

Divided Minds and the Nature of Persons

(From Colin Blakemore & Susan Greenfield, *Mindwaves*, Basil Blackwell, 1987)

It was the split-brain cases which drew me into philosophy. Our knowledge of these cases depends on the results of various psychological tests, as described by Donald MacKay.¹ These tests made use of two facts. We control each of our arms, and see what is in each half of our visual fields, with only one of our hemispheres. When someone's hemispheres have been disconnected, psychologists can thus present to this person two different written questions in the two halves of his visual field, and can receive two different answers written by this person's two hands.

Here is a simplified imaginary version of the kind of evidence that such tests provide. One of these people looks fixedly at the centre of a wide screen, whose left half is red and right half is blue. On each half in a darker shade are the words, 'How many colours can you see?' With both hands the person writes, 'Only one'. The words are now changed to read, 'Which is the only colour that you can see?' With one of his hands the person writes 'Red', with the other he writes 'Blue'.

If this is how such a person responds, I would conclude that he is having two visual sensations—that he does, as he claims, see both red and blue. But in seeing each colour he is not aware of seeing the other. He has two streams of consciousness, in each of which he can see only one colour. In one stream he sees red, and at the

¹ See MacKay's contribution, chapter 1 of *Mindwaves*, ed. Colin Blakemore and Susan Greenfield (Oxford: Basil Blackwell, 1987), pp. 5-16.

same time, in his other stream, he sees blue. More generally, he could be having at the same time two series of thoughts and sensations, in having each of which he is unaware of having the other.

This conclusion has been questioned. It has been claimed by some that there are not *two* streams of consciousness, on the ground that the sub-dominant hemisphere is a part of the brain whose functioning involves no consciousness. If this were true, these cases would lose most of their interest. I believe that it is not true, chiefly because, if a person's dominant hemisphere is destroyed, this person is able to react in the way in which, in the split-brain cases, the sub-dominant hemisphere reacts, and we do not believe that such a person is just an automaton, without consciousness. The sub-dominant hemisphere is, of course, much less developed in certain ways, typically having the linguistic abilities of a three-year-old. But three-year-olds are conscious. This supports the view that, in split-brain cases, there *are* two streams of consciousness.

Another view is that, in these cases, there are two persons involved, sharing the same body. Like Professor MacKay, I believe that we should reject this view. My reason for believing this is, however, different. Professor MacKay denies that there are two persons involved because he believes that there is only one person involved. I believe that, in a sense, the number of persons involved is none.

The Ego Theory and the Bundle Theory

To explain this sense I must, for a while, turn away from the split-brain cases. There are two theories about what persons are, and what is involved in a person's continued existence over time. On the *Ego Theory*, a person's continued existence cannot be explained except as the continued existence of a particular *Ego*, or *subject of experiences*. An Ego Theorist claims that, if we ask what unifies someone's consciousness at any time—what makes it true,

for example, that I can now both see what I am typing and hear the wind outside my window—the answer is that these are both experiences which are being had by me, this person, at this time. Similarly, what explains the unity of a person’s whole life is the fact that all of the experiences in this life are had by the same person, or subject of experiences. In its best-known form, the *Cartesian view*, each person is a persisting purely mental thing—a soul, or spiritual substance.

The rival view is the *Bundle Theory*. Like most styles in art—Gothic, baroque, rococo, etc.—this theory owes its name to its critics. But the name is good enough. According to the Bundle Theory, we can’t explain either the unity of consciousness at any time, or the unity of a whole life, by referring to a person. Instead we must claim that there are long series of different mental states and events—thoughts, sensations, and the like—each series being what we call one life. Each series is unified by various kinds of causal relation, such as the relations that hold between experiences and later memories of them. Each series is thus like a bundle tied up with string.

In a sense, a Bundle Theorist denies the existence of persons. An outright denial is of course absurd. As Reid protested in the eighteenth century, “I am not thought, I am not action, I am not feeling; I am something which thinks and acts and feels.” I am not a series of events but a person. A Bundle Theorist admits this fact, but claims it to be only a fact about our grammar, or our language. There are persons or subjects in this language-dependent way. If, however, persons are believed to be more than this—to be separately existing things, distinct from our brains and bodies, and the various kinds of mental states and events—the Bundle Theorist denies that there are such things.

The first Bundle Theorist was Buddha, who taught ‘anatta’, or the *No Self view*. Buddhists concede that selves or persons have ‘nominal existence’, by which they mean that persons are merely combinations of other elements. Only what exists by itself, as a

separate element, has instead what Buddhists call ‘actual existence’. Here are some quotations from Buddhist texts:

At the beginning of their conversation the king politely asks the monk his name, and receives the following reply: “Sir, I am known as ‘Nagasena’; my fellows in the religious life address me as ‘Nagasena’. Although my parents gave me the name ... it is just an appellation, a form of speech, a description, a conventional usage. ‘Nagasena’ is only a name, for no person is found here.”

A sentient being does exist, you think, O Mara? You are misled by a false conception. This bundle of elements is void of Self. In it there is no sentient being. Just as a set of wooden parts receives the name of carriage, so do we give to elements the name of fancied being.

Buddha has spoken thus: “O Brethren, actions do exist, and also their consequences, but the person that acts does not. There is no one to cast away this set of elements, and no one to assume a new set of them. There exists no Individual, it is only a conventional name given to a set of elements.”²

Buddha’s claims are strikingly similar to the claims advanced by several Western writers. Since these writers knew nothing of Buddha, the similarity of these claims suggests that they are not merely part of one cultural tradition, in one period. They may be, as I believe they are, true.

What We Believe Ourselves to Be

Given the advances in psychology and neurophysiology, the Bundle Theory may now seem to be obviously true. It may seem uninteresting to deny that there are separately existing Egos, which are distinct from brains and bodies and the various kinds of mental states and events. But this is not the only issue. We may be

² For the sources of these and similar quotations, see my *Reasons and Persons* pp. 502-3, 532. (Oxford: Oxford University Press, 1984).

convinced that the Ego Theory is false, or even senseless. Most of us, however, even if we are not aware of this, also have certain beliefs about what is involved in our continued existence over time. And these beliefs would only be justified if something like the Ego Theory was true. Most of us therefore have false beliefs about what persons are, and about ourselves.

These beliefs are best revealed when we consider certain imaginary cases, often drawn from science fiction. One such case is *teletransportation*. Suppose that you enter a cubicle in which, when you press a button, a scanner records the states of all of the cells in your brain and body, destroying both while doing so. This information is then transmitted at the speed of light to some other planet, where a replicator produces a perfect organic copy of you. Since the brain of your Replica is exactly like yours, it will seem to remember living your life up to the moment when you pressed the button, its character will be just like yours, and it will be in every other way psychologically continuous with you. This psychological continuity will not have its normal cause, the continued existence of your brain, since the causal chain will run through the transmission by radio of your 'blueprint'.

Several writers claim that, if you chose to be teletransported, believing this to be the fastest way of travelling, you would be making a terrible mistake. This would not be a way of travelling, but a way of dying. It may not, they concede, be quite as bad as ordinary death. It might be some consolation to you that, after your death, you will have this Replica, which can finish the book that you are writing, act as parent to your children, and so on. But, they insist, this Replica won't be you. It will merely be someone else, who is exactly like you. This is why this prospect is nearly as bad as ordinary death.

Imagine next a whole range of cases, in each of which, in a single operation, a different proportion of the cells in your brain and body would be replaced with exact duplicates. At the near end of this range, only 1 or 2 per cent would be replaced; in the middle, 40 or

60 per cent; near the far end, 98 or 99 per cent. At the far end of this range is pure teletransportation, the case in which all of your cells would be ‘replaced’.

When you imagine that some proportion of your cells will be replaced with exact duplicates, it is natural to have the following beliefs. First, if you ask, “Will I survive? Will the resulting person be me?”, there must be an answer to this question. Either you will survive, or you are about to die. Second, the answer to this question must be either a simple “Yes” or a simple “No”. The person who wakes up either will or will not be you. There cannot be a third answer, such as that the person waking up will be half you. You can imagine yourself later being half-conscious. But if the resulting person will be fully conscious, he cannot be half you. To state these beliefs together: to the question, “Will the resulting person be me?”, there must always *be* an answer, which must be all-or-nothing.

There seem good grounds for believing that, in the case of teletransportation, your Replica would not be you. In a slight variant of this case, your Replica might be created while you were still alive, so that you could talk to one another. This seems to show that, if 100 per cent of your cells were replaced, the result would merely be a Replica of you. At the other end of my range of cases, where only 1 per cent would be replaced, the resulting person clearly *would* be you. It therefore seems that, in the cases in between, the resulting person must be either you, or merely a Replica. It seems that one of these must be true, and that it makes a great difference which is true.

How We are Not What We Believe

If these beliefs were correct, there must be some critical percentage, somewhere in this range of cases, up to which the resulting person would be you, and beyond which he would merely be your Replica. Perhaps, for example, it would be you who would wake up if the

proportion of cells replaced were 49 per cent, but if just a few more cells were also replaced, this would make all the difference, causing it to be someone else who would wake up.

That there must be some such critical percentage follows from our natural beliefs. But this conclusion is most implausible. How could a few cells make such a difference? Moreover, if there is such a critical percentage, no one could ever discover where it came. Since in all these cases the resulting person would believe that he was you, there could never be any evidence about where, in this range of cases, he would suddenly cease to be you.

On the Bundle Theory, we should reject these natural beliefs. Since you, the person, are not a separately existing entity, we can know exactly what would happen without answering the question of what will happen to you. Moreover, in the cases in the middle of my range, it is an empty question whether the resulting person would be you, or would merely be someone else who is exactly like you. These are not here two different possibilities, one of which must be true. These are merely two different descriptions of the very same course of events. If 50 per cent of your cells were replaced with exact duplicates, we could call the resulting person you, or we could call him merely your Replica. But since these are not here different possibilities, this is a mere choice of words.

As Buddha claimed, the Bundle Theory is hard to believe. It is hard to accept that it could be an empty question whether one is about to die, or will instead live for many years.

What we are being asked to accept may be made clearer with this analogy. Suppose that a certain club exists for some time, holding regular meetings. The meetings then cease. Some years later, several people form a club with the same name, and the same rules. We can ask, "Did these people revive the very same club? Or did they merely start up another club which is exactly similar?" Given certain further details, this would be another empty question. We could know just what happened without answering this question.

Suppose that someone said: “But there must be an answer. The club meeting later must either be, or not be, the very same club.” This would show that this person didn't understand the nature of clubs.

In the same way, if we have any worries about my imagined cases, we don't understand the nature of persons. In each of my cases, you would know that the resulting person would be both psychologically and physically exactly like you, and that he would have some particular proportion of the cells in your brain and body— 90 per cent, or 10 per cent, or, in the case of teletransportation, 0 per cent. Knowing this, you know everything. How could it be a real question what would happen to you, unless you are a separately existing Ego, distinct from a brain and body, and the various kinds of mental state and event? If there are no such Egos, there is nothing else to ask a real question about.

Accepting the Bundle Theory is not only hard; it may also affect our emotions. As Buddha claimed, it may undermine our concern about our own futures. This effect can be suggested by redescribing this change of view. Suppose that you are about to be destroyed, but will later have a Replica on Mars. You would naturally believe that this prospect is about as bad as ordinary death, since your Replica won't be you. On the Bundle Theory, the fact that your Replica won't be you just consists in the fact that, though it will be fully psychologically continuous with you, this continuity won't have its normal cause. But when you object to teletransportation you are not objecting merely to the abnormality of this cause. You are objecting that this cause won't get you to Mars. You fear that the abnormal cause will fail to produce a further and all-important fact, which is different from the fact that your Replica will be psychologically continuous with you. You do not merely want there to be psychological continuity between you and some future person. You want to *be* this future person. On the Bundle Theory, there is no such special further fact. What you fear will not happen, in this imagined case, *never* happens. You want the person on Mars to be you in a specially intimate way in which

no future person will ever be you. This means that, judged from the standpoint of your natural beliefs, even ordinary survival is about as bad as teletransportation. *Ordinary survival is about as bad as being destroyed and having a Replica.*

How the Split-Brain Cases Support the Bundle Theory

The truth of the Bundle Theory seems to me, in the widest sense, as much a scientific as a philosophical conclusion. I can imagine kinds of evidence which would have justified believing in the existence of separately existing Egos, and believing that the continued existence of these Egos is what explains the continuity of each mental life. But there is in fact very little evidence in favour of this Ego Theory, and much for the alternative Bundle Theory.

Some of this evidence is provided by the split-brain cases. On the Ego Theory, to explain what unifies our experiences at any one time, we should simply claim that these are all experiences which are being had by the same person. Bundle Theorists reject this explanation. This disagreement is hard to resolve in ordinary cases. But consider the simplified split-brain case that I described. We show to my imagined patient a placard whose left half is blue and right half is red. In one of this person's two streams of consciousness, he is aware of seeing only blue, while at the same time, in his other stream, he is aware of seeing only red. Each of these two visual experiences is combined with other experiences, like that of being aware of moving one of his hands. What unifies the experiences, at any time, in each of this person's two streams of consciousness? What unifies his awareness of seeing only red with his awareness of moving one hand? The answer cannot be that these experiences are being had by the same person. This answer cannot explain the unity of each of this person's two streams of consciousness, since it ignores the disunity between these streams. This person is now having all of the experiences in

both of his two streams. If this fact was what unified these experiences, this would make the two streams one.

These cases do not, I have claimed, involve two people sharing a single body. Since there is only one person involved, who has two streams of consciousness, the Ego Theorist's explanation would have to take the following form. He would have to distinguish between persons and subjects of experiences, and claim that, in split-brain cases, there are *two* of the latter. What unifies the experiences in one of the person's two streams would have to be the fact that these experiences are all being had by the same subject of experiences. What unifies the experiences in this person's other stream would have to be the fact that they are being had by another subject of experiences. When this explanation takes this form, it becomes much less plausible. While we could assume that 'subject of experiences', or 'Ego', simply meant 'person', it was easy to believe that there are subjects of experiences. But if there can be subjects of experiences that are not persons, and if in the life of a split-brain patient there are at any time two different subjects of experiences—two different Egos—why should we believe that there really are such things? This does not amount to a refutation. But it seems to me a strong argument against the Ego Theory.

As a Bundle Theorist, I believe that these two Egos are idle cogs. There is another explanation of the unity of consciousness, both in ordinary cases and in split-brain cases. It is simply a fact that ordinary people are, at any time, aware of having several different experiences. This awareness of several different experiences can be helpfully compared with one's awareness, in short-term memory, of several different experiences. Just as there can be a single memory of just having had several experiences, such as hearing a bell strike three times, there can be a single state of awareness both of hearing the fourth striking of this bell, and of seeing, at the same time, ravens flying past the bell-tower.

Unlike the Ego Theorist's explanation, this explanation can easily be extended to cover split-brain cases. In such cases there is, at

any time, not one state of awareness of several different experiences, but two such states. In the case I described, there is one state of awareness of both seeing only red and of moving one hand, and there is another state of awareness of both seeing only blue and moving the other hand. In claiming that there are two such states of awareness, we are not postulating the existence of unfamiliar entities, two separately existing Egos which are not the same as the single person whom the case involves. This explanation appeals to a pair of mental states which would have to be described anyway in a full description of this case.

I have suggested how the split-brain cases provide one argument for one view about the nature of persons. I should mention another such argument, provided by an imagined extension of these cases, first discussed at length by David Wiggins.³

In this imagined case a person's brain is divided, and the two halves are transplanted into a pair of different bodies. The two resulting people live quite separate lives. This imagined case shows that personal identity is not what matters. If I was about to divide, I should conclude that neither of the resulting people will be me. I will have ceased to exist. But this way of ceasing to exist is about as good—or as bad—as ordinary survival.

Some of the features of Wiggins's imagined case are likely to remain technically impossible. But the case cannot be dismissed, since its most striking feature, the division of one stream of consciousness into separate streams, has already happened. This is a second way in which the actual split-brain cases have great theoretical importance. They challenge some of our deepest assumptions about ourselves.⁴

³ At the end of his *Identity and Spatio-temporal Continuity* (Oxford: Blackwell, 1967).

⁴ I discuss these assumptions further in part 3 of my *Reasons and Persons*.