S & CB (1989) 1, 27–39 0954–4194

DONALD M. MACKAY

In What Sense can a Computer 'Understand'?

***Computers** can in principle carry out many and perhaps all of the functions of the **brain**. This does not mean that they can think or **understand**. In debate with **behaviourism** and with **J. Searle** it is maintained that brains do not understand. Understanding is something that **agents** do. It is however possible to devise artificial agents embodied in a computer, but it is the agent and not the computer that understands. This gives no reason for claiming that such artificial agents can think, or are conscious centres of awareness as we are.*

Computers were originally aids to human calculation. Human calculation is a mental activity, in which *people think*: people start from certain 'data' and mentally 'work out' what implications those data have under given conditions: 'If 2 is added to 3 it makes 5' and so on. To save time and mental effort we can sometimes consult printed arithmetic tables, so that our only thinking consists in identifying the appropriate entry, and recognizing the meaning of the printed 'answer'. Alternatively we can use a computer: a mechanical device that displays the answer automatically after it has been manipulated in a manner determined by the data and the nature of the implication we want to know.

Computers fall into two fundamentally different classes, according to the type of automatic internal process that leads from the initial manipulation by the user ('specifying the data and the inference required') to the production of the 'answer'. In the familiar 'digital' class, the process is one of *symbol manipulation according to rule*. In the same sense as a Jacquard weaving loom can have the selection of its coloured threads specified according to a rule embodied in the pattern of holes punched in a card, a digital computer can have the selection of the symbolic tokens it manipulates, and the operations it performs, specified according to rules embodied in a similar pattern. In modern machines the pattern is made up of minute and easily altered blobs of stored magnetic or electrical energy, with huge advantages in speed; but the principle of

'manipulation according to rule' is the same in the computer as in an automatic loom or in the distributor of a motorcar engine.

Note that the same principle could also be implemented by using a *human* slave to consult a printed book of rules and do the necessary fetching, exchanging and delivering of symbolic tokens. As was often pointed out in the early days of computing, such a slave need have no comprehension of the meaning of the calculating process he is helping to instantiate. Conversely, if the slave here were replaced by an automatic lockup mechanism with equivalent functions, it would be quite baseless to credit that mechanism with any more understanding of the process of calculation than the slave had.[1] Searle,[2] whose recent Reith lectures innocently recapitulate many of the arguments on this topic that circulated in the 1950s, confuses the issue in this connection by claiming that 'in the *human* sense', computers don't follow rules at all: 'they only act in accord with formal procedures'. Although, as will soon emerge, I agree with Searle that computers don't think, I doubt whether he has found the best way to express the distinction we need. The sense in which an automatic loom 'follows rules' is close enough to the human for a human operator to take the place of the automatic card-reader and follow the same rules. What Searle should have said, I think, is that in a *psychological or mentalistic* sense computers don't follow rules; or better, that the mental process we call 'understanding the meaning of a rule and obeying it' is not embodied in the rule-following operations of a computer—any more than it is in those of a Jacquard loom or a motor engine. If (as above) we replace the automatic manipulative machinery of the computer by a human slave required to consult a book of rules, then of course we do have a mental process of rule-following going on in the slave; but as we saw, this is irrelevant to the cognitive status of the computing operations, which the slave need not understand as such. The slave may be mentally busy thinking (and acting) according to the meaning of the rules in his book; but there is nobody there thinking (working out mentally) the stages of the calculation.

Searle makes a similar point using the example of a room in which a non-Chinese speaker successfully follows syntactic procedures for the manipulation and exchange of Chinese symbols so that Chinese

1 MacKay, D. M. (1965) 'A Mind's Eye View of the Brain'. *Cybernetics of the Nervous System* (Norbert Wiener and J. P. Schade, eds). *Progress in Brain Reseach* **17**, pp. 321–332.
2 Searle, J. (1984) *Minds, Brains and Science*. B.B.C. London.

speakers can engage in meaningful communication with the system he operates, yet he himself understands no Chinese. He 'acts as if' he understands; but the appearance is deceptive. Searle, too, argues that if the manipulative functions of the non-Chinese speaker were successfully taken over by a digital computer, there would be no rational grounds for claiming that the computer understood Chinese. (Searle tries to make the argument sound syllogistic: 'Syntax alone is not sufficient for semantics, and digital computers have syntax alone'. But as we shall see, so many terms here are ill-defined—e.g. what does it mean for a computer to 'have syntax alone'?—that I don't think the formalization helps his claim to have 'simply and decisively refuted' the opposition.)

In the second class of computer, the outcome of the computational process is selected on the basis of a *physical experiment*. For example, if I want to know the sum of 2 and 3 I can pour 2 ml and 3 ml of liquid together into a graduated measuring jar, and read off the final level (5 ml) at the meniscus. This would be a typical 'analog' computing operation, where the physical operation of mixing is *analogous* to the mathematical operation of adding. The link between the physical and the mathematical levels here is on the process of *measurement*, which is normally a conscious mental process on the part of the user of the computer; but in many applications this too can be automated so that an appropriate numerical symbol is illuminated at the end.

I think it is clear that there are no more convincing arguments for attributing mental activity to such an analog device than in the digital case. If I use a mixing jar to indicate to me the sum of 2 + 3 it is in order that I need not do the arithmetical thinking, but not in order that it (or anyone) should 'do my thinking for me'[1]. The animistic superstition that would attribute consciousness to trees and hills fails to capture most of us, not on the ground that we can 'decisively refute' it, but simply on the ground that no valid evidence demands it. My argument is not that computing devices cannot be conscious (how could I know?) but that the ability (however prodigious) of computing devices to save my mental efforts offers no rational ground whatsoever for crediting them with thinking.

In an important variant of the second class of computing process, the physical experiment used is rather like the tossing of a biassed die or the rolling of a ball down a complex pin-board. For a simple

illustration[3] imagine a ball rolling down an inclined runway that tapers to a knife-edge, so that eventually the ball must drop to left or right. If the runway can be tilted around its long axis, we have a simple example of a device where the *probability* of the outcomes left/right can be modulated according to the angle of tilt. Automata constructed of such elements (usually electronic in form) are called 'stochastic' computers, and they offer a rich repertoire of spontaneous but statistically disciplined activity of the kind that is often required in exploration and trial-and-error learning. There is, incidentally, much evidence that stochastic processing of this kind feature in the working of the human central nervous system. Here again, however, although the element of spontaneity might encourage the superstitious to feel more justified in attributing mental capacities to such artefacts, no rational case could be made for doing so on the ground that they can imitate, or are useful in informing, the behaviour of a thinking human being.

Summing up so far, I have argued (along well-worn lines) that as long as by a 'computer' we mean a device that helps us in our thinking by presenting us with outcomes dependent on specified data according to known principles, there are no rational grounds for crediting such computers with the sort of mental activity that would be required if we (as conscious cognitive agents) were to work mentally through the steps leading from those data to those outcomes. As a manipulator of symbols, a computer has no more claim to understand what it is doing than a manipulator of apples or oranges. Used as an aid to human thinking, its operations replicate things the thinker might do with his *hands* (writing down symbols, looking up tables etc.) rather than with his *mind*. And in agreement with Searle and many others before him, let me point out that this argument in no way depends on any assumed limitations to future progress in computer science.

Do Brains Understand?

But it is time we cleared the air by asking just what it is we claim in the case of human beings when we say *they* think, feel or understand. For Searle[2] the claim is straightforward: 'Of course, our *brains* can think' (p. 36). 'Pains and other mental phenomena are features of the *brain*' (p. 19). 'We can say of a particular *brain*, this *brain* is

3 MacKay, D. M. (1952) 'Mentality in Machines' (Third paper in Symposium). *Proc. Aristot. Soc. Suppt.*, XXVI, pp. 182–198.

conscious or this *brain* is experiencing this or that . . .'[4]. Mental processes, he argues, are biological phenomena, as biologically based as the secretion of bile. 'The substance of the brain (is) conscious'[5]. On this view to say that human beings think, understand or the like is to claim that their *brains* think or understand. One of Searle's reasons for doubting whether digital computers think or understand is that they are not (in present technology) biological structures. 'Brains are biological engines; their biology matters.'

To this latter argument we shall return. But first we must examine the presupposition, so frequently made explicit by Searle, that it makes sense to attribute thinking, feeling or understanding to the *brain*. Are brains the right kind of thing, in the right category, to go in for such activities as thinking and understanding?

Talk of thinking and understanding arises in the first place as part of the 'I-story' that each of us in principle can tell (and is in principle obliged to admit) about ourselves as conscious agents—people who experience as a matter of brute fact, a succession of sights, sounds, thoughts, desires and the like which we would be lying to deny. Not all these data of our conscious experience may be easy to describe in words; some may indeed prove inexpressible or ineffable. But in general we can say that talk of seeing, hearing, thinking, wanting, understanding arises when we seek to bear witness as truthfully and explicitly as we can to the immediate data of our conscious experience. The subject of each verb in this category is 'I': I-see-this, I-hear-that, I-think-the-other, I-believe-that-such-and-such. As a useful visual aid, we may picture such entries from the I-story as forming a long column on the lefthand side of a blackboard with a line down the middle. The vital point is that all these verbs belong to the category appropriate to *personal witness-bearing* by conscious cognitive agents. Personal witness-bearing is something we all know how to do—we learn it as an early skill—and can best define ostensively. Its negative is false witness-bearing or lying, which we also know directly how to do, but normally avoid. Both are the intentional activities of *people*.

Where then does the brain come into this picture? Among the immediate data to which each of us can (and if truthful must) bear witness are our experiences of seeing-our-limbs-move, reading-books-about-neurophysiology, looking-at-pictures-of-the-brain etc.

4 Searle, J. *op. cit.*, p. 22. All italics mine.
5 Searle, J. *op. cit.*, p. 23.

On this basis we may reasonably believe that for each of the immediate data to which our I-story seeks to bear witness, there is some correlated entry in the 'brain-story' that an all-seeing super-physiologist could in principle tell about our central nervous system. The correlation need not take the form of one-to-one correspondence. There is plenty of evidence of neural 'pooling' of activity such that the same conscious experience could result from any one of several patterns of neural firing. What can be said (though of course only as a working assumption, not a guaranteed fact) is that if you were *not* having the given experience (of seeing-x, thinking-y, or whatever), then certain specific entries in your brain-story would have to be *different*. It is in this (quite strong) sense that they are the correlates of the entries in your I-story. For each immediate fact of experience that you must acknowledge as a person-with-an-I-story-to-tell, there are states or activities of your brain which would *have to be significantly different* if that immediate fact were otherwise.

Lest this sounds like a rather special and heavily-qualified type of correlation, we may note that exactly the same qualifications apply to the correlation between the mathematical significance of what a computer is doing, and the physical activity in its circuits. Computers are carefully designed to tolerate variations in power supply, and in the characteristics of ageing components, so that a quite wide range of physical states of a circuit may have one and the same symbolic significance. On the other hand, if it is the case that the number represented in a given register is 4321, certain things must be true about its physical activity which would not be true if the number were 1234. The same one-to-many relationship holds generally beween symbols and their physical embodiment, e.g. in handwriting, or in gestures such as spoken words.

I stress that in relation to the human brain this is only a working hypothesis, the evidence for which is impressive but incomplete.[6] My purpose is not to argue for it (though I accept it) but to work out some of the implications if it were indeed true. It implies that (at least here on earth) certain brain-states and activities are necessary if we are to enjoy/suffer certain experiences. It does not imply that every human brain activity has some correlate in conscious experience, even where the activity in question is clearly goal-directed. (Think, for example, of the neural mechanisms that regulate the

6 See for example Buser, P. A. and Rougeul-Buser, A. (1978) *Cerebral Correlates of Conscious Experience*, Elsevier, Amsterdam, New York, Oxford.

diameter of the pupil of your eye, normally without contributing anything to your conscious experience.)

My question is whether on this working assumption (which is all that Searle or anyone else has to go on) there is any case to be made for ascribing your thinking, feeling or understanding to your brain. Assume if you will that the form of your physiological brain processes determines the content of your conscious experience of thinking, feeling and understanding. (Most people have heard of the cases studied by the neurosurgeon Wilder Penfield, who found that he could to some extent shape the conscious experience of his (epileptic) patients by injecting electric currents into their exposed brain tissue at operation.) You say, for example, 'I understand now what this tax law will mean for me', and you give examples. Our superphysiologist who knows the complete pattern of correlations is able at once to find a corresponding entry for his story about your brain. To make things simple, assume that the brain process he describes is both necessary and sufficient for you to have the corresponding experience. He can then confidently write down his description on the right hand side of our imaginary blackboard, as the cerebral correlate of your entry just cited. His decription will certainly have to use categories at a level more abstract than that of individual nerve cells, probably defined in terms of the global flow and processing of information; but in the context of the whole it leaves no relevant ambiguity as to what must be going on in your brain if it is true that you have the feeling that you understand what you claim to.

Here then we have two correlated stories, one (the I-story) about you and your immediate experience, the other (the brain-story) about a physical structure and the physiological interactions that (we suppose) must be taking place in it if the I-story is true; from the one, we may even infer the other. To suppose that this tight correlation justifies attributing activities from the one story to entities in the other is, however, a baseless error. Indeed a clear counter-example is offered by the case of computing machinery itself. Suppose we set up a computer to go through a specific program. At any point in its operation there is a strict correlation (of the same one-to-many kind) between the mathematical or logical function being executed and the physical activity of its transistors and other elements. Someone who knows the machine in detail can in principle derive the 'machine story' corresponding to every entry in the functional story. But of course this strict correlation in no way implies that the story about the machine and its mechanistic work-

ings *states the same facts* as the story about the mathematical and logical functions embodied in those workings, for the two stories are in quite different categories.

The fact that a quadratic equation with two roots is embodied in a piece of electronic hardware in no way implies that the hardware at any level of description as hardware 'has two roots'. The notion makes no sense, even though the existence of two roots has well-defined hardware implications.

There is a further objection to the attribution of understanding to the brain rather than to the person (the conscious cognitive agent with an I-story to tell). As we saw in the case of a computer, the fact that tokens are manipulated or exchanged *as if* there were somebody there who understands their meaning does not justify, or even make meaningful, the conclusion that the physical *machinery* understands. If we were to lift the lid of a man's brain and observe similar token-handling behaviour going on in its depths, this in itself would make it no more sensible to credit the *biological machinery* we are observing with thoughts, understanding or feelings. If the lid that has been lifted happens to be our own, we may be able with the help of mirrors and other instruments to check on many of the correlations assumed by brain science (though there are some startling hazards attending this form of self-observation).[7] We of course have first-hand evidence in this case that *there is someone there* thinking, feeling and understanding. But the 'someone' who can bear witness to this is ourself, not our brain or any other part of the object-world, however tightly correlated its activity may be with our conscious personal experience. If it be alleged that I 'simply am' the biological structure in question, I would argue that our evidence shows only that I am *embodied in* that biological structure, and that the things I get up to as a cognitive agent—desiring, planning, observing, understanding, acting—are not (most of them) meaningfully defined if attributed to biological structures. To put it positively, it is as *agents* that we think and understand, and have first-hand knowledge of what this means. In dialogue with our human fellow agents we are rationally convinced that they enjoy or suffer a similar flux of conscious experience. We may have open minds as to how far down

7 MacKay, D. M. (1955) 'Man as Observer–Predictor'. *Man in His Relationships*. Routledge, London, pp. 15–28.
Also: MacKay, D. M. (1971). 'The Proper Study of Oneself'. *The Proper Study* (G. N. A. Vesey, ed.) Royal Institute of Philosophy Lectures, vol. 4, Macmillan, London, pp. 48–63.

the phylogenetic tree our non-human fellow creatures also enjoy something similar; but if we think of them as doing so it is as centres of subjective awareness and intentional agency, presumed to be embodied in their biological structures as we are in ours. Nowhere do we find any meaningful basis for attributing the experiences of conscious agency to the physical structures in which they are embodied.

Artificial Agency

So far this must have seemed a very conservative approach to the powers of artefacts. I have dismissed claims that computers 'think' or 'understand', on grounds which I have also argued rule out talk of human *brains* doing so. I have argued that 'understanding' enters our conceptual vocabulary as something *agents* do. But this raises a reasonable question. Is it not possible to devise artificial agents? And if so, may we not claim for some of them that they understand?

In an important sense, the answer to the first question is 'yes'. Artifical agency requires three basic ingredients:

(1) A *repertoire* from which alternative actions can be selected.

(2) An *evaluator* that attaches a 'value' to states of affairs according to criteria which may be either given (as in a thermostat) or self-modifiable or self-set.

(3) An organizing (computing) system set up to select actions calculated to increase positive evaluation, or diminish negative evaluation, of the evaluated state of affairs.

Without going into details, it is generally known that artificial agents on these basic lines can now replace or work alongside human agents on production lines, as pilots of aircraft and the like, and can increasingly be trained to communicate with their human partners in ordinary language. This immediately makes operational a possibility that is all too easily realized. *Artificial agents can misunderstand* what is said or shown to them. If a human supervisor shouts 'WAIT', and his artificial partner responds by placing the object it is holding on a weighing machine to ascertain its weight, then by all kinds of obvious criteria there has been a misunderstanding.

But if it makes sense to say that such agents can misunderstand, there is surely a corresponding sense in which at their best they *understand* the commands and information they receive? I think this

conclusion is irresistible, and that Searle is confused in claiming that such systems have no way of getting from the syntax to the semantics of their informational traffic. For any agent (whether natural or artificial) the meaning of an item of information is ultimately encashed in terms of the contribution it makes to the agent's total state of conditional readiness for action and the planning of action.[8] Thus if I tell you that the bottle is empty, your understanding of my meaning implies (if you believe me) that in relevant circumstances you would be ready to reckon with an empty bottle—not planning or trying to fill your glass from it if thirsty and so forth. If I gave an artificial agent the same information it might lack some of the meaning it had for you (assuming for example that artificial agents do not drink from bottles); but in principle its meaning-for-each-agent is defined in terms of its impact on the total conditional readiness of each agent for action.

What I am arguing is that for an agent, whether natural or artificial, the basic descriptive dimensions of the semantic frame are those of its *conditional repertoire of action*. The meaning of any item for a particular agent is defined by the constraints it implicitly specifies on the organization of that agent's conditional readiness for action. It is thus far from obvious (and in an important sense, patently false) that no artificial system can have semantics. By equipping an artificial agent with a repertoire that interacts with a field of action, I would argue, we provide it implicitly with a semantic frame even before questions of syntax arise; indeed if the system operates on an 'analog' basis we may never have to face questions of syntax at all, though my point applies regardless of the nature of the computing processes involved. An artificial agent correctly understands the meaning of an item of information when and if the item brings about the changes in its total state of conditional readiness required to match the semantic content of the item.

Nothing said here contradicts my earlier claim that neither computers nor brains understand. Even if someone argues that they would be prepared to extend the boundaries of a 'computer' to include its peripheral organs of agency, I would reply that the artificial agent we have been discussing is something *embodied in* their extended 'computer', in the sense in which a quadratic equation

8 MacKay, D. M. (1954) 'Operational Aspects of Some Fundamental Concepts of Human Communication'. *Synthese* **9**, pp. 182–198.
Also: MacKay, D. M. (1969) *Information, Mechanism and Meaning*. M.I.T. Press, Cambridge, Mass.

could be embodied in the same computer, and that we are not entitled to attribute the capacities of embodied agents to their embodiment as such, unless we have arguments in justification. Computers do not and cannot understand, even if incorporated in active robots, for the same reason that brains do not and cannot understand—namely that understanding is something that *agents* do, and not the brains or computers in which their agency is embodied.

To sum up, Searle's claim that 'If I am the computer inside a robot, I have no way of getting from the syntax to semantics' is valid enough (though not at all original) as showing that it is not the robot's computer that should be credited with any understanding that goes on; but I have argued that in any case it is not computers or brains but *agents* to whom it makes sense to attribute understanding. In the case of artificial agents, even if embodied in a structure that includes a computer, I see no reason to deny them the capacity to understand or to misunderstand the meaning of items of information. In their case, Searle's dilemma over 'getting from the syntax to semantics' is illusory. For any agent, semantics comes in with the constraints imposed by the structure of their field of action on the planning and execution of evaluated agency. To have a repertoire of evaluated agency is, in principle, to 'have semantics', even though a complex learning process may be needed to establish what particular signals mean—i.e. what constraints they impose on the conditional readiness to act.

Consciousness

But it is time to face one remaining question which I hope it is clear that my argument so far has left open. Granted (if you wish) that artificial agents can both understand and misunderstand, and in an important sense 'have semantics', does this mean that we must credit them with conscious experience? Do they think, feel, desire as we do? Is each of them a centre of awareness with an I-story to tell?

Nothing in the foregoing discussion, it seems to me, justifies any such conclusion. Agency, even evaluative and goal-pursuing, can be implemented in the human body itself without necessarily forming a part of what we consciously do. I have elsewhere[9] speculated as to

9 MacKay, D. M. (1951) 'Mindlike Behaviour in Artefacts'. *Brit. J. Phil. of Sci.* **II**, pp. 105–121. Also **III**, pp. 352–353. (1953).
Also: MacKay, D. M. (1980). 'Conscious Agency with Unsplit and Split Brains. *Consciousness and the Physical World* (B. D. Josephson and V. S. Ramachandran, eds), Pergamon, pp. 95–113.

the special 'supervisory' features of brain organization that may be essential to conscious as opposed to unconscious agency. I do not here want to argue whether the implementation of such special features in an artificial agent could be sufficient to give that agent conscious experience. Personally, I am inclined to doubt it; and (like Searle) I have in the past suggested that specific biological constituents of our embodiment may (for all we know) be essential[1]. Certainly only the crudest of behaviourism could be satisfied with the idea that behaviour-as-if-conscious is all we need require.

If we take this line, however, we must recognize a certain cost. The question whether A understands X, believes Y, or desires Z then becomes separable from the question whether A is conscious. Can 'unconscious understanding' make sense? Freud (for good or ill) persuaded many of his generation to accept the notion of 'unconscious desires'. How far can we go towards a complete psychology of agency divorced from the presupposition that the agent is conscious?

The plain truth is that in these days of increasingly sophisticated automata we have an urgent practical need for just such a conceptual framework. A relatively trivial example would be when playing a game, such as chess, against an artificial opponent. Not only understanding and misunderstanding but threatening and perceiving threats as such, feinting, setting traps and the like are forms of behaviour which the artificial opponent can not merely simulate as an end-product but *generate on the same principles* as a human opponent might. Much more serious fields of application can readily be envisaged in such contexts as international business competition or military posturing.

The agnostic methodology needed to make insightful sense of artificial agency is, I suggest, of the same family as the discredited behaviourism that an earlier generation of psychologists tried to foist on students of human behaviour. As applied to human beings, extreme behaviourism deserved its contemptuous dismissal as 'feigning anaesthesia', since each self-styled behaviourist must have had direct knowledge, which he would have been lying to deny, of the extent to which his own behaviour was not only accompanied, but also shaped by his conscious mental activity. Moreover, any attempt to define psychological categories purely in terms of externally observable behaviour would be as inept in the case of artificial agents as it proved to be in that of natural agents. Our only hope of developing a useful psychology of artificial agency will lie in analysis of the *springs* of action—the internalized processes of evaluation,

updating of states of readiness to match sensory demands, selection of action calculated to increase positive or diminish negative evaluation, and the like—from which the behaviour in question issues as a natural product according to the same principles as in human agents. The discipline of 'A.I.' (artificial intelligence) attempts to make such generative principles explicit. I happen to believe that in using *rule-following* as its model for all human cognitive processes, contemporary A.I. is missing an important constituent of the generative process underlying human thinking. In the brain, as I mentioned earlier, state-transitions are often determined not by consulting rules but by what amount to local physical experiments, usually with a stochastic ingredient which can give rise to spontaneous (though statistically not nonsensical) turns of events.[10] But insofar as the aim of A.I. is to model the generative process, and not merely to simulate its product, it would be inept to dismiss it as of no potential explanatory value.

What made behaviourism lose respect (after a surprisingly, and instructively, long innings) was its deliberate failure to reckon with the brute fact of consciousness in the case of its human subjects, and its Procrustean efforts to tie all its categories to externally manifest bodily movements. Since artificial agents in general are not known to have conscious experience, and since we usually have direct access to the internal evaluative and other processes that determine their conditional readiness for overt behaviour, I see no reason why a reformed and chastened behaviouristic approach, focusing on the generators of behavioural readiness, should not in their case have more permanent value. Indeed I would personally hope that just such an approach might in the end converge with such disciplines as neuropsychology and cognitive neuroscience to provide a much-needed theoretical basis for the understanding of normal and pathological human behaviour.

That we, unlike our artefacts, are conscious centres of awareness will remain a challenging antidote to behaviouristic complacency; but it need not, I think, rob us of the insights into the behaviour of artificial agents that come from recognizing their genuinely psychological capacities.

The late Professor Donald MacKay was Professor of Communication and Neuroscience at the University of Keele, England.

10 MacKay, D. M. (1954) *op. cit.* and MacKay, D. M. (1985) 'Machines, Brains and Persons'. Zygon **20**, 401–12.