

УДК 130.31

P.M.S. Hacker

ON HAVING A MIND, HAVING A BODY, AND BEING A PERSON*

The basic ideas of P.M.S. Hacker's writings are considered in the Introduction. The main subject of the considerations is Philosophy of the Human Being. The article of P.M.S. Hacker that is presented in English is devoted to the subject.

Keywords: mind, body, person, culture.

§1. *The emergence of the concept*

The concept of a person is central to our thought about ourselves, our nature, and our moral and legal relations. That human beings are persons is not a trivial tautology, but a fundamental claim about our moral status and our singularity in the order of nature. 'A *person*', Kant wrote, 'is a subject whose actions can be imputed to him. *Moral* personality is therefore nothing other than the freedom of a rational being under moral laws'. 'Whereas', he added, 'psychological personality is merely the ability to be conscious of one's identity in different conditions of one's existence'¹. No philosopher has placed greater emphasis on the *ethical* character of the concept of a person and on its essential link to the concepts of *freedom* (hence *reason*) and *responsibility*, or had greater influence in allocating such a pivotal role in the characterization of our conceptual scheme to the concept of a person thus conceived.

It was not always so. The concept of a person originated from the word *persona* – the mask (*prosopon*) worn by an actor in the ancient world to signify the role he was playing. Thence it was naturally extended to the role a human being has in life. In the writings of the Stoics, the idea of *persona* was associated with morality not via the notion of man as possessor of the powers of practical reason, but rather via that of social role – what would later become known as the ethics of 'my station and its duties'. Epictetus insisted that 'to play well the assigned role (*prosopon*) is your business but to choose the role, the business of another'. Similar sentiments are expressed by Cicero in *De Officiis*, where he compares the stage personages created by the playwright with the character (role, *persona*) that nature has assigned to each of us².

The association of *persona* with the notion of a role spread to the law courts. For the plaintiff and defendant are, by analogy with the actor on stage, playing clearly defined roles, and 'persona' became used for these. The expression was rapidly extended to encompass human beings standing in other characteristic legal

* This paper was delivered as a Zeno Lecture at the University of Utrecht on 23 November, 2007. For a much more detailed treatment of these themes, see my *Human Nature: the Categorical Framework*. Oxford: Blackwell, 2007.

¹ Kant. *The Metaphysics of Morals* 6:224.

² Cicero. *De Officiis*, I–XXVIII, cp. XXXI.

relations, in particular qua possessors of rights. So in the *Institutes* I §8 it is observed that ‘every right of which we treat pertains to persons or to things or to law-suits’, and §9 adds: ‘And indeed the highest division of persons in law is this, that all men are either freemen or slaves’. Here ‘person’ stands in contrast to things, and extends the original notion of *role* to signify *human beings* in general. (Thus used, the later Kantian proposition ‘Human beings are persons’ would be a trivial tautology.) However, Roman law rapidly evolved to exclude slaves from the category of persons. Justinian’s translators of the *Institutes* characterize slaves as *aprosopos*, and from Justinian’s codifications onwards only the legally qualified man is a person. Slaves were characterized as things, and a person was defined as a man of civil status.

Boethius, in the sixth century, defined ‘person’ as ‘an individual substance of rational nature’. This detached the concept of a person from that of a role, and associated it with the array of capacities Aristotle had emphasized as distinguishing humans from other animals. This conceptual nexus was adopted (and modified) by Aquinas and transmitted to early modern philosophy.

In philosophy of the modern era, the concept of a person has been the locus of much confusion (*vide* Descartes, Locke, Hume). What precisely is a person? how is a person related to the human being that is one? how is the person related to the mind and to the body that a human being is said to have?

§2. *Having a mind*

To grapple with these questions, one must first clarify what it is to have a mind and to have a body, and what warrants ascribing to a creature both a mind and a body. To clarify the concept of the mind, we need first to examine how the expression ‘the mind’ is used. English idioms of the mind are clustered around four foci:

(i) *Memory*: To hold something in mind is to retain it in memory; to keep something in mind ensures that one will not overlook it. To bear something in mind is to remember it, so that one can take it into account.

(ii) *Thought*: To have a thought cross one’s mind is for something to have occurred to one. Ideas come to mind, flash through one’s mind, or lurk, just out of reach, at the back of one’s mind – as one tries to think of something. To turn one’s mind to something is to begin thinking about it. To have something on one’s mind is to be preoccupied with it.

(iii) *Opinion*: To know one’s mind is to have formed one’s opinion, and to tell one’s mind is to express it. To be of one mind with another is to agree in opinion or judgement. To speak one’s mind or to let someone know one’s mind, is to say what one opines. To give someone a piece of one’s mind is to tell him, harshly, what one thinks of him.

(iv) *Intention*: To be minded or to have it in mind to do something is to be inclined or to intend to do it. To have half a mind to do something is to be tempted to do it, and to be in two minds is to waver between alternatives. To make up one’s mind is to decide, and to change one’s mind is to reverse one’s decision or judgement.

English idiom ramifies further. At the most general level, the mind is associated with intellectual faculties: a person is said to have a powerful, agile, subtle or devious mind if he is skilful, quick and ingenious at problem solving, or if his solu-

tions, plans and projects display subtlety and cunning. Hence too, it is connected with appropriate intellectual virtues and vices: a person has a tenacious, idle, vigorous, judicious or indecisive mind according to the manner in which he grapples with problems requiring reflection and according to the typical upshot of his reflections. A person is of sound mind if he retains his rational faculties, and out of his mind if he thinks, proposes or does things that are irrational. He is not in his right mind if he is distraught, and he has lost his mind if he is bereft of his rational faculties. Other uses fade off in other directions: we speak of people being broad- or narrow-minded, small- or petty-minded, having a dirty mind or a mind like a razor, displaying presence of mind or lacking peace of mind.

We can derive some morals from these reminders of linguistic idiom.

First, each such use is readily paraphrasable without the word 'mind'. All that is needed is a corresponding psychological predicate predicable of the human being. In this sense, reference to the mind is eliminable without loss. Talk of the mind is merely a convenient *façon de parler*, an oblique way of speaking about human faculties and their exercise. That is not to say there are no minds. On the contrary, it is to say that there are, only that they are not kinds of things. The mind is not a *nothing* – but it is not a *something* either.

Secondly, it is obvious that in the various idioms one is not speaking of one and the same

thing, called 'the mind'. When we say that someone changed his mind, that he has a dirty mind, and that he has turned his mind to a certain question, we do not imply that there is one thing: a mind, which has changed, is dirty and has been turned. We are indeed speaking of the very same thing, *namely the human being*, and from case to case, and phrase to phrase, we are saying different things *of* the human being.

Thirdly, it is evident that the question of whether the mind is identical with the brain, under one interpretation, becomes transparently absurd, since the mind is not a kind of entity that might be identical with anything.

Fourthly, the very distinctive range of idioms appertaining to the mind suggests that it is not because we have minds that psychological predicates apply to us, but rather because a fairly specific sub-set of psychological predicates apply to us that we can be said to have minds. That fish can see and hear, that birds can be frightened of a cat, that cats can want to catch a bird – all of which Descartes denied – does not show that he was wrong to deny that they have minds.

Fifthly, although we speak of someone's having a quick, agile, original mind, or a slow, idle, and dull one, it is a categorial mistake to suppose, as Reid (and others) did, that the mind is 'that in [man] which thinks, imagines, reasons, wills'. We ascribe agile, devious, imaginative minds to human beings, but do not, on the whole, speak of the mind as if it were an agent, although we do sometimes speak of it as if it were a patient. It is not my mind that thinks, imagines, reasons or wills, it is I. On the other hand, my mind may go blank, be in turmoil, or wander – but these are neither acts nor activities of the mind. It is not my mind that makes up its mind or decides, that changes its mind or reverses its decision, that has half a mind, or is inclined, to do something – it is I, a human being. It was a Cartesian confusion to ascribe the whole range of psychological attributes to the mind. That incoher-

ence is multiplied by present-day materialists identifying the mind with the brain, and ascribing the same range of predicates to the brain.

What then is the mind? If a definition is meant to provide a rule for the current use of a word, then there is no useful definition of 'the mind', although from idiom to idiom we can explain what is meant. But if we cannot fruitfully define 'mind' in *this* sense of 'definition', we can ask what must be true of a creature for it to be said to have a mind. It is evident that it is above all human beings that are the subjects of this array of idioms. We do sometimes extend the odd idiom to an animal, as when we say, of the horse or dog that will not do what we want it to do, that it has a mind of its own. But this, although it need not be crude anthropomorphizing, is just a natural extension of language from our own case.

What then is it about human beings that inclines us to invoke the panoply of the language of the mind? It is the ability to reason and to act for reasons. To have a mind is to have an intellect and a rational will. It is to be able to reason, to apprehend things as affording reasons for thinking, feeling and acting. It is to be able to deliberate, decide or choose what to do or believe, and to modify one's feelings and attitudes, in the light of reasons.

§3. *Having a body*

So, our customary talk of the mind is no more than an abstraction from our talk of human beings, of their intellectual, mnemonic and volitional abilities and their exercise. Our parallel talk of our body is in a like manner largely derivative from our talk of human beings. We have and exercise an array of cognitive, cogitative and volitional powers that we exhibit in our behaviour. As agents of this kind, we are said *to have a mind*. We likewise have a variety of corporeal determinables – a *physique*, a *figure*, physical *appearance*, susceptibility to *sensation*, degrees of *fitness* (ability to engage in *physical activity*) and *health*. As such agents, with a mind that is manifest in corporeal activities and reactions, we are also said *to have a body*, and can be described as having *such-and-such* a body (powerful, muscular, lithe, graceful). We describe ourselves thus when we wish to focus upon such attributes, and to manifest our *attitudes* to them. For we may be proud of our body (if it is fit and beautiful) or ashamed of it (if it is fat, flabby and ugly), admire the beautiful bodies of the young and be shocked by the mutilated bodies of the injured. We may be comfortable with our body, or ill at ease with it (e.g. if we are pathologically obsessed with our physical appearance and condition). Being self-conscious creatures, much preoccupied with sexual attraction and commonly given to degrees of hypochondria, it is hardly surprising that we should think a great deal about our physique, aesthetic appearance, attractiveness, our physical condition.

In the sense in which 'a body' is a material spatio-temporal continuant, then, of course, human beings *are* bodies – although this is a quasi-technical use of 'body', and not something that can be said unmisleadingly or inoffensively in extra-philosophical discourse. Material spatio-temporal continuants may be inanimate or animate. Those that are animate are organisms of varying degrees of biological complexity. Human beings are highly complex organisms, and in this sense too are bodies. Being a body in this sense, being an organism, implies being a body in the most general sense. We further speak of ourselves as *having* a healthy, ageing, weak or strong body. To be sure, only a creature that is a body in the previous

sense can thus be said to *have* a body. The moot question is: what has to be true of a creature that is an animate spatio-temporal continuant for it to be said to have such-and-such a body? After all, we do not say of trees or insects or fish that they have bodies; nor do we think of dead trees, insects or fish as corpses. The most coherent answer is: only if it can also be said to have such-and-such a mind. No doubt we sometimes extend the idiom to other animals, by analogy with ourselves. We may well say of an injured horse that its poor body was terribly lacerated, just as we may say of a stubborn horse that refused to do what we want it to do that it has a mind of its own. Such uses are harmless extensions, but they show little about the rich web of connections in the paradigmatic case. One might also say that the idiom of *having a body* is a formal mark of advanced sentient life that also leaves a corpse on death. It is no coincidence that human attitudes towards their dead are distinctive, and that the treatment of the corpse is informed by ritual.

So, talk of our body is no less derivative from, or parasitic on, talk of the human being as a whole than is talk of our mind. We say of another: 'She has a beautiful (athletic, clumsy, elegant, dumpy) body, sweet (sour, attractive, expressive, poker-) face, lively (dull, inquisitive, quick, slow) mind and kind (soft, hard, broken) heart'. Until we stop to think about it, it would not occur to anyone that we are doing anything other than talking about a single human being – specifying a variety of different modes that our complex species may exemplify in a variety of complex ways. But, like Augustine, once someone asks us, once we stop to think, we become entangled in the web of words.

What then is the relationship between the mind and the body? The mind is not an entity that could stand in a *relationship* to anything. As we have seen, all talk of the mind that a human being has and of its characteristics is talk of the intellectual and volitional powers that he has, and of their exercise. The body that a human being *is*, the living organism, has and exercises those distinctive intellectual and volitional abilities that we speak of when we speak of people's minds. But the body that a human being is said to *have*, when we speak of human beings as having beautiful or athletic bodies, is not the kind of thing that could be said to possess intellectual and volitional abilities. For we speak of a human being's body thus when we focus upon corporeal characteristics that the human being in question has. These characteristics, the very distinctive range of corporeal features of a human being that we assign to the body he has, are not the kinds of thing that could make up their minds, call things to mind or change their minds.

The mind a human being has and the body he has are not the kinds of things that could stand in *any* relationship to each other in the sense in which Jack and Jill (or London and Paris, or a man and his property) may stand in various relationships to each other. One might compare the question of how my mind is related to my body with the question of how the value of a £5 note is related to its colour. For here too one might explain that a £5 note is green and has a value of five pounds, but the colour green does not stand in any *relationship* to the value of five pounds. Green does not *have* a (monetary) value, the value of five pounds has no colour, and the £5 note itself does not stand in any *relationship* either to its value or to its colour. So too, my body does not have a mind (what would my body do with a

mind?). Moreover, it does not stand in a relationship to my mind either. *I* have a mind – and a body, but ‘having’ here does not signify a relationship.

§4. *Changing bodies and switching brains – puzzle cases and red-herrings*

Since time immemorial, human beings have fantasized about metempsychosis, the transmigration of souls. A variant of the idea is patent in Locke’s supposition of the intelligibility of the prince awakening ‘in the body’ of the cobbler, having retained all his memories. This thought has been given rein in fiction, both in Anstey’s *Vice Versa* and Conan Doyle’s short story ‘The Keinplatz Experiment’, and in more extreme form in fairy-tales, such as ‘Beauty and the Beast’, and surrealist fiction, such as Kafka’s ‘Metamorphosis’, where Gregor Samsa finds himself ‘in the body’ of a giant beetle. These fantasies are supposed to illustrate the transference of the mind, indeed, of the person, conceived as identical with his mind, from one body to another. Amusing or terrifying as these fictions are, it is doubtful whether they make sense. Like Escher etchings, it is not easy to see where the incoherence is rooted. That is part of their charm. In *this* sense of ‘imaginable’, imaginability is not a criterion of logical possibility. Whether the tale makes sense has to be submitted to at least as careful a scrutiny as Escher’s etchings demand. To do the former is considerably more difficult than to do the latter.

In the sense in which human beings *are* bodies, then it is clear that a human person can no more *change* bodies than can a house, a tree, a cat or a dog. Ascription of properties to the body of a living human being is merely an alternative way of describing the corporeal properties of the human person. In the sense in which human beings *have* bodies, then talk of the body a human being has is a *façon de parler* behind which lie the corporeal attributes of a human person. Hence there is a trivial sense in which a human being can change his body. If his body is fat, he can diet and lose weight – then he will have a thin body. If her body is ungainly and her nose unsightly, then she can have plastic surgery and emerge from it with a shapely body and an attractive face. So, in this sense, human beings *can* change their bodies, i.e. change from having a fat and flabby body to having a fit and athletic one.

Obviously, this is not what is meant. Could one’s mind not be transferred from one organism to another? Would this not involve changing one’s body? I can give another the same pain I have. If I have a stomach-ache from eating bad oysters, and I give you the same oysters to eat and you then have a stomach-ache that you describe in the same terms as I describe mine, then you have the same pain as I. I can give you my thoughts and beliefs by telling you what I think and believe and giving you good reasons for thinking and believing likewise. Then, if you are reasonable, you will have the same thoughts and beliefs. But that, no matter how extensive, involves nothing that could be described as ‘changing bodies’. For one thing, giving another person the pain I have no more means that I cease to have the same pain than persuading another to think what I think means that I cease to think thus. But even if I were to cease having the pain I have when I ensure that you have the same pain, or if I were to cease thinking what I persuade you to think, it would not follow that you were gradually becoming me – no matter how many of my mental attributes or modes of my consciousness you acquire in *this* sense. It would merely

mean that you were becoming more like me. But numerical identity is not a limiting case of similarity. Not even my psychological *Doppelgänger* is me.

Of course, when philosophers speak thus of transferring thoughts or memories from one person to another, they do not mean transferring *what* is thought or remembered. Rather, they entertain the fantasy of transferring the thinking or remembering. But it makes no sense to speak of transferring *my* thinking that *p*, *my* expecting that *q*, or *my* having a headache – understood as events, processes or states. For thus understood, these items are *dependent particulars*. The event of my feeling a twinge, as opposed to the twinge I feel (the sensation), is identifiability-dependent upon me – i.e. on the substance change in which is constitutive of the event in question. The event of another person feeling the very same twinge is a different event, for it is a different person undergoing change. One can no more transfer a person's experiences thus than one can *transfer* a person's smiling or gesturing.

Why can one person's mind not be transferred into another person's body? – After all, Locke, Anstey and Conan Doyle imagined it. The mind, we have argued, is not an entity of any kind, but an array of distinctive powers and their exercise. So it is not a kind of thing that could *migrate* or be *transferred* from one substance to another like an appendage; nor are character traits, tendencies and pronenesses of thought, feeling and behaviour that characterize individual human beings. It is possible to transform chemical elements, but, contrary to the alchemists of old, one cannot, *logically*, turn lead into gold by getting the yellowness of a quantity of sulphur and the shininess of brass to *migrate* into a bar of lead. The only sense in which my mind, understood as my mental attributes, can be 'transferred' is that one can bring it about that another has whatever psychological attribute one has oneself (I can, without difficulty, bring it about that you believe something that I believe, that you have the same pain as I have, and so forth) – and perhaps one may fantasize about bringing it about that another has *all* one's psychological attributes (perhaps at the cost of the destruction of one's own brain). But if this is (remotely) intelligible, then there is no reason why there should not be multiple 'transfer' – not to one other human being, but to any number. But if multiple transfer is logically possible, then identity is excluded¹.

Similarly, the body I have is not an entity that I inhabit or *occupy*. To speak of the body I have is to speak of a certain range of my corporeal characteristics – and characteristics are not 'inhabitable' – certainly not inhabitable by minds, which are neither inhabitants nor occupants. The features of the body I have are *my* corporeal features – features of *the body I am*. But I cannot *inhabit* what I *am*. Nor can I change the body I am save in the innocuous sense mentioned above. The thought that one 'inhabits' the body one has arises from failure to grasp the logical character of the idioms of *having a body*. If one thinks of oneself as *inhabiting one's body*, it becomes difficult to resist the temptation of supposing that one stands in a direct causal relation to one's body. But although one may be the *indirect* cause of one's corporeal features (e.g. one's body may be fat because one is self-indulgent, or athletic because one takes exercise; one's leg may hurt because one stubbed

¹ See Wiggins's 'Only *a* and *b* rule' in his *Sameness and Substance Renewed*. Cambridge: Cambridge University Press, 2001. P. 96ff.

one's toe, and one's body may be sunburnt because one spent too much time on the beach), one cannot be the *direct* cause of attributes of one's body. One may bring it about that one's body is covered in perspiration, or in good trim – but to bring something about is always a matter of indirect causation. One may bring it about that one's arm moves, but one does so by moving it with one's other hand or by pressing it against the wall and then standing back and letting it rise. But to raise one's arm is not to cause it to rise. If one further confuses the body one is with the body one has, i.e. confuses the two different forms of description, one may also think that one stands in a direct causal relation to the body one is. But that too is confused. For I am not the cause of my doing what I do. I may move, but in so doing I do not cause the body that I am, i.e. myself, to move. Of course, I may make myself do things I don't want to do – but that signifies determination in the face of weariness, reluctance and disinclination, and is not a matter of causing anything.

Most of the philosophical fantasies of the late twentieth-century differed from their ancestral forms, philosophical and fictional alike. Perhaps inchoately aware of the incoherences of 'transferring' memories and minds, perhaps convinced that *science* fiction is better than fiction, and more impressed by men in white coats than by God, the fantasy has involved transplanting brains, hemispheres of the brain, and other bits and pieces of cortical tissue. These seem to ensure a firm material, causal, basis for the transference of memories and other features encompassed by Locke's 'consciousness'. For it was this material base that was conspicuously absent in the tales thus far considered. Transferring hemispheres was proposed in jest by Swift in *Gulliver's Travels* (Part III, 'Voyage to Balnibarbi'). In the latter half of the twentieth century as much serious effort was spent by philosophers on speculating about brain transfers and hemispherectomies as the mediaevals are said to have spent on speculating how many angels can dance on the head of a pin, and to rather less purpose. How did this come about?

The idea of an immaterial substance lost favour among philosophers. Where Descartes had conceived of the person as identical with the mind, it now became common to identify a person with his brain. The dependency of all the characteristic functions of human beings on their brain inclined many to suppose that the mind *is* the brain or that it stands to the brain as the software to the hardware of a computer. We can lose our legs and remain the same person, lose our arms and remain the same person – the one thing we cannot lose is the brain (the limiting case of a mutilated human being). *Or so it seemed*. If the brain survives, appropriately nourished, in a vat – surely the person survives – *or so it was argued*. We are not, it was conceded, brains in vats – although we could be. But we are in fact *brains in skulls*. Philosophers, cognitive scientists, and cognitive neuroscientists proceeded to ascribe to the brain a large range of psychological predicates, e.g. to think, believe, compute, calculate, perceive, remember, want, decide. Hence ancient fantasies about transferring minds or souls from one body to another took a materialist turn. Surely, it was argued, we can imagine transplanting brains from one skull to another, and if the patient wakes up after the operation with all his memories and traits of character intact, then he is the same old person in a different body.

Brain transplants are biologically impossible. Let us waive that, since they are not *logically* impossible. Nevertheless, the science fiction does not show that if A's brain is transplanted into B's decerebrate skull, then the resultant person, let us call him Aby, is actually A. That would be true only if a person were an embodied (enskulled) brain. But the brain is a part of the human being – an organ, not an organism. It makes no sense to apply psychological predicates such as 'thinks', 'remembers', 'perceives', 'imagines', 'wants', 'decides', 'feels happy', and so forth, to the brain. It is a mistake to apply predicates that signify attributes of an animal as a living whole to its parts – a mereological (or homunculus) fallacy. Just as it is a mistake to ascribe the property of *keeping time* to the fusée of an eighteenth century table clock, or the property of *flying* to the engine of an aeroplane, so too it is a mistake to ascribe to a person's brain the property of *thinking that it is raining* or of *deciding to take an umbrella*. Aristotle sapiently wrote 'to say that the *psuchē* is angry is as if one were to say that the *psuchē* weaves or builds. For it is surely better not to say that the *psuchē* pities, learns or thinks, but that the man does these things with his *psuchē*'. This insight was echoed by Wittgenstein in *Investigations* §281 'only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind; hears; is deaf; is conscious or unconscious'. It is human beings, not their brains, that have a stomach-ache; it is human beings who put their eye to the keyhole to see what is in the locked room, who cock their ear to listen better, who sniff to smell – not their brains. It is human beings in the stream of life who manifest their powers in what they do and say, who exhibit their character traits in their behaviour. It makes no sense to ascribe such attributes to brains, for brains are not animals, and there is nothing that brains can do that could warrant ascribing such attributes to them.

What is true is that human beings would not have the powers they do but for their brains' functioning normally. It does not follow that the brain is the real possessor of those powers. We could not walk but for the normal functioning of our brain, but the brain is not the real walker. Nor is it the real thinker. Magritte might well have painted a brain on a pedestal and entitled it *Le Penseur* – but that would have been a joke. Human persons are not enskulled brains (they weigh more than 3 lbs, and are taller than 7 inches). Contrary to Descartes, we are not identical with our minds – but we are not identical with our brains either. Substituting the brain for the Cartesian mind remedies none of the deep conceptual incoherences in Cartesianism – it only conceals them under a deceptive veneer of misunderstood neuroscience. A brain removed from a skull is not a limiting case of a mutilated human being.

Could we not coherently imagine a living brain in a vat, connected up to various bits of electronic hardware, seeing what is going on around it in the laboratory by means of prosthetic eyes, and responding to our questions by means of a computerized voice? Perhaps – but then it would not be the brain that sees, that reflects before it speaks, and speaks thoughtfully. It would be the complex being consisting of the living brain of a dead human being and the electronic machinery to which it is attached – a cerebroid. If one removes A's cortex, leaving his thalamus, basal ganglia, cerebellum and brainstem intact, then A will be in an irreversible vegetative state. If one removes A's brain from his skull, then A is dead – as is evi-

dent if one examines A's remains on the operating table. Suppose that one could, miraculously, insert A's brain into B's skull, and that when Aby awoke, he asserted that his name was 'A', that he recollects doing a variety of things (that we know A did), that he has various plans for the future (that we know A had). Does this not show that Aby is A? No – no more so than if B had been hypnotized to forget everything about his own life and been told that he was the subject of the biography that we know to be A's. But suppose that he also appeared to recognize all the people that A knew, and to know his way about all the places that A frequented – would that not show that Aby is really A? If you woke up tomorrow and seemed to recognize all the people and places I know, seemed to recollect everything I recollect – would that make you me? The transplantation of cerebral tissue is, so to speak, a red-herring – we are so mesmerized by the bogus science that we follow a false trail.

What is true is that our concept of a person is tailored for creatures like us. We can coherently imagine all manner of circumstances in which it would lose its grip, circumstances under which we would no longer know what to say. The Jekyll and Hyde phenomenon might become common. Multiple personality cases, with no mnemonic connection between the different personalities, might become ubiquitous. Pairs of human beings might find that they, as it were, alternate personalities and memories every day. Complete 'memory portfolios' (i.e. what is remembered, not the remembering of it) might be transferable from one human being to another at the press of a button. And so on. Our concept of a person would not be very useful to such creatures. But all this shows is that the concept of a person, like all our concepts, is of utility only against a background of constancies.

There are some methodological morals to this tale of philosophical confusion that resulted from science-fictional prestidigitations. In neuroscience, the investigation of brain injuries leading to varying degrees of catastrophic impairment of human faculties is a primary method of investigating the brain and its activities. Breakdowns shed light upon the normal case. The interminable debate within late twentieth-century Anglophone philosophy about brain transplants and personal identity laboured under a serious illusion. It was assumed that in *conceptual* investigations light can be thrown upon a problematic concept by close examination of cases in which the application of the concept becomes questionable. But that is mistaken. We cannot clarify the normal use of a word, and shed light on the concept it expresses, by looking at its abnormal use – at its contentious and problematic application to extreme cases that never occur. The most that we can hope to discover, if we are fortunate enough not to become confused, are various breakdown conditions. In the extreme case, neither the conditions for applying the predicate nor the conditions for withholding it are satisfied. So, for example, in cases of so called blindsight one can neither say that the patient sees nor that he does not see. This does not clarify the concept of vision, but only describes conditions under which it crumbles.

Similarly, the debates about brain transplants involved extensive consulting of what goes under the name of 'intuitions'. We are supposed to contemplate various empirically impossible circumstances of brain transplants and to consult our intuitions as to whether the resultant individual is identical with the person from whose

body the transplanted brain was removed. It might be transplanted into the skull of his twin brother, or into that of a man of his age, or into the body of the emperor of China, or the body of a woman, a little boy or girl – and by consulting our intuitions as to whether the resultant person is identical with the person from whom the brain was taken, we shall obtain valuable data that will clarify the nature of personal identity.

It is far from obvious what ‘intuitions’ are supposed to be. Are they simply hunches? Or guesses? And why should hunches about truth be of any value in philosophy? Why should guesses about how to apply a rule to circumstances for which it was not designed be supposed to shed light on its application to cases for which it was designed? Is there any truth of the matter here to guess? It is one thing to determine whether someone has mastered a concept by asking him to decide its applicability to normal (logically possible) cases, but quite another to try to determine whether a problematic case is indeed logically possible, i.e. whether it makes sense, by asking someone who has mastered the concept to decide on its applicability in such abnormal, imaginary cases by consulting his intuitions.

Light can be shed upon a concept by careful examination of the correct use of the word that expresses that concept in normal conditions of its application – not by consulting one’s hunches about the truth or falsity of its application in abnormal cases. There is no room for intuitions in determination of a concept. And a competent speaker’s judgement about the correct use of a word in normal circumstances is no hunch or guess, any more than is the competent chess player’s judgement about the rules of chess – he *knows* how the game is played. If in abnormal scenarios, such as science fiction fantasies, we do not know whether to apply or withhold an expression (e.g. ‘same person as’), that is not a matter of ignorance that could be remedied by consulting one’s intuitions. It is itself a *result* – namely that the rules for the use of the expression do not go this far. What is then needed is not conceptual elucidation but reasoned legislation – which, for science fiction fantasy, is pointless.

§5. *The concept of a person*

The seventeenth and eighteenth century debate provided the framework within which the concept of a person was discussed by most philosophers, including analytical philosophers, until the Second World War and beyond. That framework was radically defective, and it continues to be at the root of extensive confusions, some of which we have just touched upon. To avoid these incoherences, one must start afresh, rather than trying to weave the rotten strands of the Cartesian-empiricist tradition into yet another threadbare tapestry. Informed by our previous examination of the concepts of the mind and the body that human beings have, and well apprised now of the pitfalls that lie in the path, we shall start with a reminder of some commonplaces. Human beings are the only persons we know or are ever likely to know. Our concept of a person evolved above all as a concept applicable to human beings – social beings who are members of a moral community. It is clear enough that it is not a concept belonging to the physical sciences that concern themselves with matter in motion; but that does not mean that it belongs to a science of immaterial substances. It is not a concept that finds its place in the biological sciences that study the behaviour of animals in general. It belongs to the moral

sciences – to the study of man as a moral, social and cultural being. And it is at home in our daily discourse about ourselves and our fellow human beings.

Human beings are living organisms of a given type. We are language-using, culture-creating, self-conscious creatures, that have a mind and a body. The latter peculiar mode of self-characterization, we have argued, is no more than a manner of talking of distinctive attributes of intellect, will and affection consequent on being a language-using animal, on the one hand, and of corporeal characteristics pertaining to appearance, physique, health, and sensation, on the other. We are not two one-sided things, as Descartes suggested. But nor are we, as Strawson suggested, one two-sided thing – like a penny. We are multi-faceted, like cut stones – partly opaque and riven with flaws. Being self-moving creatures with cognitive and volitional two-way powers, we can voluntarily act, take action and engage in activities. We could not be thinkers unless we were also doers, could not be observers unless we were also experimenters, could not be spectators of the passing scene unless we were also players in the unfolding play.

Being rational, we can reason and act for reasons. So we have intentions, plans and projects that we pursue. Having a language, our cognitive powers endow us with the ability to retain the complex forms of knowledge that we can and do acquire. This includes experiential or personal memory. So we possess, as it were, an autobiography – we can tell the tale of our life as we remember it. Each human being traces a unique autobiographical route through the world, and the combination of genetic endowment, variable responsiveness to individual experience, and memory gives each human being a *personality* – with a unique combination of character traits, behavioural tendencies, an awareness of a unique past (pertaining both to inner and outer life) and of projects for the future. Our consciousness of our past itself typically incorporates an awareness of our family and social group, and of the form of life that contributes to the generation of our social identity and plays a crucial role in our conception of ourselves. So we have a sense (weaker or stronger) of our own identity, which may be more or less individualistic, or more or less tribal.

Being animate, we have needs and a good. Being rational, we can act for the reason that something is beneficial or detrimental, either for ourselves or for others. Being social, we grow to maturity within a community bound by norms of conduct and by common values. Being capable of reasoning about what is right or good, we can judge shared norms and values as reasonable or unreasonable, genuine or perverted. So we have, or are normally capable of having, knowledge of good and evil. So we are capable of doing good and evil. So we are persons, bearers of duties and obligations, and of rights and related normative powers. Because we internalize norms of conduct and can exercise our power of reasoning concerning the right and the good, we (normally) possess a conscience. A corollary of this is that we are subject to moral emotions of self-assessment, such as guilt, shame and remorse, as well as other-regarding emotions, such as resentment and indignation. We are answerable for our deeds. A condition of answerability is the possibility of explanation and self-understanding in rational terms, i.e. by reference to the categories of reason-giving explanations.

The concept of a person is not a substance-concept, as is the concept of a human being. It has been said by Peter Strawson to be a 'primitive concept'¹. In so far as this means that it is not analysable as a combination of a subject of predicates of states of consciousness (P-predicates) and a different subject of corporeal predicates (M-predicates), i.e. of a mind and a body thus conceived, then it is surely primitive. For it is one and the same thing that is the subject of both kinds of predicate. But, of course, P-predicates are ascribable to a large part of the animal kingdom, to a large range of different substances (elephants, apes, monkeys, cats, dogs, etc.), none of whom, save for human beings, are persons. So, one may add that persons, unlike other developed members of the animal kingdom, are not only subjects to whom P- and M-predicates are ascribable, they are also *self*-ascribers of P-predicates (and M-predicates). Hence they are self-conscious, language-using creatures. This is true (of naturally developed, undamaged persons) – but then so too is much else characteristic of human beings that flows from mastery of a language. And it is doubtful whether the capacity for self-ascription of P-predicates can be hived off from the numerous other powers characteristic of persons that we have discussed above. Furthermore, the dichotomous division of predicates into P- and M-predicates is overly Cartesian. Many predicates ascribable to plants cannot be ascribed to inanimate objects (e.g. grows to maturity, flourishes, produces seed, dies), and others cannot be applied to animals either (e.g. grows in acid soil, bears fruit in autumn, has pink blossom). Many classes of predicates applicable to animals are inapplicable to plants (perceives, feels pain and pleasure, has desires, runs swiftly), and others are inapplicable to any known persons (e.g. has eight tentacles, lays a clutch of five eggs, has bright green plumage, hibernates, has kittens). The class of predicates uniquely applicable to human persons ranges far wider than the domain of psychological predicates, and many of the sub-class of psychological predicates that human beings apply to themselves without criteria are applicable to other animals on the basis of similar criteria as they are applied to other humans. It is true that human beings are subjects of P- and M-predicates and uniquely self-ascribers thereof. But it is perhaps more illuminating to say with Aristotle that we are unique possessors of a *rational psuchē*, over and above the powers of the sensitive *psuchē* that we share with other animals, and of the vegetative *psuchē* that we share with all living things.

Although we know of no persons other than human beings, usage does not restrict the application of this categorial term to human beings alone. If there be other creatures possessing the appropriate range of language-dependent rational powers grafted on to an appropriate animal nature, then they too are persons. To be a person is not to be a certain kind of animal, but rather to be an animal of one kind or another with certain kinds of abilities. The nature of a person is rooted in animality, but transformed by possession of intellect and will. So the concept of a person qualifies a substance-concept of an animal of such-and-such a kind, earmarking the individual of the relevant kind as possessing (or as being of such a nature as normally to possess) a distinctive range of powers, a personality, and the status of a moral being.

¹ Strawson P.F. *Individuals*. London: Methuen, 1959. P. 101f.

Locke was mistaken in supposing that experiential memory is a necessary condition of personal identity (since amnesia does not exclude continued identity). But since the concept of a person is linked with that of a rational being possessed of intellect and will, memory in general and experiential memory in particular are among the faculties such a being must, unless damaged, be endowed with. Locke confused this constitutive feature with a necessary condition of personal identity, which it is not¹. Only if a creature can retain knowledge acquired, can it act for the reason that things are as it previously learnt they are. And only if such a creature can recollect its doings and undergoings, can it learn from its own experience, amass first-hand evidence, reason from its past experience to anticipation of future eventualities, and act for the reason that things were thus and so with it previously. Without experiential memory, we would have no genuine autobiography to tell, but at best only factual and secondhand knowledge of our past. We would be like victims of severe aphasia who had to learn about their lives from records and tales of others, and (from day to day) could never remember learning it. We could not enjoy the intimate personal relationships that turn on shared experience recollected, on reciprocities and debts of reciprocity. Experiential memory is an open-ended matter – as one event is recollected in shared reminiscence, others flood forth, sometimes in delight and merriment, sometimes in embarrassment or shame. Shared experience recollected is constitutive of shared lives. Those with whom we share experience that we thus recollect are, in a deep sense, a part of our lives. A person's *sense of identity* depends on his remembering a substantial part of his autobiography. It is not only normal, but the *norm*, for a person to have a sense of his own identity and a more or less substantial autobiography.

It is the norm that human persons do not completely change their looks and characteristic manner of expression, gesture, voice and intonation contour. It is the norm that human persons have a determinate personality, manifest in action and reaction, that recognisably persists throughout change. It is the norm that human beings possess rational powers of thought and action, have emotions that are to a greater or lesser degree determined by reasons and to a greater or lesser degree amenable to reason. Being by nature social creatures, being born into families and growing up under parental guidance, human beings naturally exist within a web of social and moral relations and commitments. All this, and much more, is familiar platitude – but it is platitude that needs to be borne in mind in reflection on the concept of a person. These complex regularities and these typical entanglements characteristic of humanity are the background against which we identify human persons. These features provide the characteristic grounds for reidentifying human persons, as we do, on the basis of their appearance, voice, behaviour and recollections. These normal persistences are part of what gives point to the concept of a person that has evolved to meet our needs.

St John's College, Oxford

¹ See Wiggins. *Sameness and Substance Renewed*, chap. 7.