

Faith and Thought

A Journal devoted to the study of the inter-relation
of the Christian revelation and modern research

Vol. 90

Number 2

Autumn 1958

D. M. MACKAY B.SC., PH.D.

Brain and Will¹

PART I

LOGICAL VERSUS PHYSICAL INDETERMINACY

The classical debate

Everyone admits that some human actions may sometimes be determined by the physical state of the brain. No one doubts that the convulsions of epilepsy or the tremors of Parkinson's disease have, as we say, a physical cause; and most of us would admit that many of our less spectacular actions could probably also be traced back continuously to the physical action of our central nervous system. At least it would not worry us if it were so.

It worries nobody, as long as the actions concerned are not of a kind to which we attach moral significance. But as soon as we come to acts of choice in which questions of responsibility might arise, we find ourselves in the middle of a well-trodden battlefield. On the one hand, there are those who believe that if my choice is to be morally valid, the physical activity of my brain must at some point 'change its course' in a way which is not determined by purely physical factors. They do not mean only that the change would be too complicated to work out in practice—though in fact it probably would be. They believe that even with unlimited powers of calculation, and complete physical information about every part of the brain, it would be impossible to know the change in advance, because, they would say, the change does not depend only on physical factors. If it did, then the choice would not be a morally valid one.

According to this view, then, the brain is to be thought of as an instrument often likened to a pianoforte, with at least a few controlling keys open to influences of a non-physical kind. I shall refer to it, for short, as the 'open-system' view.

Over against this view we have a strong body of opinion, particularly among scientists, which maintains that even when I make a moral choice, the physical changes in my brain depend entirely on the physical

¹ This paper is based on two B.B.C. talks which were reprinted in *The Listener* issues dated 9 and 16 May 1957.

events that lead up to them. On this view there would be no discontinuity in the chain of physical cause and effect. A complete knowledge of the immediately preceding state, it is believed, would always be sufficient in principle to indicate beforehand which choice would be made. No openings are admitted for any non-physical influences to disrupt the expected pattern. We may refer to this as the 'closed-system' view of the brain.

On both sides there are plenty of varieties of opinion. Some who hold the 'open-system' view would maintain that each morally valid choice—each choice for which I may properly be held responsible—requires a miraculous physical change to take place in the brain. Others, such as Dr E. L. Mascall in his recent Bampton Lectures, hold that the well-known indeterminacy of small-scale physical events, first formulated by Heisenberg, could allow the brain to respond to non-physical influences without disobeying physical laws.

In the 'closed-system' camp there are even more varieties of opinion about the 'mental' aspect. Some robustly deny that there are any morally valid choices. They agree with the 'open-system' people that a choice could not be valid unless it falsified or went beyond what was indicated beforehand by the state of the brain—but they do not believe that human choices do so. Others, again, would argue that questions of moral validity are 'meaningless'; and so we could go on.

A Prior Question

But I am not concerned here to come down on one side or the other of this traditional fence. I simply do not know—nobody knows—to what extent the processes going on in the brain are physically determined. We are gradually accumulating evidence which suggests that brain tissue does behave according to the same physical principles as the rest of the body; and we now know also that no behaviour-pattern which we can observe and specify is beyond the capabilities of a physical mechanism. On the other hand, it is undeniable that some processes in the brain might occasionally be affected by physically indeterminate events of the sort which Heisenberg's Principle allows.

No, what I want to do is to undercut all discussion of this kind by raising a group of prior questions which might profitably have been asked before sides were picked on the traditional ground. The central question is: Could I be *excused* from responsibility if a choice of mine did *not* involve any physically indeterminate changes in my brain?

At first sight the answer may seem obvious. 'Surely', we may say, 'a choice which is uniquely indicated beforehand by the state of the brain cannot be called a "free" choice? If you could in principle predict how I shall choose before I make my choice, surely my choosing has no moral validity?' In one sense this is obvious. We should all agree that if we could be given a description of our action beforehand, and had no power to help or hinder its fulfilment, then we should have to admit that this action was not 'free' but involuntary. A sneeze, for example, at a sufficiently advanced stage, might be judged involuntary by this criterion. So would a simple reflex action like a knee-jerk or an eye-blink.

But—and this is the point—even supposing that the necessary brain-processes were determined only by physical factors, are we sure that what we normally call a 'free choice' could be described to us in advance? I think not. In fact I believe that whether the brain-mechanism is physically determinate or not, the activity which we call 'making a free choice' is of a special kind which could never be described to us with certainty beforehand. Suppose we are asked to choose between porridge and prunes for breakfast. We think: 'Let's see: I've had prunes all last week; I'm sick of prunes; I'll have porridge.' We would normally claim now to have made a 'free choice'. But suppose that some super-physiologist has been observing our brain-workings all this time; and suppose he declares that our brain went through nothing but physically determinate actions. Does this mean that he could have told us in advance that we would certainly choose porridge? Of course not. However carefully calculated the super-physiologist's proffered description of our choice, we would know—and he would know—that we still had power to alter it.

Logical Indeterminacy

No matter how much he tried to allow in advance for the effects of his telling us, we could still defy him to give us a valid description of what our choice would be. This is our plain everyday experience of what most people mean by our free choice: a choice which nobody could (even in principle) describe to us in advance. My point is that this vital criterion of freedom of choice, which we shall see later can be extended and strengthened, would apply equally well whether the brain were physically determinate in its workings or not. In either case, the state of our brain after receiving his description would not (and

could not) be the state on which he based his calculations. If he were to try to allow beforehand for the effects of his description upon us, he would be doomed to an endless regression—logically chasing his own tail in an effort to allow for the effects of allowing for the effects of allowing . . . indefinitely. This sort of logical situation was analysed some years ago in another connection by Professor Karl Popper, and the conclusion I think is watertight. Any proffered description of our choice would automatically be self-invalidating.

It is necessary, however, to carry the argument a stage further. One might get the impression from what I have said that our choice could not be proved free in this sense unless we succeeded in actually falsifying a would-be description of it. But this is not so. If we are supposing that our super-physiologist has access to all our brain-workings, then our freedom to nullify predictions of our choices can in principle be established simply by examining the structure—the blueprint, so to speak—of those brain-workings. It is not necessary actually to make the experiment of presenting us with an alleged 'prediction', in order to verify that the basis of the prediction would be invalidated. The point is simply that the brain is always altered by receiving information; so that the brain which has received a description of itself cannot possibly be in the state described. Provided that the parts of our brain concerned with receiving and understanding the information are linked up with the mechanisms concerned with our taking the decision (and nobody doubts this even on the 'closed-system' view), then it is logically impossible to give us—or even to make ourselves imagine—a valid description of a decision we are still deliberating, whether on the basis of advance observation or anything else. It is not that we are unable to ascertain the true description. It is that for us *there is no true description to ascertain*. For us the decision is something not to be ascertained but to be *made*. In fact, any description would be for us *logically indeterminate* (neither true nor false) because it would be self-referring in a contradictory way, rather like the statement: 'This sentence I am now uttering is false.'

It is this *logical* indeterminacy, of statements predicting our decisions, which has tended in the past to be confused with *physical* indeterminacy, as something which was thought to be necessary if a choice were to be morally valid. We all feel intuitively that there is something queerly 'undetermined' about the decisions we take—that there is something absurd and self-contradictory in trying to believe or even consider as 'true now' any advance description of them. I hope I have shown that

this intuitive feeling is entirely justified—but on grounds which have nothing to do with physical indeterminacy in the matter of our brains. We appear to be so constructed that any would-be prediction of our voluntary actions becomes for us merely an invitation to choose how to act. This is not only theory, but also empirical fact. If anyone tries to predict to us that we are about to choose porridge rather than prunes, no matter how scientific the basis of his statement, we can easily verify that he is simply giving us a fresh opportunity to make up our minds. Whether we decide in the end to fall in with his would-be prediction or to contradict it, we know—and he knows—that it has lost any scientific validity by being offered to us.

'I Knew You'd Choose That'

But, we may well ask, what if our super-physiologist does *not* tell me of his prediction? What if he just keeps his mouth shut and watches how I choose, and then says, 'Aha, I knew you'd choose that'? We must admit straight away that we should feel rather upset if anybody could do that to us every time we made a choice; and I must agree that I do not believe it could ever be done consistently in practice. Consistent success would be possible only if our brains *were* physically determinate, and if the super-physiologist could know the whole of our brain-workings, together with all the influences which would act on them from the outside world. The first supposition is doubtful and the second is certainly impossible on practical grounds of sheer complexity; and between them I think these considerations are enough to account for—and justify—our feeling of incredulity.

But suppose for the sake of argument that it *were* so: that although we can defy anyone to tell us how we are going to choose, yet a successful prediction of our choice could in principle be made by someone who keeps quiet about it. What then? Could we excuse ourselves from responsibility for our choice on these grounds? I do not think so. If we had no power to falsify his prediction, we might indeed excuse ourselves. But in this case there is no doubt that we have the power. Our silent observer is only denying us the opportunity to demonstrate it. He knows, as well as we, that in fact his prediction is only conditionally 'certain': certain just so long as we do not know it; and it is rather an odd sort of 'certainty' that you have to hide from someone in case it turns false! Clearly even when he kept quiet the sense in which his prediction was 'certain' would be a rather limited one.

As a matter of fact, the great majority of our choices day by day could be predicted with great success without even opening our heads, by anyone who knows us sufficiently well; but it never occurs to us to question our responsibility for them on these grounds. At least if it does I do not think it ought to, for all it means is that we make most of our choices 'in character'; not that we could not have chosen otherwise (if confronted with the allegedly 'certain' prediction), but simply that we were not inclined to—and might not have felt so inclined even if the prediction had been offered to us.

In short, the super-physiologist's knowledge, if our brain-workings accurately reflected what we were thinking, would do no more than enable him to make predictions as if he knew what was going on in our minds. In that case it would be surprising if he were not successful, so long as he kept quiet; but we could never appeal to his evidence in order to excuse ourselves from responsibility for such choices, for at most it could only offer *confirmation*—and not contradiction—of the mental processes in terms of which our moral responsibility would be judged.

To sum up thus far, I believe that brain-processes may well include some events which are physically indeterminate as well as many which are not. But I am suggesting that our responsibility for moral choices rests not on any physical indeterminateness of our brains, but on the logical indeterminateness to us of any advance description of our decisions. It is the unique organisation of our brains which gives this peculiar status to our decisions—not anything physically queer about their workings. If there is any physical indeterminacy, its effects will be entirely different, as we shall now see.

PART II

THE SIGNIFICANCE OF PHYSICAL INDETERMINACY

Heisenberg's Principle

It is just over thirty years since Heisenberg enunciated his principle of indeterminacy, asserting that the motions of atomic particles can never be predicted exactly from the physical data available to us. Laplace's dream of a clockwork universe was gone; in fact, according to Eddington, just half of the data which we would require for a complete prediction of the universe are not available until after the change

we want to predict. But if this came as a blow to the classical physicists, it was welcomed with open arms in other quarters. To those who felt that the dignity of man was being threatened by the creeping spread of physical causality to the very mechanism of the brain, Heisenberg's principle seemed a God-send. Here, surely, was the solution to the problem of free will. 'If atomic particles are physically indeterminate in their movements, then, since my brain is made up of atomic particles, its activity is not physically determined, and my will is free'—so the argument ran.

I have been arguing that the kind of 'freedom' which physical indeterminacy would give us is not required in order to establish moral responsibility: that on the contrary, whether my brain were *physically* determinate or not, my choosing is for me a unique and *logically* indeterminate activity for which I could not escape full moral responsibility. We must now take a look at the other side of the picture; for I have no wish to deny that physically indeterminate events may sometimes take place in our brains; and it is interesting I think to see what kind of effects these events could have upon the delicate and complex processes going on in our heads.

The first thing to keep in mind is that the degree of physical indeterminacy allowed by Heisenberg's principle becomes more and more negligible, the bigger and heavier the objects we are studying. Indeed it is only with the smallest objects of all—electrons, for example—that it is really serious. A nerve cell may be a tiny object by everyday standards; but it is roughly a million million million times heavier than an electron; so the chances of its suffering appreciably from Heisenberg indeterminacy are small indeed. Even if we suppose that the controlling part of a nerve cell weighs only one-millionth of the whole, we are still thinking on a scale a million million times larger than that of the electron.

There are about 10,000,000,000 nerve cells in each of our brains; so the chance that some *one* of these should be disturbed by a physically indeterminate event is correspondingly greater. But this brings us to the second point. The brain is not like a wireless set, in which a single valve-failure is enough to upset the whole performance. The nerve cells in the brain seem to be organised on a principle of team-work, often with hundreds or even thousands of cells working together on any one job—rather like the individual strands in a rope. Even if one of our brain-cells were put out of action altogether, the chances are that it would make no significant difference. Only a most unusual

combination of circumstances could allow the behaviour of the brain as a whole to be affected.

One further point needs to be made before we discuss the implications of all this. The brain has to carry on its business in the face of all manner of physical disturbances besides those which Heisenberg has discussed. There are random vibrations due to the heat of the brain-tissue for example, random fluctuations in blood supply, and random disturbances reaching the brain from the outside world. These are not indeterminate influences in principle, but in practice they are far too complex to be predictable; and their effects are much larger than those due to Heisenberg indeterminacy, though similar in other respects. Yet, surprisingly enough, in spite of all those unpredictable influences, the brain still manages to work. It is in fact marvellously designed to be unaffected by disturbances of this kind. It follows that if the brain is at all affected appreciably by the physically indeterminate 'Heisenberg' variety of disturbance, this ought to be a much rarer occurrence than the other sorts, which are not absolutely unpredictable. Hardly any of the disturbances which do have significant effects are likely to be of the feeble Heisenberg type.

Effects of Physical Indeterminacy

What, then, could we expect to be the effects of such unpredictable disturbances? In the first place, they would undeniably introduce a certain kind of 'freedom' into the brain's activity. But I suggest that this would not be the freedom characteristic of rational moral choice and responsibility, which we have seen to be something different. It would rather be of the kind we should call 'spontaneity' or even sometimes 'mental aberration'—according to the part of the brain affected by it. In most cases it would mean the interruption of a normal train of thought by an 'unbidden idea', as we would say, or by some 'unaccountable lapse'. Perhaps this really does happen on occasions. If it does, it raises the interesting question whether the person concerned could properly be held responsible for what has happened. So far from enhancing his responsibility, such undetermined events would seem if anything to lessen it. We may be reminded of the fact that great composers and artists have often disclaimed responsibility for their inspirations, saying that they 'received them unbidden', though I am far from suggesting that originality is only a matter of random disturbances in the brain. I only want to emphasise that in most cases the

unpredictability produced in this way would not seem to enhance responsibility for the resulting action.

But now, it may be asked, what if I were deliberating a choice between two possibilities which was so finely balanced that I could find no reason for favouring one rather than the other—like Buridan's donkey, which starved to death, we remember, because it could not choose between two equally tempting bundles of hay: might not the outcome ultimately be settled by one of these unpredictable disturbances? I think this might well be so, and that the resulting choice might be unpredictable even to a super-physiologist who knew all that was going on in our heads—and kept his mouth shut. But what would be our own view of such a choice? Would we want to give it a higher moral status than one in which the right issue was clear to us and we decided unwaveringly on principle? I doubt it. Indeed I think that to such a finely balanced choice I would attach if anything a lower moral significance—rather as if I had settled it by mentally tossing a coin.

There are, however, more subtle effects which unpredictable disturbances could have. When we make a choice, we take into account all the pros and cons we can think of, weigh them up, and decide accordingly. All of this, I believe, requires physical activity in our brains, which in a sense indicates or represents what we are thinking. Suppose that I make some choice which seems to me straightforward on the evidence I have considered. I see no reason to doubt that the corresponding physical activity in my brain might be equally 'straightforward'—in other words, it might well have nothing physically discontinuous or 'queer' about it. But now, how did I come to consider the evidence I did? Obviously, I could never think of all the factors that might conceivably be relevant. There is an unconscious selection of evidence, which I believe also involves a physical brain-process; and if this process were to suffer one of these unpredictable disturbances, I might well have no conscious awareness of it at all. It would mean simply that some factor, affecting my decision, would come to mind, or fail to come to mind, as a result. There would be nothing to indicate to me that anything unusual had occurred. And yet, in consequence of this disturbance, the different selection of factors might sometimes lead me just as clearly to the opposite decision.

In either case, I think I would be fully responsible for my decision. But in the second case it would have an unexpectedness, from the observer's angle, which it would lack if there had been no disturbance of the process by which the evidence was brought to my conscious

attention. To sum it up, I am suggesting that although physical indeterminacy in the brain is not necessary for moral responsibility, there is some evidence that occasional brain disturbances may be physically unpredictable, and that a small minority of these could be physically indeterminate. Such discontinuities, however, would show themselves more as a kind of originality or spontaneity, than in connection with a deliberate moral choice; and it is only if they affected the unconscious selection of evidence that they might be said to play any significant part in such a choice. Their general effect would be, if anything, to weaken rather than strengthen responsibility for any action which resulted.

From all this you will gather that I have not much hope of Heisenberg's indeterminacy as a gateway through which the mind acts on the brain. Perhaps it would be only fair to try to indicate how I think the two are related, for I believe most seriously in both the spiritual and the physical aspects of our human nature.

'Subject-language' and 'Object-language'

The trouble here, I believe, is that we have two different and entirely legitimate languages which we use about human activity, but that these tend to get mixed up in illegitimate ways. On the one hand there is what we might call 'subject-language', to which belong words describing mental activity, like thinking, choosing, loving, hating, and so forth. All of these are words defined from the standpoint of myself as the actor in the situation. From the standpoint of an observer of the situation, on the other hand, we can define an entirely different vocabulary, making up what we might call 'object-language'. To this belong words like 'brain', 'nerve cell', 'glandular secretion', 'electric current', and so forth.

The problem is to discover how descriptions in these two languages can be related. I think out some decision, let us say, and at the same time a scientist observes certain physical events in my brain. Are we to say that my decision causes the physical events, or that the physical events cause my decision, or is there some different way of relating the two? My own view, for what it is worth, is that my decision neither causes nor is caused by its immediate physical concomitants. For we can only say 'A causes B' when A and B are two activities (two separate events or sets of events). And my suggestion is that the mental activity I describe in 'subject-language' and the corresponding

brain-activity described in 'object-language' are not two activities, but two aspects of one and the same activity, which in its full nature is richer—has more significance—than can be expressed in either language alone, or even in both together.

I am not suggesting that mental activity is 'nothing but' an aspect of brain-activity: this would be the attitude which I call 'nothing buttery', and one might equally fallaciously maintain the converse. The idea is rather that each is a descriptive projection, so to say, of a single complex unity which we can call simply my-activity. An observer can describe my-activity under the aspect of brain-activity; I myself can describe it under the aspect of mental-activity; but each, and any, descriptive projection, however exhaustive in its own language, can do only partial justice to the complex and mysterious reality that is my activity as a human being.

As a crude illustration of what I mean by 'doing partial justice' imagine the two descriptions which a physicist and a telegraphist might give of a morse signal, sent by flash-lamp from ship to shore. The physicist might exhaustively record the duration and intensity of every light flash, without ever mentioning the message. The telegraphist might exhaustively record every word of the message without ever mentioning the intensity of the light. Each description, exhaustive though it is, requires to be complemented by the other in order to do justice to the significance of what took place. The two, as we say, are logically complementary. We do not debunk the one by claiming that the other is exhaustive, nor do we justify the one by trying to find discontinuities or gaps in the other.

The Unity of Mental and Physical

It would follow from this view that there is no need—indeed it would be fallacious—to look for a causal mechanism by which mental and physical activity could act on one another. Their unity is already a closer (and a more mysterious) one than if they were pictured as separate activities in quasi-mechanical interaction, one of them visible and the other invisible. Yet it is a unity which safeguards rather than threatens my responsibility for my choosings; for it makes nonsense of any suggestion that my body, rather than I myself, could be held responsible for them. This would be simply to muddle up the two languages—rather like asserting, or denying, that when a man feels in love, his brain-cells feel in love. Such a statement is neither true nor

false, but meaningless, because feeling in love is an activity of subjects, not of objects; and when a man is feeling in love, his brain-cells are presumably fully occupied doing something physically describable in 'object-language' as the correlate of this mental condition.

I would suggest indeed that the theory of mental activity as an 'extra' which interacts with the brain, is not only unnecessary, but also open to two serious objections. First, it hangs the whole of morality on an unsupported physical hypothesis—namely, that brain activity shows discontinuities, in the right places, which would require non-physical influences for their explanation. Even in the present primitive state of our knowledge this hypothesis now looks more improbable with every advance in the science of the brain. Secondly, the theory would deny my responsibility for any choices which did not entail physical discontinuity in my brain, even although I made them deliberately, and could defy anyone to describe them to me beforehand. This I believe to be flatly immoral, and a menace to a human being's right, as we say, to 'know his own mind'. If there were any question that someone's brain were disordered—prevented from functioning properly—then it might be legitimate to deny his responsibility. This could in principle be settled by examining the structure of his brain; but it would be fallacious to describe a brain as disordered merely because it failed to show any physical discontinuities, or because one could discern some of the pattern of physical cause-and-effect which was the necessary correlate of the man's mental activity. I believe that this represents a fallacy to be guarded against particularly in much of our contemporary thinking about the penal code. If I am right, there is need for a radical rethinking of the role of psychiatric evidence especially, in the assessing of moral, if not legal, responsibility.

But to follow this now would take us too far. I would just repeat once more the main contention of this paper—that to hang moral responsibility on theories of physical indeterminacy in the brain is both misguided and immoral: misguided, because my responsibility is adequately nailed to my door if my choice is *logically* indeterminate until I make it—which could be true even if my brain showed no physical discontinuities; immoral, because a reliance on physical indeterminacy would deny responsibility for choices (whether good or bad) for which I think a man has a right to claim responsibility. This is no less distressing because those who hold such views do so in the name of human dignity. But I believe that our true dignity lies in having the

humility to see ourselves for what we are: and I am convinced that the Christian doctrine of man at any rate, in all its fullness, requires no licence for his brain to suffer non-physical disturbances. There is, as I have said, a profound mystery in our human nature; but it stands wholly apart from any scientific puzzles that we may find in the brain. It will be in our wisdom to avoid any temptation to confound the two.