Reprinted without change of pagination from the
Proceedings of the Royal Society, B, *volume* 161, *pp.* 376–383, 1965

REVIEW LECTURE

The evolution of mind*

BY D. O. HEBB
McGill University

(*Lecture delivered* 9 *April* 1964—*Received* 30 *June* 1964)

It was Darwin himself who first raised the question of an evolution of behaviour in his *Expression of the emotions in man and animals* in 1872. The book had little direct effect on psychology, which at the time concerned itself solely with normal adult man as a unique being, mentally if not physically. But Darwin's ideas were as revolutionary for psychology as for zoology and could not be escaped indefinitely. In this century psychology has been reluctantly 'biologizing' itself, slowly absorbing the full implications of the idea that man's behaviour (and therefore the mind that controls it) is as much the product of evolution as his erect posture or the structure of his hand.

Revolutionary ideas in science are often accepted slowly, but in this case the resistance was extreme and continued long after Darwin's ideas were accepted in other fields. The reason may be partly that it took long to work out biological or behavioural theory to the point where it was viable, but the reason is partly that here the conflict between theory and common sense was unusually sharp. The situation draws attention to a feature of scientific thought which may be obscured in other fields by the tremendous successes of physics, for example, or biochemistry or genetics. Psychology has given no such dazzling performance, and thus may allow us to see more clearly the nature of scientific thought. It should be added that the problem of scientific thought is a proper professional concern of the psychologist, and the history of science an important source of information about man's thought processes, for it is a record of real problem-solving as contrasted with the more or less artificial problem-solving one may set up for study in the laboratory. Thus in the present discussion I am concerned with the scientific method in itself, as well as with the effort by psychologists to apply it rigorously to the problem of mind.

Biologizing led to two propositions: (*a*) that psychology must be objective—in other words, we do not know our own minds (directly) and must eschew introspection as a method—and (*b*) that dualism must be rejected, just as vitalism is and

* This is the substance, more or less, of a review lecture given before the Society on 9 April 1964. The main omission concerns the motivation of scientific thought, which as others have suggested is not unlike that of the poet, and the importance of its aesthetic characteristics for creative achievement (Hadamard, *The psychology of invention in the mathematical field*, 1945).

Aid in preparing this paper is acknowledged from the U.S. Public Health Service, grant no. MH–02455–06, and from the Defence Research Board of Canada, grant no. 9401–11.

for essentially the same reasons. These propositions seemed, and may still seem, nonsensical. Everyone knows that he has direct knowledge of his own thoughts, and that mind and matter are wholly different things.

However: what weight should one give such common sense, as scientific criticism? Huxley defined science as 'organized common sense' but a better definition would be an organized *attack* on common sense: a wild intellectual adventure, in Bridgman's terms 'doing one's damnedest with one's mind, no holds barred'. If common sense gets in the way, so much the worse for common sense—which if one examines it closely turns out to be only an accumulation of ancient theories, whose origins are forgotten. Like any theory then we might expect it to be partly sound, partly unsound. As far as psychology is concerned, at any rate, though the coinage is worn smooth René Descartes's stamp can still be detected on psychological dualism, and John Locke's on introspection. Descartes and Locke were brilliant men, but they cannot of course be allowed any power of veto on subsequent speculations—as they would have if common sense was allowed to call the tune.

How much weight is given to common sense elsewhere in science? The notions of physics are indubitably successful, but they hardly bear inspection from a common-sense point of view: particles that are waves (some of the time), and waves moreover in a medium that does not exist; particles with discontinuous existence in space or time; and now anti-matter. Physics of course is logical, as a science must be; but not in any common-sense way. Some 30 years ago certain theoretical assumptions about elementary particles were not confirmed by experiment. What to do when theory and data disagree? What we are taught is, discard the theory. But physics said, nonsense, the theory is perfectly good theory, and on the spot invented the neutrino, whose sole *raison d'être* was to make an equation balance and preserve the law of conservation of energy. Add that the neutrino is a particle with no mass, and one gets more of the flavour of this operation.

But it has already been said that such logical operations *work*, are validated by their success. I am not aspersing physics, but showing how little its conceptions have to do with common sense. There are many other examples throughout science, and one must conclude that the thought of the scientist is very different from what we have supposed. Science makes essential, inveterate use of preposterous ideas—at first exposure—to which, when they have been in existence long enough, one becomes habituated so that one can forget how intrinsically improbable they are. Darwin's idea that a one-cell organism could spontaneously elaborate itself into bird, snake or monkey, and the geneticist's current notion that the whole template of man is contained in a few molecules of *DNA*, would be further splendid examples. Barber (1961) provides supporting evidence for the view that science operates with improbable notions, by showing how often ideas that are now unquestioned were at first rejected by scientists of the highest ability: Agassiz wholly unconvinced by Darwin; Helmholtz resisting Planck's ideas; the brilliant conceptions of Mendel rejected by everyone (it is not, as I had thought, that he was unknown: he was just not accepted); Heaviside ignored for 25 years; and Lord Kelvin, finally, unable to accept Maxwell's theory of light or Rutherford's theory of the atom, and regarding Roentgen's X-rays as a hoax.

What I am leading up to is evident: if behaviourism seems preposterous to common sense it may still be on the right path. Now a friend, Professor P. R. Wallace, reminds me that nonsensical ideas may be a necessary but not a sufficient condition and the fact that psychologists are crazy is not enough to show that they are good scientists. Conceding that point, I must next ask what evidence there is of value in the behaviouristic approach.

Biological theory of mind

So, where do the biological assumptions of modern psychology lead us? It is clear, first, that the prime function of the nervous system is to serve as a communications network. When sensory and motor cells first became specialized in the primitive multicellular animal they necessarily became separated in space. Light-sensitive cells to warn of the approach of a predator are of no value unless the word gets to muscle cells at some distance, so flight can occur; development of a sensitive nose, detecting food at greater and greater distances, is useless if the legs are not excited into motion and guided in the right direction. This is the first and most important role of nerve cells, to establish a reflexive sensory-motor communication: and not only in the primitive animal but in the most highly evolved, for the delicate adjustments of reflex function are essential moment by moment to the life of each one of us.

But what of other, 'higher', functions? If the nerve cell is simply a conductor of excitations from one point in space to another, how can it account for mind, thought, consciousness? The classical answer is, it cannot: an answer well expressed by Sir Charles Sherrington (1941). Until about 1940 the central nervous system was regarded by everyone, except the great Ramón y Cajal and his pupil Lorente de Nó, as an essentially reflex system, doing nothing but connect receptor with effector, directly or indirectly. Learning was of course possible but the pathways of learning through the cortex were thought of as alternative or additional routes, for making more connexions of the same kind. If this was so, the conclusion was inevitable that some other agent, some non-neural principle, must account for those higher functions of man on whose existence we are all agreed.

Now evolution certainly produced the higher animal's reflex function, a marvellously sensitive, but automatic and limited adjustment to the environment. In the course of doing so, however, it seems to have produced something else. The higher levels of the brain are not merely a collection of alternative sensory-motor paths, one-way streets, or in–out connexions like those of an old-fashioned telephone switchboard (which, we may note, did require an 'other agent'—an operator—different in kind from the switchboard itself). The brain instead is full of anatomically demonstrated closed circuits which must have some other function than direct sensory-motor connexion. These central circuits must be self-re-exciting and self-modifying. Instead of a direct transmitter of information, this is a system which can hold a sensory message by allowing it to travel round and round in closed circuits; can re-order its components in time when transmitting it; can suppress one component and replace it with another held over from a former message: any or all of these. It may then suppress the whole thing, or may transmit the modified

information to the muscles in a form that means a new response to the environment, with a large element of unpredictability about its operation.

In a reflex system, the same message leads to the same response time after time; a reflexive organism is sense-dominated, an automaton fully controlled by the environment (for any given state of the body humours). In theory at least, the closed circuits in the brain of a higher animal can detach him from such control. What the animal does depends not on the present sensory input but on a synthesis of this input with the prior activity of the central circuits.

Biological theory equates that central activity with thought; the transformations that permit new ways of responding to the environment then become creativity, and the capacity to withhold response or not (and to choose between different modes of response) is free will. Whether this approach is ultimately found satisfactory or not, there is no basis yet for saying that brain function cannot account for this or that feature of man's nature. The new theoretical possibilities opened up by neuroanatomy and neurophysiology are such that we have hardly begun to explore them.

It is worth observing that behaviourism as developed by Watson was the direct descendant of British associationism: the latter a narrow theory if there ever was one, and Watson could not have been rejected in Britain because he too was narrow. As Humphrey (1951) has shown, apart from the question of dualism every criticism directed at behaviourism was a repetition of attacks, such as Bradley's, on associationism. We can now see that both theories were as wrong as could be, but we must see also that this was the fertile line of thought. It was the earlier clarification achieved by subjective associationism that made Watson's objective theory possible; and the advances in knowledge made in this century have been consistently related to Watson's ideas and the closely related ones of Pavlov and Thorndike—either to show experimentally that they were wrong, or again to defend and develop them. A theory is a tool, and there are times when a 'wrong' theory (in hindsight) can be a more powerful tool than a right one.

In psychology, at any rate, the result of those behaviouristic implausibilities is an access of knowledge and understanding that is independent of the theory. We have now criteria, not very precise but objective and intelligible, for the presence of higher processes—in short, criteria of mind or consciousness. 'Mind' by these standards is not all-or-none, it does not appear suddenly in the phyletic scale, and (since much of man's behaviour is wholly reflexive) it does not enter into all the behaviour of the organism that possesses it. So I do not attempt to say where it appeared in phylogenesis; but the criteria of freedom from sense-dominance and capacity to hold and re-order sensory information make it possible to give objective meaning on the one hand to the proposition that cat or dog or porpoise has a mind (or is conscious); and, on the other, give good reason for denying that the ant is conscious, that the bee has true language (which is purposive), or that the earthworm can suffer pain (which has an emotional component).

Maybe such theory is nonsense, in the true sense. I do not argue for its truth, however, but for its experimental power and its capacity to show us 'new' characteristics of man as a species. The perception of evidence is a function of theoretical

preconceptions, so it need not surprise us that as theory develops we see things we did not see before, even if they were always present and even if theory cannot yet account for them adequately.

Concerning motivation

Consider the nature of man's fundamental motives. Despite the frequency of generous actions in almost any human society, there is a long tradition by which all human behaviour is traced at bottom to selfishness. Kindness is not in the child's nature but imposed by training and maintained at maturity by social pressures. In the same tradition it is taken for granted that man is fundamentally averse to work, mental or physical; when he works, it must be for some extrinsic reward, except when habit gets him in its clutches—then, stupidly, he may keep on working without needing to.

But observation of lower animals tells us another story, on both points. 'Altruism', a disinterested concern for others, has repeatedly been seen in the chimpanzee and I believe can be seen in the dog. There are circumstances in which the laboratory rat will prefer to reach food via a maze problem rather than a shorter, direct route; H. F. Harlow has shown that the monkey will work for hours at solving simple puzzles without extrinsic reward; and the chimpanzee if he likes a learning task will work at it when he is not hungry, rejecting the food reward offered by the experimenter. No animal, including man, is altruistic all the time and some individuals may never be; none the less, altruism is a species characteristic of some of the higher mammals and most marked in man. Similarly, the higher animal is likely to avoid work imposed by others but work of his own choice and at his own pace is characteristically sought, not avoided. 'Play' is a most misleading term; what it refers to is work, physical or mental, that is done for its own sake, and one of the most striking facts about man as a species is the amount of time and effort spent in intellectual as well as physical play.

Seen from this comparative point of view, as the high point in evolutionary development, man is as remarkable for characteristics of motivation and emotion as for his intellect. It is usual to think only of an evolution of 'intelligence' or of learning and problem-solving, but in fact it is difficult to show any steady progression of these abilities as one goes from 'lower' to 'higher' species. An evolution of motivational characteristics, however, becomes evident as soon as one looks for it.

Anger and fear provide good examples. Here man and chimpanzee are close relatives—far closer than in their capacities for problem-solving—with well-developed susceptibility to a great variety of causes of emotion and the same characteristic modes of expression. The young chimpanzee's temper tantrum, for example, is recognized at once by anyone who has brought up human children, and to the human observer there is nothing strange about the acute emotional disturbance shown by an adult chimpanzee who found a live worm in the biscuit she was eating. The text-book discussions of man's emotional characteristics have been extraordinarily myopic, based apparently on observation of man alone, without comparative perspective (and dominated as well by outdated theory). They

used to teach that there is no fear of the dark; if there seems to be, the child must have been conditioned by some untoward event in a dark place—and as for the adult subject, such a thing is never mentioned, from which I can only conclude that writers of text-books have never been alone in the deep woods at night. Loneliness and homesickness are left unmentioned, along with the notorious disturbances at the sight of blood, surgical operations (even when blood flow is fully controlled), or gross mutilations of the human body. Even when the phenomena in the two species are not identical, the chimpanzee's reactions draw attention to the extraordinary range of the causes of emotional upset in man.

Man as we know him in this society is an unemotional being, or at least less emotional than young children or wild animals, but this is only because he is able most of the time to avoid situations that evoke strong emotion and, in what we call civilized society, has created an insulated behavioural environment in which his emotional susceptibilities are well concealed—even from himself. The chimpanzee in *his* preferred environment is also a placid animal. In captivity he is as unpredictably explosive as a fireworks display, sometimes viciously aggressive without cause and capable of being angered by trivial things; terrified at the sight of a toy animal or a model of a human head and greatly disturbed by others of a long list of visually perceived objects; and in these and other ways reminding us of human fears, hostilities and abhorrences, each of which is familiar by itself but which in their totality make a picture of man that we have not seen clearly.

As for theory, two points should be mentioned. First, the correlation of emotionality with phylogenetic level, or presumed level of intelligence, suggests that in emotion we are dealing with some sort of transient breakdown of orderly function, some instability of transmission in the large brain, which can occur more easily as cerebral circuits become more complex. The suggestion gains weight when it is observed that it is the older subject, thus presumably with more complex intellectual processes, who is most easily disturbed: the variety of irrational fears is greater in the adult chimpanzee than in the infant, and it is the older human subject who is upset by snakes, spiders and mice. It is not the young but their elders who are likely to be bothered by horror films on T.V., or tales of giants who eat little children and grind the bones to make their bread.

The second point is *de rigueur* in any such discussion as this. It concerns A.R.A.S., the 'ascending reticular activating system' of the brain stem. It is very popular nowadays, even if the theories such as mine that relate it to human behaviour are still a trifle fanciful. Some degree of activity of A.R.A.S. is necessary to cortical function and to consciousness (it is impaired, for example, in sleeping sickness and in coma following head injury). It has been proposed that there is an optimal level of A.R.A.S. activity; too low a level is boredom, a state in which the subject tends to seek sources of excitement; too high a level is fear or anger or some other emotional state in which the subject tends to act in a way that decreases excitement, by avoiding or forcibly suppressing the excitant. Something of the sort certainly seems to be what happens, though the details are far from clear; and in any case such ideas have the great merit of drawing our attention to another aspect of human motivation. They imply that the same situation that produces avoidance

(when arousal level is too high) will also attract (when arousal is low, and the excitant is not too strong). Such an ambivalence is very marked in the higher animal and most important for understanding certain aspects of human behaviour.

We have already seen that man both seeks and avoids work. Another ambivalence concerns fear. Ski-ing, mountain-climbing and automobile-racing are sports in which, at least some of the time, fear is deliberately courted. (Common-sense theory is sure that fear is always avoided, so we speak instead of seeking 'thrills' or 'adventure'.) Man is empathic, tending to identify with his fellows, so he is often altruistic and disturbed by the danger or pain of others; but some people are capable of getting a thrill from seeing others tortured, and most people are fascinated by tight-rope walkers and trapeze artists, the more so when their acts are performed without a safety net (so the fascination is not merely in seeing an exhibition of skill). The person who might faint in the dissecting room avidly reads details of highway accidents and airplane crashes. Add the ambivalence that lies in avoidance of the obscene and enjoyment of *risqué* jokes, and it becomes possible to justify the generalization that what repels man tends also to attract him, as the above interpretation of A.R.A.S. function suggests.

Conclusion

These are sufficient examples perhaps to make my case. It might possibly have been wiser, in trying to show some validity in 'biologizing' psychology, to stick to details of experimental work. The behaviouristic approach has resulted in a new capacity for fertile experiment, with a large body of factual data whose value is independent of the theory that gave rise to them. There are such studies as Lashley's on the nature of thought and language, Skinner's on the control of learning, or Harlow's on 'learning sets' (learning how to learn), all directly behaviouristic in origin and all helping us to understand the human mind and behaviour. Our understanding of the heredity–environment relation in behaviour (showing that *both* are more important than once was thought) has greatly increased. Perceptual processes are studied directly in man with concepts derived from animal work. And so on. It may be foolhardy to try to establish points of contact between Penfield's ideas about the centrencephalic system of the brain stem and man's behaviour under stress; between what is known about neural circuits or synaptic function and creativity in the scientist; or between the chimpanzee's behaviour and human generosity or human prejudice toward those who differ in language, ideas, or skin colour. But making such attempts, asking such questions, does sometimes lead us to see what we did not see before and may perhaps show us how to ask better questions and plan better experiments.

It is certainly too soon for the experimentalist to attack some of these grandiose problems directly, or to spend all his time thinking about them. The primary business of the psychologist, as of other scientists, is his daily bread-and-butter research, small-scale and often perhaps trivial but which he hopes will add up to something bigger. But it is important also to keep one's eyes open to the possible wider significance of one's research in detail, because doing so may lead one to see how to do that research better. As far as I can see, timidity in dealing with ideas

is contrary to the spirit of the scientific method, and the thinker who is afraid of looking like a fool has tied one hand behind his back before entering the ring. For our present purposes, at any rate, it has been important to show that 'biologizing', and 'behaviourism', eschewing any form of vitalism and denying that self-study is the best way to understand oneself, need not mean closing one's eyes to the subtleties of mind and may even mean that one can see better the full extent of the problem.

References

Barber, B. 1961 *Science*, **134**, 596–602.
Humphrey, G. 1951 *Thinking*. London: Methuen.
Sherrington, Sir Charles 1941 *Man on his nature*. New York: The MacMillan Co., also Cambridge University Press.

PRINTED IN GREAT BRITAIN AT THE UNIVERSITY PRINTING HOUSE, CAMBRIDGE
(BROOKE CRUTCHLEY, UNIVERSITY PRINTER)