn about through play. They play because it pleases them to do so. How then shall they learn about the feelings associated with experiences which are in no way pleasurable? Many of the feelings most pertinent to the modelling of the behaviour of others in the social group are in one way or another disagreeable to the animal who has them - fear, anger, pain, jealousy, grief.... But these are the very feelings which a young animal, left to itself, is likely to do its best to avoid. If play, on the whole, plants pleasant flowers in the garden of a child's mind, what - or who - plants the tares and weeds?

My answer may surprise you. I think that, often enough, it is the child's parents. Biologically it is in the interests of parents to increase the fitness of their offspring in whatever ways they can. Ethologists have long recognized that this is the reason why parents so often take a hand in their children's education, giving them lessons in how to do things and, of course, being active partners in their play. But there has been very little discussion of how parents might help their children by abusing them "for their own good". Let me illustrate the principle with a happening I witnessed not long ago on the train to Cambridge. A woman sat opposite me in the carriage with her 4-year-old daughter. The little girl asked her mother an innocent question. The mother pretended not to notice her. The girl repeated her question, adding plaintively "Mummy, please tell me". "I'm not your mummy", said the woman, "Your mummy got off at the last station". The girl began to look anxious. "You are my mummy. I know you're my mummy." "No I'm not. I've never seen you before." And so this strange game, if you can call it such, continued until eventually the bewildered little girl broke down in tears. A wicked, heartless mother? I thought so at the time - but maybe it was an unfair judgement. That little girl was in the truest sense being taught a lesson, the lesson of what it

feels like to be mystified and scared. She perhaps learned more of real importance in those few unhappy minutes than I myself have ever learned from the hundred books I have read on train journeys.

Now I believe such parental abuse of children may be much more widespread than ethologists have either noticed or perhaps cared to admit. And, following my present line of argument, I believe that its biological function may often be to educate children in the knowledge of disagreeable feelings. Children, as apprentice psychologists, need to know about being frightened, so parents frighten them; they need to know about jealousy, so parents do things to make them jealous; they need to know about pain, so parents hurt them; they need to know about feeling guilty, so parents contrive to catch them doing wrong. And so on. If you were to press me for further specific examples, I should probably continue to refer chiefly to the actions of human beings. But there is one general category of parental abuse which is well known to occur in other social animals than man. That is the "parent-offspring conflict" which occurs in relation to weaning. There are, of course, alternative theories of why mothers become progressively more hard-hearted to their children around the time of weaning, but I would suggest that at least one of the functions of the mother's behaviour is purely educational - it is in the child's best interests that it should have first-hand experience of frustration, rejection, hunger and loneliness.

The third way by which young animals may acquire their ideological grounding as psychologists is by exposing themselves to purely imaginary experiences. I mean by *dreaming*. Dream experience is clearly in a different class to the experience provided by play or parental manipulation; yet I would argue that as a means of introducing the animal to a range of novel feelings it is potentially as powerful. True, there may seem at first sight to be a fundamental problem here: whereas through play or parental manipulation *real* things happen to the infant animal and *real* feelings are aroused, in dreams *unreal* things happen and, presumably, *unreal* feelings are aroused. But it is a mistake to talk of "unreal" feelings. All feelings, whatever context they occur in, are internal creations of the subject's mind. Although they may be - and usually are - evoked by external happenings, it is not the external happenings as such which evoke them, but the subject's perception of and belief in those external happenings. For a feeling to occur it is a *sufficient* condition that the subject should

have the appropriate perceptions and beliefs - that he should "think" himself to be undergoing the relevant experience. Thus for me to feel fear it is sufficient that I should think I am being chased by a crocodile: my fear will be the same whether the crocodile is an objective physical crocodile or a subjective crocodile conjured up in my imagination.

If you yourself have never dreamed of being chased by a crocodile, or if - as I hardly think likely - you doubt altogether the possibility of feelings being induced by fantasy experience, go and watch a stage hypnotist at work. Better still, go up on the stage and allow him to use you as one of his subjects: the hypnotist will, perhaps, suggest that there is a spider crawling up your neck and you will find yourself shuddering with genuine horror.

What the hypnotist does to his subjects on the stage the dreamer can do to himself as the subject of his self-generated fantasies. In the freedom of the dream he can invent extraordinary stories about what is happening to his own person and so, responding to these happenings as if to the real thing, he discovers new realms of inner experience. If I may speak from my own case, I have in my dreams placed myself in situations which have induced in my mind feelings of terror and grief and passion and pleasure of a kind and intensity which I have not known in real life. If I *did* now experience these feelings in real life I should recognize them as familiar; more important, if I were to come across someone else undergoing what I went through in my dream I should be able to guess what he was feeling and so be able to model his behaviour.

Although I have been talking now more of people than of other social animals, I have intended that most of what I have said should apply to animals as well. In people, and people alone, however, the biological mechanisms for providing ideological instruction have been supplemented in important ways by *culture*. All three mechanisms - play, parental manipulation and dreaming - have parallels in human cultural institutions. The play of individual animals has its counterpart in organized games and sports where youngsters, besides enjoying themselves, are encouraged to compete, co-operate, take risks, set their hearts on winning, and discover what it means to lose. Abuse by individual parents has its counterpart in "initiation rites" where adolescents are frequently subjected to bodily mutilation, to fearsome ordeals, and sometimes to forced isolation from the social group. And dreaming has its counterpart in drama and

public story-telling where the actors - and their audience too - get drawn into elaborate fantasies. I am suggesting not merely an analogy but a functional homology between the cultural and the biological phenomena. I believe it could be shown that members of a society who have, for example, been put through a brutal initiation ceremony make better introspective psychologists than others who lack the experience. At another extreme I believe that nineteenth-century readers of Dickens's serial novel *The Old Curiosity Shop*, who cried in the streets when they heard of the death of Little Nell, may have been better able to understand the behaviour of their neighbours when a real child died.

I do not for a moment mean to say that this is *all* there is to these cultural institutions, any more than a sociobiologist would say that the avoidance of inbreeding is all there is to the incest taboo. But if, as I have argued, greater insight into other people's behaviour is one of the benefits of subscribing to a cultural institution, then almost certainly it is one of the factors which keeps that institution alive.

So much for how I think that Nature's psychologists proceed. Let me turn to more purely philosophical implications of the theory. I promised at the start of this paper to say something about the evolution of *consciousness*.

I take it to be the case that what we mean by someone's conscious experience is the set of subjective feelings which, at any one time, are available to introspection, i.e. the sensations, emotions, volitions, etc., that I have talked of. Our criterion for judging that someone else is conscious is that we should have grounds for believing that he has subjective *reasons* for his actions - that he is eating an apple because he *feels hungry*, or that he is raising his arm because he *wants* to. If we had grounds for believing that a dog had similar subjective reasons for its actions we should want to say the dog was conscious too. In proposing a theory about the biological function of introspection I am therefore proposing a theory about the biological function of consciousness. And the implications of this theory are by no means trivial. If consciousness has evolved as a biological adaptation for doing introspective psychology, then the presence or absence of consciousness in animals of different species will depend on whether or not they need to be able to understand the behaviour of other animals in a social group. Wolves and chimpanzees and elephants, which all go in for complex social interactions, are probably all conscious; frogs and snails and codfish are probably not.

There may be philosophers who protest that it is nonsense to talk of a biological "function" for consciousness when, so Wittgenstein tells us, conscious experience does not even have a "place in the language game".⁵ But what Wittgenstein demonstrated is that there are logical problems about the *communication* of conscious experience - and it is not proposed by the theory that consciousness had any direct role in communication between individuals; I am not saying that social animals either can or should *report* their subjective feelings to each other. The advantage to an animal of being conscious lies in the purely private use it makes of conscious experience as a means of developing an ideology which helps it to model another animal's behaviour. It need make no difference at all whether the other animal is actually experiencing the feelings with which it is being credited; all that matters is that its behaviour should be understandable on the assumption that such feelings provide the reasons for its actions. Thus for all I know no man other than myself has ever experienced a feeling corresponding to my own feeling of hunger; the fact remains that the concept of hunger, derived from my own experience, helps me to understand other men's eating behaviour. Indeed, if we assume that the first animal in history to have any sort of introspective consciousness occurred as a chance variant in an otherwise unconscious population, the selective advantage which consciousness gave that animal *must* have been independent of consciousness in others. It follows, *a fortiori*, that the selective advantage of consciousness can never have depended on one animal's conscious experience being the "same" as another's.⁶

Maybe this sounds paradoxical. Indeed, if it does not sound a little paradoxical I should be worried. For I assume that you are as naturally inclined as any other introspective animals to project your conscious feelings onto others. The suggestion that you may be wrong to do so, or at least that it does not matter whether you are right or wrong, does I hope arouse a certain Adamite resistance in you. But allow me to elaborate the argument.

I think no one of us would object to the claim that a piece of magnetized iron lacks consciousness. Suppose now that an animal - let us call it one of "Nature's physicists" - wanted to model the behaviour of magnets. I can conceive that it might be helpful to that animal to think of the north pole of a magnet as having *a desire* to approach a south pole. Then, if the concept of having a desire was one which the animal knew about from its

own inner experience, I should want to argue that introspective consciousness was an aid to the animal in doing physics. The fact that the animal would almost certainly be *incorrect* in attributing feelings of desire to magnets would be irrelevant to whether or not the attribution was heuristically helpful to it in developing a conceptual model of how magnets behave. But if this is conceivably true of doing physics, all the more is it true of doing psychology. Notwithstanding the logical possibility that every other human being around me is as unconscious as a piece of iron, my attribution of conscious feelings to them does as a matter of fact help me sort out my observations of their behaviour and develop predictive models.

Ah, you may say, but you are not really saying anything very interesting, since it can only be helpful to attribute feelings to other people - or magnets - in so far as there is *something* about the other person or the magnet which corresponds to what you call a feeling: the attribution of desire to magnets is heuristically valuable if, and only if, there exists in reality an electromagnetic attractive force between a north pole and a south pole, and the attribution of a feeling of hunger to a man is valuable if, and only if, his body is in reality motivated by a particular physiological state. Quite so. But the magnet does not have to know about the electromagnetic force and the man does not, in principle, have to know about the physiological state.

Magnets do not need to do physics. If they did - if their survival as magnets depended on it - perhaps they would be conscious. If volcanoes needed to do geology, and clouds needed to do meteorology, perhaps they would be conscious too.

But the survival of human beings *does* depend on their being able to do psychology. That is why, despite the sophisticated doubts I have just expressed, I do not consider it to be even a biological possibility - let alone do I really believe - that other people are not as fully conscious of the reasons for their actions as I know that I myself am. In the case of frogs and snails and cod, however, my argument leads me to the opposite conclusion. Let me say it again: these non-social animals no more need to do psychology than magnets need to do physics - ergo they could have no use for consciousness.

Somewhere along the evolutionary path which led from fish to chimpanzees a change occurred in the nervous system which transformed

an animal which simply "behaved" into an animal which at the same time informed its mind of the reasons for its behaviour. My guess is that this change involved the evolution of a new brain - a "conscious brain" parallel to the older "executive brain". In the last few years evidence has at last begun to emerge from studies of brain damage in animals and man which makes this kind of speculation meaningful.

To end my paper I want to talk about a monkey called Helen.

In 1966 Helen underwent an operation on her brain in which the visual cortex was almost completely removed. In the months immediately following the operation she acted as if she were blind. But I and Professor Weiskrantz with whom I was working were not convinced that Helen's blindness was as deep and permanent as it appeared. Could it be that her blindness lay not so much in her brain as in her mind? Was her problem that she did not *think* that she could see?

I set to work to persuade her to use her eyes again. Over the course of seven years I coaxed her, played with her, took her for walks in the fields - encouraged her in every way I could to realize her latent potential for vision. And slowly, haltingly, she found her way back from the dark valley into which the operation had plunged her. After seven years her recovery seemed so complete that an innocent observer would have noticed very little wrong with the way she analysed the visual world. She could, for example, run around a room full of furniture picking up currants from the floor, she could reach out and catch a passing fly.⁷

But I continued to have a nagging doubt about what had been achieved: my hunch was that despite her manifest ability Helen remained to the end *unconscious* of her own vision. She never regained what we - you and I - would call the sensations of sight. Do not misunderstand me. I am not suggesting that Helen did not eventually discover that she could after all use her eyes to obtain information about the environment. She was a clever monkey and I have little doubt that, as her training progressed, it began to dawn on her that she was indeed picking up "visual" information from somewhere - and that her eyes had something to do with it. But I do want to suggest that, even if she did come to realize that she could use her eyes to obtain visual information (information, say, about the position of a currant on the floor), she no longer knew how that information came to her: if there was a currant before her eyes she would find that she knew its position but, lacking visual sensation, she no longer saw it as being there.

It is difficult to imagine anything comparable in our own experience. But perhaps the sense we have of the position of parts of our own bodies is not dissimilar. We all accept as a fact that our brains are continuously informed of the topology of the surface of our bodies: when we want to scratch an ear we do not find ourselves scratching an eye; when we clap our hands together there is no danger that our two hands will miss each other. But, for my own part, it is not at all clear *how* this positional information comes to me. If, for example, I close my eyes and introspect on the feelings in my left thumb I cannot identify any sensation to which I can attribute my knowledge of the thumb's position - yet if I reach over with my other hand I shall be able to locate the thumb quite accurately. I "just know", it seems, where my thumb is. And the same goes for other parts of my body. I am inclined therefore to say that at the level of conscious awareness "position sense" is not a sense at all: what I know of the position of parts of my body is "pure perceptual knowledge" - unsubstantiated by sensation.

Now in Helen's case, I want to suggest that the information she obtained through her eyes was likewise "pure knowledge" for which she was aware of no substantive evidence in the form of visual sensations. Helen "just knew" that there was a currant in such-and-such a position on the floor.

This, you may think, is a strange kind of hypothesis - and one which is in principle untestable. Were I to admit the hypothesis to be untestable I should be reneging on the whole argument of this paper. The implication of such an admission would be that the presence or absence of consciousness has no consequences at the level of overt behaviour. And if consciousness does not affect behaviour it cannot, of course, have evolved through natural selection - either in the way I have been arguing or any other. What, then, shall I say? If you have followed me so far you will know my answer. I believe that Helen's lack of visual consciousness would have shown up in the way she herself conceived of the visually guided behaviour of other animals - in the way she did psychology. I shall come back to this in a moment; I think you will be more ready to listen to me if I first refer to some remarkable new evidence from human beings.

In the last few years Weiskrantz and his colleagues at the National Hospital, and other neurologists in different hospitals around the world, have been extending our findings with Helen to human patients.⁸ They have studied cases of what is called "cortical blindness", caused by

extensive destruction of the visual cortex at the back of the brain (very much the same area as was surgically removed in Helen). Patients with this kind of brain damage have been described in most earlier medical literature as being completely blind in large areas of the visual field: the patients themselves will say that they are blind, and in clinical tests, where they are asked to report whether they can see a light in the affected area of the field, their blindness is apparently confirmed. But the clinical tests - and the patients' own opinion - have proved to be deceptive. It has been shown that, while the patients may not *think* that they can see, they are in fact quite capable of using visual information from the blind part of the field if only they can be persuaded to "guess" what it is their eyes are looking at. Thus a patient studied by Weiskrantz, who denied that he could see anything at all in the left half of his visual field, could "guess" the position of an object in this area with considerable accuracy and could also "guess" the object's shape. Weiskrantz, searching for a word to describe this strange phenomenon, has called it "blindsight".

"Blindsight" is what I think Helen had. It is vision without conscious awareness: the visual information comes to the subject in the form of pure knowledge unsubstantiated by visual sensation. The human patient, not surprisingly, believes that he is merely "guessing". What, after all, is a "guess"? It is defined in Chambers's Dictionary as a "judgement or opinion without sufficient evidence or grounds". It takes consciousness to furnish our minds with the sensations which provide "evidence or grounds" for what our senses tell us; just as it takes consciousness to give our mind the subjective feelings which provide "evidence or grounds" for our eating behaviour, or our bad temper, or whatever else we do with the possibility of *insight* into its reasons.

So if Helen lacked such insight into her own vision, how might it have affected her ability to do psychology? I do not think that Helen's particular case is a straightforward one, since Helen was already grown up when she underwent the brain operation and she may well have retained ideas about vision from the time when she could see quite normally. I would rather discuss the hypothetical case of a monkey who has been operated on soon after birth and who therefore has never in its life been conscious of visual sensations. Such a monkey would, I believe, develop the basic capacity to use visual information in much the same way as does any monkey with an intact brain; it would become competent in using its eyes

to judge depth, position, shape, to recognize objects, to find its way around. Indeed, if this monkey were to be observed *in social isolation from other monkeys*, it might not appear to be in any way defective. But ordinary monkeys do not live in social isolation. They interact continuously with other monkeys and their lives are largely ruled by the predictions they make of how these other monkeys will behave. Now, if a monkey is going to predict the behaviour of another, one of the least things it must realize is that the other monkey itself makes use of visual information - that the other monkey too can *see*. And here is the respect in which the monkey whose visual cortex was removed at birth would, I suspect, prove gravely defective. Being blind to the sensations of sight, it would be blind to the *idea* that another monkey can see.

Ordinary monkeys and ordinary people naturally interpret the visually guided behaviour of other animals in terms of their own conscious experience. The idea that other animals too have visual sensations provides them with a ready-made conceptual framework for understanding what it "means" for another animal to use its eyes. But the operated monkey, lacking the conscious sensations, would lack the unifying concept: it would no longer be in the privileged position of an introspective psychologist.

In the days when we were working with Helen, Weiskrantz and I used to muse about how Helen would describe her state if she could speak. If only she could have communicated with us in sign language, what profound philosophical truths might she have been ready to impart? We had only one anxiety: that Helen, dear soul, having spent so long in the University of Cambridge, might have lost her philosophical innocence. If we had signalled to her: "Tell us, Helen, about the nature of consciousness", she might have replied with the final words of Wittgenstein's *Tractatus*: "Whereof one cannot speak, thereof one must be silent." Silence has never formed a good basis for discussion.

Too often in this century philosophers have forbidden the rest of us to speak our minds about the functions and origins of consciousness. They have walled the subject off behind a Maginot line. The defences sometimes look impressive. But biologists, advancing through the Low Countries, should not be afraid to march around them.

REFERENCES

1. K.J.W. Craik, *The Nature of Explanation*, Cambridge, 1943.
2. N.K. Humphrey, The social function of intellect, in P.P.G. Bateson and R.A. Hinde (eds.), *Growing Points in Ethology*, Cambridge, 1976, pp. 303-17.
3. N. Chomsky, Review of B.F. Skinner, *Verbal Behavior*, in *Language*, 35, 26-58, (1959).
4. R.A. Hinde, *Animal Behaviour*, McGraw-Hill, New York, 1970.
5. L. Wittgenstein, *Philosophical Investigations*, Blackwell, Oxford, 1958, Part I, §293.
6. *Ibid.*, §272.
7. N.K. Humphrey, Vision in a monkey without striate cortex: a case study, *Perception*, 3, 241-55 (1974).
8. L. Weiskrantz, E.K. Warrington, M.D. Sanders and J. Marshall, Visual capacity in the hemianopic field following a restricted occipital ablation, *Brain*, 97, 709-28 (1974).

Discussion

RAMACHANDRAN:

You point out that consciousness permits social interaction. I agree that my direct conscious experience of non-neutral (and emotionally coloured) states, such as pain, hunger, sex, etc., does improve my ability to interact effectively with someone experiencing similar states; especially when I assume that the other person is also *conscious* of these states in the same intense way that I am conscious of them. But I do not see how this argument applies to *neutral* states such as elementary sensations (e.g. reds, greens, etc.). How would my knowledge that the other person was *consciously* seeing these (rather than merely reacting to them) influence my behaviour towards him? If a person were consistently to report red when confronted with such and such a wavelength then I can at once begin effective communication with him. It is quite irrelevant to me whether he is actually conscious of it (like I am) or not. If this is true, then why did "redness" emerge into awareness at all instead of "behaviour towards red" remaining a subconscious and neutral event like the pupillary light reflex? It seems to me that what you have given us is a theory of *emotions* rather than a theory of consciousness.

I see a partial answer to some of these questions in your example of the monkey Helen, who was (presumably) not conscious, although her visual behaviour could be restored; but would you like to elaborate? Supposing I met a man whose visual performance was indistinguishable from normal (i.e. an extreme example of the kind of patient reported by Weiskrantz) but who lacked visual *consciousness*. Would this knowledge make any difference to my understanding him or communicating with him? If not, where does your argument stand?

HUMPHREY:

Your question about the function of "neutral states of consciousness" raises problems which, I am bound to say, I have not fully thought through. Certainly the hypothesis I've presented lends itself more readily to explaining why someone should be conscious of affective states (emotions, motives, etc.) than to explaining why they should be conscious of neutral states such as simple auditory or visual sensations. But I did not mean in my paper to side-step the latter issue altogether, and I hope that what I say about "blindsight" does suggest where the answer lies. On pages 73-4 of my paper I do indeed discuss the question which you now put to me: "In what way would someone who lacked visual consciousness (e.g. after removal of the striate cortex) prove biologically defective?" And I answer it by suggesting that, in one respect at least, such a person would prove to be a poor psychologist, because he would find it difficult to conceive that the behaviour of *another* person was guided by what we call "sight" (I don't say that he could never arrive at the concept, but it might well take him a long time to catch on). A parallel of a sort is provided by the difficulty zoologists have had in accepting the existence of "alien" sensory systems, such as the electric sense in fish or the magnetic sense in birds, of which a human being can have no introspective knowledge. More pertinent still, perhaps, is the case of so-called pheromones: it now seems quite probable that human beings are, *without being consciously aware of it*, influenced by chemical signals from other human beings - but the idea of pheromonal communication remains strange to us because (I would argue) we cannot fit it into a conceptual framework informed by our own consciousness. Radical behaviourists did, in the early days, actually attempt to develop models of both human and animal behaviour which, borrowing nothing from human insight, made no reference

to the existence of different sensory "modalities"; ordinary people, however, being disinclined to cut off their intuitive noses to spite their psychological faces, have always made life easier for themselves by relying on the phenomenology of their own conscious experience to generate the (genuinely) useful concepts of "sight", "hearing", "taste" and so on.

RAMACHANDRAN:

Is the distinction between ordinary consciousness and self-consciousness important to your argument?

HUMPHREY:

By ordinary consciousness or "raw consciousness" I mean sensations, desires, etc., existing as primitive mental events. Self-consciousness or reflexive consciousness, on the other hand, involves inward observation of what is happening on the level of raw consciousness: it is thus logically dependent on the existence of raw consciousness, although it might be argued that the converse is not true, i.e. that raw consciousness is not logically dependent on the existence of reflexive consciousness. However, I know of (and can imagine) no reason to suppose that raw consciousness does *as a matter of fact* ever exist without reflexive consciousness: indeed, if raw consciousness were present in a subject who was unable to reflect on it, he could not (by definition) notice it, remember it, think about it or, *a fortiori*, tell any one else about it. Further, I am not convinced that raw consciousness as such has, or could have, any independent biological function; my own view is that raw consciousness probably evolved to provide the substrate for reflexive consciousness.

JOSEPHSON:

While we are discussing reflexive or self-consciousness, it is worth pointing out that according to some people there are two kinds of "self" involved. There is the individual self, which is the accumulation of the individual's own experiences, and the higher or transpersonal self concerned with creative insights and spiritual experience, which have the appearance of coming from a source beyond the individual and being unrelated to memory. While contact with a higher self is usually stated to be an exclusively human experience, possibly behaviour involving insight, as occurs with monkeys, indicates that they too possess this ability to a limited degree.

VESEY:

As you may know, philosophers spend a lot of their time talking about meaning. There are radically opposed views, some with quite a history to them. For instance, there is the empiricist view, held by people like John Locke, J.S. Mill, and, more recently, Bertrand Russell and A. J. Ayer. Roughly, they say that a word has meaning by being a name given to an experience. For instance, someone has a pain, gives the name "pain" to it, and then uses the same word again when he has an experience he recognizes as being similar to the one to which he first gave the name. (That is a one-sentence summary of what Mill says in Book I, Chapter 3, of his *System of Logic*, 1843.) This seems an attractively simple

account of meaning, but there is a problem connected with it. If "pain" is a name I give to one of my experiences, and re-give when I have a similar experience, what can I mean when I say that *someone else is in pain*? It's a bit like knowing what it means to say that it is afternoon, when one is in Houston, Texas, and then being expected to understand the remark when one is half-way to the moon. The conditions of meaningfulness have been removed. There is no zenith for the sun to be past, no horizon for it not to be past. Similarly with talk about someone else being in pain, if one accepts the empiricist account of meaning. The condition of meaningfulness, that the sensation can be recognized as similar to the one first named, no longer holds.

It seems to me that a basic presupposition of your argument is the correctness of the empiricist view of meaning. Do you have a solution to the problem I've indicated?

HUMPHREY:

Let me try to make my argument clearer with an example. Then maybe the problem you raise about meaning will be easier to resolve.

Suppose that each and every one of us owns a whistling kettle, and that it is important to be able to predict the "behaviour" of these kettles (to anticipate their whistling, etc.). The external facts I observe about my own and other people's kettles are, say, of the following kind: (i) the kettle when filled with cold water and put on the stove begins to whistle within about 5 minutes, (ii) the kettle takes less time to whistle when filled with hot water, (iii) the kettle takes more time to whistle when salt is added to the water, (iv) the kettle takes less time to whistle on top of a mountain, (v) if the kettle is filled with liquid nitrogen instead of water it whistles without being put on the stove, (vi) if the kettle is filled with treacle it doesn't whistle at all, and so on.

I suggest that, if these external facts were all I had to go on, the behaviour of the kettles might seem puzzling. I would be hard put to it to develop a *theory* of the relation between what is done to the kettle and what the kettle does. But suppose that, while everybody else's kettle is made of tin, my own kettle is made of Pyrex glass so that I can see into it. I look into my kettle and observe (i) that when certain things are done to the kettle the liquid inside it boils, and (ii) that when the liquid boils the kettle whistles. I am led to regard *boiling* as an explanatory concept, an "intervening variable" which "bridges the causal gap between a set of antecedent circumstances and a set of subsequent actions-between what happens to my kettle and what my kettle does" (cf. my paper, p. 62). Thus I now *explain* the behaviour of my kettle by arguing along the following lines: the kettle whistles when the liquid boils, the liquid boils when the kettle is put on the stove, therefore the kettle whistles when it is put on the stove.

But at this point something philosophically interesting has happened. While the concept of boiling has been put into my mind by *a factual observation* (what I actually saw when I looked into my kettle), its usefulness as an *explanatory concept* does not depend on the observation's having been of any particular kind; indeed, I could have observed something quite different. Suppose, for example, that when I looked into my kettle I had observed the liquid turning a red colour under just those circumstances when in fact I saw it boil, then the concept of reddening might have come to play exactly the same role in my argument as the concept of boiling: the kettle whistles when the liquid reddens, etc. Indeed as far as my new-found theory is concerned it really doesn't matter what I have actually observed (and *a fortiori* it doesn't matter what I choose to *call* what I have observed - I might as well say the liquid in the kettle is in pain).

Now, how about other people's kettles? Since they are made of tin I cannot, of course, *observe* the liquid inside their kettles boiling (or reddening or whatever). Can I then use the concept of boiling to help myself *explain* the behaviour of their kettles? Yes. Since the usefulness of boiling as an explanatory concept is independent of any particular observation I have or could have made, the concept can play just the same role in my argument about someone else's kettle as it does in my argument about my own.

With regard to the problem of meaning, I accept that the *factual* propositions "The liquid in my kettle is boiling" and "The liquid in his kettle is boiling" are of different status (indeed the latter proposition is arguably, by positivist criteria, meaningless). But the *explanatory* propositions "My kettle is whistling because the liquid inside it is boiling" and "His kettle is whistling because the liquid inside it is boiling" are on a par.

Another example to think about: suppose that Mendel, when he was searching for a theory of inheritance, could have observed his own genes.

VESEY:

You are right: your example does make your argument clearer. It makes it clearer that it is as follows. (i) The concept of boiling is put into one's mind by what one observes on looking into kettles. Similarly, (ii) the psychological concepts one uses to explain people's behaviour-concepts like expecting, hoping, remembering, understanding, wanting, wondering-are put into one's mind by what one observes on looking into one's mind (introspecting) when one is doing these things. (iii) That one cannot look into other people's minds does not prevent one using psychological concepts to understand their behaviour.

Not only does your example make your argument clearer; it also enables me to make clear the extent and nature of my disagreement with you. I disagree with you not only about (ii) but also about (i). And the disagreement is a fundamental one, about meaning. To know the meaning of a word (= "to have the concept for which the word stands") is to know how to use the word correctly. A word's being meaningful, and there being criteria of its correct use, go hand in hand. This being so, it does not make sense to talk of concepts being put into people's minds by their observing things, either inner things or outer things. Concepts are not experiences, to be put into people's minds by pointing their eyes, or their mind's eye, in the right direction. They are abilities exercised primarily, in humans, in acts of verbal communication. And the linguistic practices involved could not, even in theory, start as private practices.

JOSEPHSON:

The dilemma can be resolved by assuming that the concepts are already there in latent form in the nervous system, waiting to be triggered off by the relevant experiences. The latter do not have to be linguistic in nature.

BARLOW:

As a result of thinking about the biological role of consciousness both Nick Humphrey and I (see next paper, "Nature's Joke") have come to the same conclusion, namely that the survival value of consciousness is very much connected with its role in the social life of gregarious animals, but there is a difference between our proposals that may be important. I argue that consciousness is impossible without some kind of social interchange, so that

mankind is driven to engage in social relations to preserve his consciousness. Consciousness is thus Nature's tool to make man social, just as pain can be regarded as her tool to make us avoid injury. The survival value of consciousness would result from social hominids leaving more offspring than solitary hominids. If I understand Humphrey correctly, he regards the gregarious nature of man as a prior fact, and sees consciousness as conferring an advantage in competing against other individuals within the same social group. Am I right in understanding him to say that consciousness improves social behaviour, but does not actually help to generate it, as I would claim?

I have another question relating to the use of the word "introspection", for I don't think we find out about others in this way. It is very likely true that you cannot understand certain aspects of other individuals' behaviour until you have yourself undergone the experience motivating that behaviour, and this is interesting and important. But this insight seems to come by a process of imitation rather than introspection, which I take to mean a conscious searching of one's own mind. Sight of a pattern of muscular movements may enable one to imitate them, and I think one's feelings can imitate the emotions that generate a pattern of behaviour in another. But I don't think there is any conscious search in one's mind for them, so I would hesitate to call this process introspection.

HUMPHREY:

1. I hope Barlow will not mind if I characterize his argument as follows. Consciousness is rather like group sex: something which is a source of pleasure to the individual but which he can't achieve on his own and so is obliged to seek through interaction with others. Thus Barlow sees the biological function of consciousness—the contribution it makes to biological survival—as the provision of an *incentive* to being social (sociality being essential to human survival). His argument rests, as I see it, on three premises: (i) people *desire* to be conscious (as, for example, they desire sex); (ii) people can only be conscious through social interaction; (iii) people would not be social if they were not made to be by this "trick" which Nature plays on them. Barlow's question relates to this last point, and he is right to think that I disagree with him here. I do not believe that people remain in social groups in order to preserve their consciousness; my view is that people would, whether conscious or not, *try* to form social groups but that if they were not conscious they would probably fail because they would be unable to understand each other. In Barlow's view, without consciousness the social group would never get together; in my view, without consciousness the social group would fall apart. But either way, surprisingly enough, we draw the same conclusion, namely that consciousness is probably a necessary condition of being a highly social animal. And indeed we agree on a more specific prediction, namely that a dysfunction in the mechanism of consciousness (as I suggest may have occurred in Helen and Barlow suggests may occur in autistic children) is likely to show up as social maladjustment.

2. Barlow has misconstrued my argument if he thinks I'm suggesting that "we find out about others" by introspection. No, we don't "find out" about them that way; we find out about them by ordinary external observation - looking at them, listening to them, etc. What introspection does is to help us *explain* what we find out about them: it provides us with the explanatory concepts in terms of which we "make sense" of what we observe. This point is elaborated in my reply to Professor Vesey. But when, for example, we explain someone else's behaviour by saying "He is crying because he is in pain" we don't have to be feeling the pain ourselves (which is what Barlow seems to be implying by his remarks about "imitation").