An inaugural lecture given at University College London on May 13 2004. © Tim Crane 2004. Please do not quote without permission.

The Soul

Tim Crane Department of Philosophy, UCL

Michael Dummett says in the preface to his book on Frege that he is always disappointed when a book lacks a preface. 'it is like arriving at someone's house for dinner' Dummett says 'and being conducted straight into the dining room'. I feel the same way about inaugural lectures. To give an inaugural lecture is in part an acknowledgement of a professional honour, and in part an opportunity to pay a personal tribute to the institution which has honoured you in this way. It is not difficult, and a pleasant task, to do this.

My professorship has no predecessor, of course, but I hope that this does not disqualify me from saying something about what I owe to UCL and to its philosophy department. The intellectual character of the department as it is now was largely shaped by the influence of the late Richard Wollheim. I am sorry to say that I did not know Richard Wollheim well, and it is a cause of great sadness for the whole department that Richard was never able to return to the department as he had planned to do before he died last autumn. But I nonetheless feel the influence he left in the department, and I would like to pay a small tribute to it here.

When I first came to UCL in 1990, one of my colleagues was Bill Hart, who had just published his under-rated book, *The Engines of the Soul*. Hart's book was highly unorthodox and did not persuade many people. But it seems to me that in its boldness, disregard of current fashions and independence of mind it represented some

¹ Michael Dummett, Frege ix.

values which the UCL philosophy department has stood for over the years. In his 1984 book *The Thread of Life* Wollheim wrote that the UCL philosophy department had always 'exemplified to a high degree the values that happen to please [him] most: audacity, toleration, a concern for tradition and a disregard for authority'. In his own book, Hart says that Wollheim created a 'department that encouraged serious thought untrammelled by narrow conventions and intellectual piety'. Under the subsequent leadership of Ted Honderich, Malcolm Budd and Jo Wolff, the department has continued to do this. It has been a great privilege – and a source of pleasure and intellectual stimulation – to belong to this department, and I hope to remain in it for as long as I can.

Bill Hart's book, *The Engines of the Soul*, was a defence of the tremendously unpopular doctrine of Cartesian dualism – Descartes's view that there are two fundamental kinds of thing in the world: minds (or souls) and body. Philosophers do not agree on much; but I would guess that most philosophers today are convinced that Cartesian dualism is false. And although most of them agree that it is false, they often disagree on *why* it is false. Some think that it is false because of what we have learned from modern science; others think it is because the view is in some way conceptually incoherent. Still others think is not merely false but 'disastrous' or 'pernicious' – but being false seems to me a big enough drawback in a theory that we can ignore these more moralistic accusations. The fact is that these days philosophers tend to be *materialists* or *physicalists*: they think the world is entirely physical in its nature.

² Quoted in W.D. Hart, *Motions of the Mind* in Savile and Hopkins: 220

Materialists reject Cartesian dualism, but what exactly is Cartesian dualism? The central idea is the idea of *substance*. This idea derives from Aristotle: substances are the most fundamental things which exist. Substances are, in some sense, the ultimate entities, the things on whose existence the existence of everything else depends. On the Aristotelian view, properties or attributes, like the colour or weight of something for example, are dependent for their existence on the substances which have these attributes. Although one of Descartes's aims in philosophy was to overthrow the Aristotelian world picture, he nonetheless retained the concept of substance, and modified it in his own special way. For Descartes, a substance is 'a thing which exists in such a way that it does not need anything else in order for it to exist'. Strictly speaking, then, there is only one substance – God – since God is the only thing which needs nothing else in order to exist. But we can distinguish further among those things which only depend on God, and divide them into two: there is the material world (material substance) and there are minds or souls (mental substances). There is only one material substance – the material world itself – but there are many mental substances. Each kind of substance has its own characteristic attribute: extension in length, breadth, and depth (that is to say, shape and size) is the characteristic attribute of matter, and thought is the characteristic attribute of mind. Descartes's *dualism* is the view that these two very different kinds of substance constitute our world.

These facts about Descartes's philosophy will be very familiar to most people here, and the deficiencies of Descartes's dualism have been exhaustively treated in most textbooks in the philosophy of mind. I don't want to labour any of these familiar

³ Principles of Philosophy section 51.

points here, but I do want to spend a little time discussing the original argument for dualism, since this will help to introduce my main theme.

In *The Engines of the Soul* Bill Hart explicitly acknowledged that the argument he used was a version of Descartes's argument in his *Meditations*. The argument is striking simple:

- (1) I can imagine being disembodied: that is, existing independently from my body
- (2) What is imaginable is possible
- (3) Therefore: it is possible for me to be disembodied
- (4) Therefore: my body and I are distinct substances

There are various ways in which one might defend the assumption that one can imagine being disembodied: maybe one can imagine having the experience described famously by A.J. Ayer, of looking down at one's body on an operating table; or one could imagine being like a ghost, floating through doors and windows.

Many people will agree that one can imagine these things, but nonetheless challenge the principle that what is imaginable is possible, in Hume's famous phrase 'nothing we can imagine is absolutely impossible'. Critics have pointed out that it is incredible that what is genuinely and objectively possible should depend on our powers of imagination. Here a defender of the Cartesian argument should respond that the principle is not about what *makes* something possible but about how we *know* what is possible. Our imagination is a guide to what is possible, just as perception is a guide to what is actual.

Now, even if we accept that imaginability is our best guide to what is possible, and also that we can imagine being disembodied, there is still a gap to the conclusion

⁴ Hume, *Treatise* 32.

that it is really possible for me to be disembodied, since this may be a case where our best route to knowledge simply is not good enough to guarantee that we have latched onto a real possibility. However, what really interests me here is a further gap in the argument: how we get from the claim that it is possible for me to be separable from my body to the conclusion that my mind and my body are actually *separate*.

Of course, since identity is a necessary relation, then if it is possible for A to exist without B, A and B cannot be identical. But non-identity is one thing, entire separateness another. What supports the move in the argument from separability to separateness is the Cartesian conception of substance as something which is capable of independent existence. If my mind and my body are capable of existing independently of one another, as the possibility of disembodied existence shows, then they must be separate substances, on this Cartesian conception. But the Cartesian conception of substance is deeply problematic, and we should not accept it. Descartes's view of substance implies that something cannot be a genuine substance if it necessarily depends for its existence on something else. For if it did depend necessarily on something else, then it would not be possible for it to exist independently of that thing. But there seem to be good reasons for thinking that some kinds of thing can depend for their existence on something else, while still remaining distinct things, fundamental kinds of entity which are unities in themselves (substances, in other words). If Saul Kripke is right, for example, that each human being has an individual essence which is determined by their actual biological origin, then none of us could exist independently of our parents. I depend metaphysically for my existence on my parents – though this is not the only reason why they are here tonight – and yet I am a substance if anything is. Maybe Kripke's view is not correct, but to refute it merely by a definition of substance seems too easy.

We should reject the Cartesian conception of substance, then. If we want to continue talking in terms of substance, we should try and approach this notion (and therefore the notion of the soul) from another direction. But before getting to this, I would like to make one more comment about the Cartesian argument for dualism. If we reject the Cartesian view of substance, what should we say about the apparent fact that I can imagine being disembodied? Why doesn't this carry some weight against the view that I am necessarily an embodied thing? Here I think the right thing to say is that although it may be true that I can imagine an experience (of a certain kind) existing without a body, all this could show – if it shows anything at all – is that experiences of some kind can exist independently of bodies. It does not show that my experiences can exist independently of my body. This conclusion, that disembodied experiences of some kind are possible, is innocuous: since few materialist philosophers think that materialism is a necessary truth, many of them will all allow that experiences of some kind can exist without a body, in some remote possible world. What they will not allow is that the experience I am currently having *now* could exist without a body; for this experience is dependent for its existence on me, and there are good reasons (for example, Kripke's reasons just mentioned) for thinking that I am essentially embodied.

Recent discussions of these so-called 'conceivability' arguments for dualism have tended to avoid the issue of substance. If what I have just said is right, then they are wise to do so. For the conceivability argument cannot establish the separability of mental and material substances unless there is something like Descartes's conception of substance in the background. Take, for example, David Chalmers's recent and much-discussed version of the conceivability argument. Chalmers addresses the issue of the distinctness of the mental and the physical not by arguing for the possibility of

disembodiment, but by arguing for the reverse: a body identical to ours with no consciousness whatsover. He calls such a body a 'zombie' and the resulting argument is known as the zombie argument. A simplified version of the argument goes like this:

- (1) Zombies are imaginable
- (2) What is imaginable is possible
- (3) Therefore zombies are possible
- (4) If zombies are possible, then materialism is false
- (5) Therefore materialism is false

The reason why the conclusion follows is that materialism holds that physically identical creatures must be mentally identical too. So if there could be creatures physically identical to us who are lacking consciousness, then materialism is false. Whatever the merits of this argument, it clearly does not show that I am a mental substance. All it shows is that the physical *properties* or *attributes* I have do not determine my mental properties or attributes: specifically conscious properties (sometimes called 'qualia').

Our two dualists, then, differ in their targets. Descartes was concerned to demonstrate what he called the 'real distinction' between mental substances (souls) and material substances; Chalmers and his contemporary followers are concerned only to demonstrate that not all properties or attributes are determined by physical properties. Chalmers would be the first to admit that his view does not imply Descartes's; in fact, he might see this as an advantage of his view. So the latest arguments against materialism do not go any way towards establishing Descartes's conclusions about the soul: they do not establish that I could be disembodied, rather than the weaker conclusion that mental and physical properties are separable in various ways.

Bill Hart's use of the Cartesian argument is certainly audacious, and it shows a respect for tradition and a disregard for the authority of orthodox philosophical opinion. But despite what we might call these UCL virtues, I am not convinced by the argument. Even putting to one side the difficulties to do with the problematic relationship between conceivability and possibility, the argument presupposes the Cartesian conception of substance, and this conception is unacceptable.

At this point some of you might feel that you were brought here under false pretences. You came to a lecture on the soul and you find me arguing against it. This also might confirm your suspicions about inaugural lectures in philosophy: they are never about what they say they are going to be about. Some of us here will remember Malcolm Budd's brilliant inaugural lecture in UCL, called 'On looking at a picture'. Some people came to the lecture thinking they would spend a restful and edifying hour looking at beautiful paintings by Poussin or Titian; instead they got a few grainy black-and-white slides and some intense analysis of angles subtended at the eye and non-conceptual content. They were disappointed.

I realise I am not going to satisfy those who believe Cartesian dualism or an immortal, immaterial soul. And those who are already convinced materialists will not need further persuasion. But does this rejection of dualism mean there is no place in contemporary philosophy for some idea of the soul? Wittgenstein once said that the human body is the best picture of the human soul. Since he plainly did not believe in the Cartesian soul, he must have had something else in mind. Was he just speaking loosely or metaphorically, or is there more going on?

By the end of this lecture, I will have suggested an answer to this question.

But I suspect that there might be another source of unease with what I have just said.

Maybe some people, on learning the title of this lecture and hearing what I have said

so far, will think that there is something misconceived about the whole discussion. Talk about the soul is really talk about the human mind. And in talking about the human mind, they will say, we are really talking about the human brain. Science has shown us that all so-called mental capacities reside in the brain, so if we want to understand these capacities, we should simply study the brain. Talk about 'the mind' as such, or 'the soul' is a relic of our pre-scientific age, and should be abandoned.

There are a number of sources of this idea. One is the doctrine known as physicalism, which is the contemporary version of the materialist view that everything is material, the version which gives the ultimate authority to physical science in telling us what there is. Physicalism can be understood in many ways. One extreme version is encapsulated in a famous phrase of Rutherford's: 'there is physics; and there is stamp-collecting'. A less extreme, and therefore more believable version, is that whatever facts there are, they have to be determined entirely by the purely physical facts (facts about atoms and molecules and so on). To use a theological image: all God would have to do to create this world, just as it is, is to create its physical nature. Everything else, including the mental realm, would come for free.

Some scientists and philosophers have drawn extreme conclusions from this dependence of the mental on the physical. Ten years ago Francis Crick defended something he called the 'astonishing hypothesis': essentially the view that materialism is true. Understood in its most charitable way, the hypothesis seems to be little more than an expression of the dependence of mental capacities on the brain, something that is not especially astonishing and would not have astonished earlier materialists like Thomas Hobbes, for example. The fact that mental capacities are dependent on the brain does not imply that they are not real, or that talk about mental capacities is in some other way illegitimate. Given this obvious fact, even physicalists

will then have to say what makes the mental special: what makes certain physical things mental. (In a paper we wrote some years ago, Hugh Mellor and I made this point by saying that even if it is true that all flesh is grass, this does not tell us the difference between carnivores and vegetarians.)

So let's put physicalism to one side; in its most plausible version it clearly does not make talk about the mental illegitimate or redundant. What's more, this most anodyne version of physicalism does not imply the view known pejoratively as 'scientism': that the only real knowledge of the world is the knowledge we gain from natural science. If scientism implies that we have no knowledge of the mind other than the kind we get from psychology and neuroscience, then scientism is false. For it is plain that we already know a lot about our own minds and the minds of others without embarking on any scientific study of the mind. The real question is not whether we have any knowledge of the mind – psychological knowledge – which is not part of science, but rather what this knowledge is, how we obtain it, and how it is related to the knowledge we get from science. (It goes without saying – and therefore should be said, like everything that goes without saying – that opposition to scientism is not opposition to science.)

We should, I think, resist the reductionist tendencies present in both the physicalist and the scientistic views of reality and our knowledge of it. One of the things which the philosophy of mind attempts to do is to outline some of the very general and asbtract elements of our psychological self-conception. The hope is that in doing this, we can uncover some of the fundamental concepts we need in order to say what kinds of beings we are. One of these concepts, for example, is the concept of intentionality, or mental representation, or the mind's direction upon its objects. I agree with those philosophers who think that a general conception of the mental

faculties – of perception, thought, sensation, will and emotion – must depend upon a general conception of mental phenomena as intentional. Yet I think it is a mistake to suppose we have not accounted for these phenomena until we have accounted for them in physicalistic, causal, evolutionary or other reductionist terms. This is not to say that discoveries about the underlying mechanisms of the mind and their history are always irrelevant to our understanding of mental phenomena; the point is only that our belief in the reality of these phenomena should not depend on there being such discoveries. And nor does this attitude imply any opposition towards the very idea of reductive explanation. Any genuine explanation is an advance in our knowledge; so if we manage to obtain a reductive explanation of some phenomenon, then this will be an advance in our knowledge too. My hostility is only towards the idea that we must always aim for reduction 'in principle' and the idea that only a reductive explanation of the mind can vindicate our belief in it.

Let's return to the idea of the soul. Remember that I rejected the Cartesian conception of the soul as an immaterial substance because I rejected the conceivability argument; and I rejected the conceivability argument because I rejected the Cartesian conception of substance. Does this mean that the concept of substance – or even the concept of a *mental* substance – is not one of the fundamental concepts which we need in order to understand what kind of beings we are? This depends of course on what concept of substance we have in mind. When Aristotle first introduced the idea of substance into philosophy, it was defined as the subject of predications – the thing of which other things can be said – which is not itself predicated of other things. In this way substances are distinguished from attributes – like being musical or being pale – since attributes *are* predicated of other things. The paradigm examples of substances – what Aristotle called 'primary substances' – are living things like human

beings and other animals. Substances are things which persist through change: what changes a substance can allow while still remaining in existence is determined by the fundamental kind of thing the substance is: whether it is a pig or a goat or whatever. This is the reason that Aristotle says that substances contain a 'source of change' or 'principle of activity' within them: the principle belonging to the kind of thing they are.⁵

Even from these brief remarks, we can see that Aristotle's concept of substance differs from Descartes's rather more abstract idea of a thing with its characteristic attribute, which depends on nothing else for its existence. Descartes's conception of substance, although it has its roots in Aristotle, is stripped of some of the things which make Aristotle interesting to us now. Descartes's conception was challenged in his own century by Leibniz. Descartes believed that the material world is one substance, and that what we think of as particular material objects within the material world are parts of that substance. This has the consequence that for Descartes (unlike for Aristotle) a particular pig, for example, is not a substance. Leibniz argued that Descartes could give no account of the real unity of a substance, of why it is that substances are genuinely unified beings. A substance must have a unity which a mere aggregation of things, like a pile of stones or a flock of birds, does not. Leibniz's view was that a substance must have a principle of unity, and that this principle must be something which explains why things happen to a substance in the way it does: 'all its actions come from its own depths'.

Leibniz drove his ideas to the conclusion that the only genuine substances are souls – which he called 'monads' – and that the apparently substantial unities of the world were a kind of appearance dependent on this. Few will want to follow him

⁵ Metaphysics Book 5: 1015a11

down this path today. Nonetheless, the idea that our world contains real substantial unities, not simply aggregates which are 'given their final touch of unity by ... thought and perception' (as Leibniz put it 6) – *this* idea is one which has had a hard time of it in contemporary metaphysics. In dispensing with Leibniz's monads and the reasoning which led to them, philosophers have been too quick to dispense too with one of its guiding assumptions: that there are substances with their own unity.

In recent years, metaphysicians have struggled with the problem of how things constitute other things. What makes up an object? How do the parts of objects get together to make one object? One might think that these questions would be answered by talking about the physical forces and chemical bonds which hold things together. However, this would give us too many objects: not only would this lectern be an object, but the lectern plus anything which was firmly fixed to it would make a new object. Yet sticking this computer, say, to the lectern just gives us an assemblage of objects – surely nothing new has come into the world when we do this. But somehow putting the parts of the *lectern* together did create a new object. So what kinds of adhesion or bonding – or more generally, composition – make a genuine object? Various proposals have been made and some extreme proposals defended. Peter van Inwagen, for instance, has answered this question by saying that only living things and their ultimate simple parts exist. So there is no lectern after all; and so the problem of how its parts compose it disappears.

In response to views like this, some go to the other extreme and say that any 'fusion' of any objects with any others itself forms an object. These philosophers follow a principle of 'unrestricted composition': that any arbitrary collection of objects makes up an object. David Lewis once told the story of someone who bought

⁶ New Essays on Human Understanding II.xxiv.1.

a six-pack of beer and was not sure whether he was able to go through the 'five items or fewer' counter. Is the six-pack six things or one? Or is it seven (the six-pack itself plus the six individual cans)? Unrestricted composition says there are all these things and more. But if we think like this, we will be at a loss to recover the genuine unities in the world. One approach gives us too few real objects; another gives us too many. The only way out of this dilemma, it seems to me, is to reject the principle of unrestricted composition and simply deny that there is any general answer at all to the question, 'what makes up an object?'. The mistake is to think that the category 'object' divides up the world at all; what our world contains is *kinds* of object, like pigs or people or prams and pushchairs, and there will be a different answer to the question of what makes these things up for all the different kinds that there are. If we could use only the notion of an 'object' then we would be at a loss to find the genuine unities in the world. The right response to this is to not to give up on unity, but to give up asking the question about objects.

These dogmatic remarks are intended merely as propaganda in defence of a more traditional, more Aristotelian conception of substances as the ultimate, unified subjects of predication. When it is mental capacities which are predicated, what is the substance which is the bearer of these capacities? The familiar Cartesian answer to this question is one answer; but what are the alternatives? The prevailing materialism of the day sometimes forces people to say that it is the brain which is the bearer of mental capacities (like the capacity for perception, for example). But when combined with the view that *I* am also a bearer of mental properties, this has the rather unsatisfactory consequence that I am my brain. Although some philosophers are happy to accept this way of talking, they are not forced to do so; even materialists can agree with Noam Chomsky's commonsense observation that 'people think, not

their brains, which do not, though their brains provide the mechanisms of thought'.⁷ Distinguishing between subjects of thought and their underlying mechanisms allows us to say something which does not do much damage to something we normally believe: that persons are the fundamental subjects of mental predications. The tempting conclusion towards which we are moving is one which was defended by P.F. Strawson forty-five years ago: that persons are substances in their own right, among the basic entities in our world, to which we can add that they are items with their own principle of unity and principle of change.

Our question then becomes, 'what is a person?'. In the contemporary debate over this question, the thesis that persons are substances has been associated with the Cartesian conception of the soul as an immaterial substance. This unappealing approach has been opposed by broadly reductionist approaches to personal identity, which attempt to construct a person's persistence over time out of more simple units, such as 'person-stages' or individual mental states and processes. This is the modern descendent of David Hume's view that each of us is 'a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement'. If we are to defend the notion of a person as a substance in its own right, we need to reject this reductionist maneuver too. The thesis that persons are substances must then occupy a middle position between the Cartesian theory and the reductionist theory. Finding space for this position is, however, easier said than done. The problem is that even if we have rejected the reductionist view that a person simply is a collection of 'person-stages', we still need to say something about what balance of psychological and bodily features is needed in order for the same person to persist over time. I cannot hope to address this question satisfactorily

_

⁷ 'Language and Nature' *Mind* 1985: 8.

in the time I have left, but I would like to close with a few remarks about this problem.

Among the approaches that treat persons as (what I am calling) substances, we can distinguish two answers to the question of the relationship between the person and the animal which each person is. One answer is to say that each person simply is an animal (this is the 'animalism' defended by Eric Olson and my UCL colleague Paul Snowdon). On this view, I am an animal and the conditions for my continued existence just are the conditions for the continued existence of the animal which I am. Unlike views which reduce the person to a succession of psychological states, this view explains how it is that I – this person here – was once a foetus, and will continue to exist even if I collapse into a persistent vegetative state. What would this view say about Locke's famous definition of a person as 'a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places'? The simple answer is that if this is what a person is, I can continue to exist even when I am not a person, since the conditions for my own existence are not essentially psychological at all. On this view, I was once a foetus even if no foetus is a person (which by Locke's criterion it clearly isn't). So the consequence is that I am not a person throughout every moment of my existence.

An alternative view is to say that what it is to be a person is to be an animal of a certain kind, but a kind which is characterised partly psychologically. This is the view which has been defended by David Wiggins, who has argued that a person is:

'an animal falling under the extension of a kind whose typical members perceive, feel, remember, imagine, desire, make projects, move themselves at will, speak, ...acquire a character as they age, are happy or miserable, are susceptible to concern for members of their own or like species [etc.]'⁸ [This is

⁸ Sameness and Substance (1980) 171.

only to mention a sample of the characteristics which Wiggins mentions in his description.]

This view can allow that an animal is a person whenever it exists. Since the relevant kind of animal is a kind which is characterized partly psychologically, it is tempting – though perhaps misleading – to describe such an animal as a *mental* substance, so long as that does not exclude the animal's also having bodily characteristics.

In any case, each of these views is a way of defending the view that the person is a thing which persists through change by being numerically identical at different times. Neither view is without its difficulties. But the difficulties with the views should not force us back into the reductionist position which gives us no explanation of the unity which we encounter in experience of ourselves and others. Sometimes the debate about personal identity is introduced in such a way that a reductionist answer to the question looks inevitable. But it is not inevitable; we are not forced into it, and we should instead examine the assumptions which encourage the illusion that it is inevitable.

The picture I have ended up with here, rather unsurprisingly, does not have any need for the idea of the soul as an immaterial (let alone an immortal) entity. And even though I took some inspiration from Aristotle's views about substance, the picture does not need Aristotle's conception of the human soul as the 'form' of the body's matter either. (The interpretation of Aristotle's own understanding of the relationship between the soul and the body is fraught with difficulties that I will not presume to address.) So I although I started off talking about the soul, I have ended up dispensing with it. Jo Wolff told me once that Richard Wollheim would give his first year lectures on the subject of immortality: having raised the question of whether we can survive death, he would use every lecture to digress onto a another philosophical

topic which this question presupposed. By the end of the lecture course, he had never returned to the subject of immortality.

I began this lecture by talking about Descartes's dualism of soul and body. I rejected this dualism because I rejected his conception of substance, as that which is capable of independent existence. Nonetheless, I did not want to dispense with the idea of substance entirely: I claimed that we should instead find a place for an Aristotelian conception of substance as a persisting thing with its own principle of unity and conditions for continued existence. Persons are substances in this sense, beings with a special balance of psychological and bodily characteristics. It is worth noting that something like this idea is suggested by some famous remarks of Descartes, which are rather out of harmony with his dualistic view of soul and body:

'I am not lodged in my body like a pilot in his ship, but, besides ... I am joined to it very closely and indeed so compounded and intermingled with my body, that I form, as it were, a single whole with it.'

If our aim in the philosophy of mind is to understand, in very general terms, our psychological nature, then one of our starting points should be this insight of Descartes's about what he called the 'substantial union' of mind and body. And if this apparent unity is a real unity, then my hypothesis is that we can only capture it by employing the traditional concept of substance, rather than some reductionist surrogate.

May 13 2004