



SUPPLEMENTARY VOLUME XLVIII
1974

THE
ARISTOTELIAN
SOCIETY

THE SYMPOSIA READ AT THE
JOINT SESSION OF THE
ARISTOTELIAN SOCIETY
AND THE MIND ASSOCIATION
AT THE UNIVERSITY OF
LANCASTER

12-14 JULY 1974

PUBLISHED BY

Methuen & Co. Ltd.
in association with The Aristotelian Society

1974

of wishes are connected with the physiological correlates of actions, any more than a normal person needs to know about the mechanism in the nervous system by which his desire for revenge is related to his behaving badly towards a friend in order to see that as his reason and to attempt not to act for that reason in future.

In the case of Elizabeth von R. it is clear that what the patient resisted and later accepted was "the group of ideas relating to her love" which had been "separated from her knowledge" (*S.E.* II, 157). It was the recovery of this group which had a shattering effect on her and which was associated with the variations in her pains to which Freud attached such significance. About the conversion of what "might have become, and should have become, *mental pain*" Freud says "I cannot, I must confess, give any hint of how a conversion of this kind is brought about. It is obviously not carried out in the same way as an intentional and voluntary action" (*S.E.* II, 166). He is here separating the story in terms of reasons from the story in terms of causes. If *he* did not know how the neurophysiological story was to be told he cannot have thought that Elizabeth must know it in order to be cured.

Excitation in the genitals and disgust in the glottis, love for a brother-in-law and pains in the legs can all be felt, just as a desire for revenge and the satisfaction of hurting can be in normal cases, but of course there is no feeling of the connection between them. There is no possibility of being aware by introspection of, or of reporting, the *production* of symptoms by fantasies any more than there is any possibility of being aware of, or reporting, the *production* of hurtful behaviour by jealous feelings. What would be measured by a libidometer is not, and is not claimed by Freud to be, unconscious in his sense; it is not the kind of thing that could be either conscious or unconscious. Neither is the flow of nerve currents which is no doubt correlated with my jealous behaviour. We cannot ask for evidence of implementation in the abnormal case any more than in the normal case; in neither do we need it and in neither does the neurophysiological mechanism enter into the explanation in terms of reasons. Freud does not use self-intimation to protect his views from refutation because in so far as he does use it evidence is always needed to support it and so obtainable to refute it, at least in principle.

ZOMBIES v. MATERIALISTS

Robert Kirk and J. E. R. Squires

I—Robert Kirk

'Man is nothing but a physical object.' This slogan will serve as the starting point for an attempt to reach a clear statement of what is involved in the materialist view of man. Though I think the resulting statement is reasonably clear, it goes beyond what is usually thought to be implied by materialism, and will probably seem too restrictive. Nevertheless, I shall argue that it is logically implied by *any* view of man which remains true to the rationale of materialism. That accounts for the first half of my paper. In the second half (secs. II-IV) I shall expound and discuss a certain objection to materialism. The form the objection takes is suggested by the arguments used in the first half. If I am right in both halves of my paper, all materialist theories of man are false. If I am right in the first half but wrong in the view I defend in the second, a materialist theory of man is true. If (to be more realistic) I am wrong in both, the discussion may still help to clarify some of the issues.

I

If men were as Descartes describes them, they would of course not be mere physical objects. For although a Cartesian man, being spatially extended, could properly be described as a physical object, he would be more than that, since part of him—his soul—would be by definition non-physical. Thus if Descartes' view of the nature of man were correct, an exact physical replica of a given man would not necessarily be an exact replica of that man—unless by natural necessity a soul were automatically assigned to it. This suggests an approach to the explanation of the materialist's dictum that man is nothing but a physical object. For the dictum seems to imply that any exact *physical* replica of a man would be in all respects a replica of that man, not because something was automatically

added to his physical replica, but because there would be nothing to him over and above what was fully describable in physical terms. I shall now try to find a clearer way to express this idea.

First, what is it for one thing to be a 'replica' or a 'physical replica' of another? One possible definition would be in terms of properties. We might say, for example, that one thing was a replica (physical replica) of another if and only if each had exactly the same properties (physical properties) as the other. But a glance at the literature suggests that it might be better to avoid the notion of a property, if that is possible. I therefore propose to make do with the reasonably clear notion of a description, hoping thereby to circumvent those foggy patches which tend to obscure talk of properties and the reduction of properties at precisely the places where clarity is most needed.

An exact replica of a given object x need not stand in the same types of causal relations as x does, nor occupy the same positions in space or time. In general, whatever descriptions logically depend, for their application to x , on x 's particular circumstances will not necessarily apply to everything which is exactly similar to x . (Note that the dependence in question here must be logically necessary. For it seems possible that some or even all of the properties of physical objects are in fact causally dependent on their particular circumstances.) We may distinguish two main classes of description which must be ruled out of consideration. The first is those whose application to the given object x depends logically on the past, present, or future existence of certain particulars other than x . These will be all those descriptions which include names, definite descriptions, or indicator words such as 'I', 'now', 'here'. But it is not sufficient to exclude only this class of descriptions, as is shown by such examples as 'exposed to atomic radiation'. We must exclude all descriptions whose application to x depends logically on the existence of things of certain types, when these are not actually components of x itself. These two conditions appear to be jointly sufficient, however, and I therefore define the class of *non-relational descriptions* of x as all those descriptions of x whose application to x is logically independent of the existence (past, present or future) both of particulars other than x , and of types of things which are not

components of x . It is important to note that this definition does not exclude descriptions of things as subject to various natural laws, and as the subjects of various law-like generalisations. (The application of any description depends logically on the existence of a language, of course. But since this is true of every description whatever, the point may be ignored.)

I can now give a further definition: y as at time t_2 is a replica of x as at time t_1 if and only if every non-relational description which applies to x at t_1 applies to y at t_2 . (It follows that no non-relational description which does not apply to x applies to y .)

In order to define the notion of a physical replica I want to use the notion of a 'purely physical vocabulary'. Instead of defining it, however, I propose to assume that some reasonably clear specification has already been provided. It might perhaps be explained as 'vocabulary sufficient for describing all the features of non-living things'; or as 'vocabulary of contemporary physics'; or the materialist might simply, if tediously, list his preferred vocabulary. I do not think that the difficulty of specifying the favoured vocabulary is serious. True, the materialist's thesis becomes vaguer the less precisely the vocabulary is defined; and the richer the vocabulary is allowed to be, the less interesting the thesis. But imagine what it would be like if materialists had managed to show convincingly that the applicability to human beings of such descriptions as 'has an after-image', 'feels faint' could, without running into philosophical difficulties, be wholly explained on the basis of contemporary physics and chemistry. That would be a marvellous achievement; and neither the fact that physics and chemistry will probably be revised, nor the absence of any very precise general account of what is to be counted as 'physics' or 'chemistry', would diminish its interest or importance. My arguments will not be significantly affected if I simply assume that the 'purely physical vocabulary' has been adequately specified.

With that understood, 'physical replica' can be defined thus: y as at t_2 is a physical replica of x as at t_1 if and only if every non-relational description in the purely physical vocabulary which applies to x at t_1 applies also to y at t_2 .

We are now better equipped for the attempt to clarify the

implications of materialism. First, notice that anyone who maintains that man is nothing but a physical object must agree that:

- (A) Any physical replica of a given man as at a given time would also be a replica of that man as at that time.

It may seem obvious that materialism has this implication, but since the point plays a vital part in my argument, some further remarks are in order. Consider the case of materialism about stones, a view to which I suppose we all subscribe: 'stones are nothing but physical objects'. One could not be said to hold that view unless one agreed that for any given stone *s*, every physical replica of *s* as at a certain time would be in all respects a replica of *s* as at that time. For to admit some actual stone was such that certain non-relational descriptions which applied to it would not also apply to every physical replica of it would be to say that there was more to the stone than there was to things which were exactly similar to it in all physical respects, things which were 'nothing but physical objects' in a very clear sense. In other words, it would be to contradict the slogan. *That* stone, at least, would be more than a mere physical object. I conclude that materialists could not avoid commitment to (A) even if they wished to (not that there is any reason to suppose they do).

However, while it is necessary, if one thinks that man is nothing but a physical object, to agree that any physical replica of a given man would be a replica of that man, it is not also sufficient. A dualist could hold that we inhabit a world in which anything answering to the purely physical descriptions of a given man as at a certain time would be by natural necessity endowed with all those non-physical features which (in his view) the man possessed at that time. But this consideration suggests the possibility that a neat explication of the materialist's slogan, and so of materialism, might result from adding to (A) the condition that it be true by logical or conceptual necessity:

- (B) The mere fact that someone is a man entails that any physical replica of him as at a given time would also be a replica of him as at that time.

This would certainly be a materialist view of man; and the

form of (B) might even serve for an adequate formulation of a materialist view of sticks and stones. However, contemporary materialists are usually willing to admit that there could conceivably have been more to a man than his purely physical features; they only maintain that in fact there never is more. (Perhaps there are *a priori* arguments which show that no other view but materialism is even logically tenable. But it would be rash to define materialism on the assumption that such arguments were forthcoming.) So we must ensure that our account permits the materialist to say, if he wishes, that men only happen to be nothing but physical objects. The following account meets this requirement:

- (C) Every non-relational description which applies to a given man at a given time is entailed by the conjunction of all the purely physical non-relational descriptions which apply to him at that time.

(C) avoids the objection just mentioned because it leaves open the logical possibility that someone should be a man and yet fail to be the subject of purely physical descriptions which actually entailed whatever other non-relational descriptions applied to him. The proponent of (C) would insist that, in the world as it is, the purely physical descriptions of any actual human being—including, of course, physiological descriptions, or their physical-language equivalents, and physical laws and lawlike generalisations—do entail whatever other non-relational descriptions also apply; but he would allow it to be a matter of contingent fact that human beings conform to physical descriptions of these kinds. Thus while (C), like (B), is not merely consistent with (A) (the requirement that any physical replica of a given man would also be his replica), but also offers an explanation of (A) by entailing it, (C) avoids the extreme of implying that materialism is necessarily true.

Now (C), or (as it may be helpful to call it) the Entailment Thesis, is itself entailed by analytical behaviourism, when that is taken to be the view that all psychological concepts are analysable in terms of mere bodily movements and dispositions to move; and it is also entailed by the conjunction of a causal analysis of psychological concepts with the highly plausible assumption (which I shall take for granted throughout

this paper) that all human bodily movements are in principle explicable in purely physical terms. But it is unnecessary for the materialist to construe 'entails' so tightly that the Entailment Thesis could be true *only* if it were possible to find satisfactory behavioural or 'topic-neutral' analyses of problematic concepts. There are two reasons for this.

One reason is that the difficulty, even the impossibility, of formulating such analyses might be a reflection not of some deep unbridgeable difference of subject-matter between the psychological and physical vocabularies, but simply of their having been arrived at in different ways, to serve different purposes. A non-psychological illustration may help. Suppose the 'purely physical vocabulary' is restricted to the vocabulary of physics and chemistry. The materialist would say that trees, for example, are 'nothing but physical objects', despite the fact that a word like 'gnarled', which applies to some trees, is excluded from his preferred vocabulary. To provide an analysis of 'gnarled' in physical or neutral terms would indeed be one way of explaining why its application did not imply that trees were anything but mere physical objects. An equally satisfactory explanation, however, even if such an analysis is impossible, would be that the applicability of the word is determined solely by what could in principle be described in the purely physical vocabulary. Conceptual considerations are clearly sufficient to account for the fact that any physical replica of a given gnarled tree would also be gnarled—even if no analysis is available. The other reason why the materialist need not take the word 'entail', as it occurs in the Entailment Thesis, to imply that all problematic expressions can be analysed in physical or neutral terms, is that epistemological or other conceptual considerations might do the necessary work indirectly. Some valid form of verificationism, for example, might rule out the possibility that something should satisfy the physical descriptions which applied to a given man without also satisfying whatever other non-relational descriptions also applied to him.

It is clear, then, that materialists are not necessarily committed to the Entailment Thesis on any narrow interpretation of entailment. Indeed, despite the fact that some (David Lewis, for example¹) have adopted positions which entail the Entailment

Thesis, one might well think that materialists as such are not necessarily committed to the Entailment Thesis in any form, even on the widest possible interpretation of entailment. However, our earlier considerations make it possible to see that this would be a mistake. Let us take 'entail' as it occurs in the Entailment Thesis in the widest possible sense, so that *p* entails *q* if and only if there are *some* logical or conceptual considerations as a result of which it is impossible that *p* should be true and *q* false, or (for the case of descriptions) impossible that *p* should apply to something to which *q* does not also apply. Then (surprising as this may appear) very little additional argument is required to show that on this interpretation the Entailment Thesis is not only sufficient, but necessary, for any materialism worthy of the name. (I am using such phrases as 'entail' and 'logical or conceptual' as blunt instruments, by the way. For those with Quinean scruples they may, I think, be acceptably paraphrased without impairing the argument. '*p* entails *q*' might, for example, be rendered 'inclusion of both "*p*" and "*not-q*" in our sentence-systems would be more than difficult: it would make the systems unworkable'.)

The additional argument required is very straightforward. The Entailment Thesis would be false if, and only if, there were a man whose non-relational physical descriptions at a certain time did not entail the totality of non-relational descriptions which applied to him at that time. But such a man could not sensibly be described as 'nothing but a physical object'. For reasons we have already glanced at, he would be *more* than that. What would undoubtedly be nothing but a physical object, and that in a transparently clear sense, is a physical replica of the man to which there applied *only* the physical descriptions and whatever they entailed—something we may conveniently dub a 'Zombie replica'. Since the man would differ from his Zombie replica, but not in ways describable in purely physical terms, he would not be 'nothing but a physical object', or not in that same transparently clear sense. So if the materialist's slogan is to be understood in that sense, it is logically equivalent to the Entailment Thesis. (I shall assume for the sake of simplicity that if Zombie replicas are logically possible, the descriptions which fail to apply to them are descriptions of sensations. Obviously such descriptions

are likely candidates, if any are; and my argument is unaffected by the exact nature of the class of descriptions in question.)

It does not matter that there are in fact no Zombie replicas of men. The point is that if such things are merely possible, it is false that *we* are mere physical objects.

The materialist might perhaps object that the phrase 'nothing but a physical object' may be understood in a different sense from the one I have given it, a sense in which materialism could be true even if the Entailment Thesis were false. But if there is such a sense—and the onus of explaining it would be on the materialist—I do not think it could help him to escape the main point of the argument. For suppose the Entailment Thesis is false, so that it is logically possible that a man should stand side by side with a Zombie replica of himself. Then to say that the man and the Zombie would *both* be 'nothing but physical objects' would leave the differences between them—differences signalled by the fact that some non-relational description which applied to the man failed to apply to his Zombie replica—wholly inexplicable in physical or neutral terms. *Ex hypothesi* these differences could not be captured in the net of the most sophisticated physical investigation: physics and chemistry would be blind to them. A dualist could happily swallow all this; but it cannot be passed off as the pure milk of materialism.

For similar reasons the materialist cannot avoid commitment to the Entailment Thesis by invoking the Identity Thesis. Suppose that (the 'havings' of) sensations, after-images, and so forth are held to be identical with certain states or processes describable in the purely physical vocabulary. (Here it is not important to know just what is supposed to be identical with what.) The materialist will either agree that these identifications are actually entailed by the purely physical descriptions, in which case he accepts the Entailment Thesis; or he will deny it. But to deny it is to admit that no logical or conceptual considerations whatever preclude the possibility that there should have been a race of beings physically indistinguishable from ourselves—our Zombie replicas—to which there failed to apply some of the non-relational descriptions which apply to us. Such a race of Zombies would indeed have been 'nothing but physical objects'.² But to say there could have been such a

race is to say that *we* are not mere physical objects. We should differ from them in having after-images, sensations, or whatever—yet physics and chemistry could provide no means for explaining or even recognising such differences. It is not as if statements of psycho-physical correlations or laws—Feigl's famous 'nomological danglers'—were either part of or (on the present assumption) entailed by physical or chemical theory. On this assumption, so far as physics and chemistry were concerned, humans might just as well have been Zombies. Therefore, if a materialist rejects the Entailment Thesis, physics and chemistry cannot do what he requires of them. They cannot provide an adequate description of what actually exists. I conclude that any materialist who wishes to defend a reasonably clear position must endorse the Entailment Thesis.

It is worth adding that if materialists generally had accepted the Entailment Thesis, they would not have needed to appeal to Occam's razor in order to justify the assertion of psycho-physical identities. Indeed, if the thesis is true there ceases to be any need for materialists to decide whether, and in what sense, the various alleged identities hold. For this reason, and also because it requires no commitment to contentious claims about analysability or translatability, the Entailment Thesis enables many current objections to materialism to be side-stepped. However, attractive though I personally find this position, reflection on the idea of Zombie replicas has led me to the conclusion that it is not in fact true. In the remainder of this paper I shall try to explain why.

II

Consider Gulliver in Lilliput. The Lilliputians were fully justified in treating Gulliver as a sentient being (i.e., for our purposes, as having sensations). After all, he was very like themselves, except for his relatively vast size. But would Lilliputian philosophers have been right if they had said that not only did Gulliver's behaviour and dispositions to behave justify the inference that he was sentient, they actually entailed that he was sentient? No. For the truth—although this has not hitherto been recorded—is that although Gulliver had started life as a human being, he was one no longer, having

met in his travels with the following misfortune. He had encountered a race of beings even tinier than the Lilliputians, and technologically much more advanced than ourselves. A team of their scientists (the 'Brain Team') had invaded his head, disconnected the nerves to and from his brain, then rendered it inert, and arranged to monitor the inputs via the afferent nerves from the remainder of his nervous system, and to send outputs via the efferent nerves so as to produce behaviour indistinguishable from that of a normal man. The Brain Team had thus taken over those functions of Gulliver's brain which governed his behaviour, and their zealous commitment to their project ensured that the entity of which they formed a part was *disposed* to behave like a man. I will refer to this entity as 'Zulliver'.

Now I doubt if anyone would want to say that Zulliver had sensations. But even if someone did maintain this view, it is clear that Zulliver's having sensations is not *entailed* by the fact that he (or it, but I shall use 'he') behaved and was disposed to behave in all respects like a man. For Zulliver was nothing but a super-puppet, with resident puppeteers. This blocks the entailment because, if we know that a certain body is being manipulated as a puppet, we know that what appear to be expressions of the feelings of the being whose body it is are nothing of the sort. (Usually they are not even expressions of the puppeteers' feelings.) Yet it is only because the body seems to express feelings that it is up for consideration as the body of a sentient being at all. This argument applies not only to pantomime horses and Zulliver, but also (for example) to the sophisticated electrical manipulation of human beings, whereby it might be possible for the experimenter to cause someone's body—under anaesthetics, perhaps—to go through an extended sequence of evolutions which, had they occurred naturally, would rightly have been regarded as expressions of various coherent thoughts and feelings.

Zulliver seems to me to be a decisive counter-example to analytical behaviourism, including that less vulnerable variety which does not insist on the possibility of piecemeal behavioural analyses of mental concepts, but only on the claim that the psychological vocabulary provides no more than ways of talking about behaviour and behavioural dispositions. But it

will be more useful for our purposes to consider whether Zulliver is also a counter-example to causal analyses such as those advocated by Armstrong and Lewis.

According to Armstrong's version of the causal analysis a mental state is, roughly, a state 'apt for bringing about a certain sort of behaviour'.³ (For certain concepts the cause of the state must also be taken into account.) Since this analysis places no restrictions on the mode of causation, Zulliver does seem to be a counter-example to it. (For a similar reason, Zulliver is a counter-example to the view that all mental states are adequately analysable as what Hilary Putnam calls 'functional states'.⁴) Lewis's version, however, may appear less straightforwardly vulnerable to the counter-example:

... the definitive characteristic of any experience as such is its causal role. The definitive causal role of an experience is expressible by a finite set of conditions that specify its typical causes and its typical effects under various circumstances. By analytic necessity these conditions are true of the experience and jointly distinctive of it. ("An Argument for the Identity Theory", pp. 19-20.⁵)

Now in Zulliver's case there *is* a state typically caused by (say) damage to his bodily tissues, and typically resulting in certain sorts of behaviour (wincing, groans, etc.). (It is a state of the Brain Team more particularly, since the Brain Team responds to the neural inputs by arranging for just those behavioural responses to be produced.) And since it does not follow from this that Zulliver feels pain, Zulliver appears to be a counter-example to Lewis's scheme of analysis. The causal functions supposed to be analytically correlated with pain are performed, yet the fact that they are performed does not entail that Zulliver feels pain. However, we must attend to the way in which the word 'typically' occurs in Lewis's scheme. The point of saying that *x* is an experience of type *T* if and only if *x* is *typically* caused by *C* and *typically* produces *E* is to avoid such objections as that of the paralysed man: a view on which total paralysis necessarily involved the total absence of experiences would be highly counter-intuitive. By not requiring that experiences of type *T* always produce *E*,

such cases can be accommodated. But this device might also be thought to rule out Zulliver as a counter-example. It is true that, if 'typical' is understood only in relation to Zulliver, a certain state of the Brain Team is typically caused by *C* and typically produces *E* (for every relevant pair *C* and *E*). But relative to the human race, it might be urged, that state, *under the description* 'state of the Brain Team', is *not* typically caused by *C* and does not typically produce *E*. When 'typical' is understood in relation to the human race, that state does not perform this causal role *typically*, because it does not occur in human beings at all. If Lewis's scheme of analysis is taken in this way, therefore, it seems that Zulliver need not be regarded as a counter-example.

But to interpret Lewis's causal scheme of analysis on these lines would be to destroy its claim to acceptability. We may indeed concede that every kind of sensation, under its own psychological description, is typically caused in certain ways and typically produces certain effects. We may even concede, for the sake of argument, that to each kind of sensation in human beings there corresponds a certain distinct type of physical state such that, under some physical-language description (e.g., a physiological one) that physical state typically fills the causal rôle specified by the Lewisian analysis of the corresponding sensation. But there can be no logical or conceptual guarantee that this is so. It is quite conceivable, for example, though perhaps false, that as a matter of fact there is *no* physically describable state which typically occurs in human beings just when they are in pain. Hence, if a scheme of analysis for sensation concepts entails that there is such a state, that scheme of analysis is invalid.

If Lewis's scheme is interpreted in the way just considered, then, it is invalid for the reason given; while if it is interpreted in the first, more natural, way, Zulliver is a counter-example. Therefore the causal analysis will not help to prove the Entailment Thesis. (This is not to deny that Lewis's scheme may serve to explain how psycho-physical identifications could come to be established. The point is that, if I am right, he has failed to show that a *materialistic* identity theory could be adequate.)

The above point can be turned against the Entailment Thesis

in general. For if Zulliver's internal constitution and functioning do not entail that he is sentient, there is no good reason to expect that modifying the contents of his head could fill the logical gap. We do indeed have good reason to believe that if Zulliver were to be restored internally to the state Gulliver was in originally, he would thereby be made sentient. But the question is whether his being sentient would be entailed by the fact that such modifications had been made. And now we must remember our ignorance: we do not know that the causal and other functions of the brain are in each one of us performed by exactly similar physical structures in exactly similar ways. On what rational principle, therefore, could we draw a line between those cases—including, presumably, all normal human beings—for which the entailment held, and those—including Zulliver—for which it did not? I suggest that the example of Zulliver shows that there is no such rational principle. If the causal analysis had been valid, we could have said that if a given sort of brain performed the relevant causal rôles, then it was the brain of a sentient being. But since, if I am right, the causal analysis is invalid, it would be futile to include some specification of physical structure or composition in the hope of closing the logical gap. After all, we knew that normal people were sentient long before we knew anything of the causal functions of their brains. Neither the discovery of those causal functions, nor the discovery of certain details of the fine structure and composition of the brain have added anything to the strength of our conviction that others are sentient. Had they done so, the fact might have tended to support the Entailment Thesis. But (for obvious reasons) they did not, and I conclude that there is no *direct* entailment from the purely physical descriptions of a man to all the other non-relational descriptions which apply to him.

III

Indirect support for the Entailment Thesis may be looked for from some form of verificationism. One suggestion might be that if there is no publicly observable difference between two individuals at a given time, there is no sense in speaking of any difference at all between them at that time, such as there

would be between a man and his Zombie replica. But this principle would be too strong. We cannot, for instance, know *a priori* that two men, physically indistinguishable at a certain time right down to the minutest details of neural structure, could not possibly be having different sensations at that time, as revealed in the usual ways. Of course the materialist will insist that this would never actually happen. (The Entailment Thesis commits him to that.) But it would be self-defeatingly unempirical of him to rule out the possibility that nature should fail to come up to his expectations.

In order to take account of such difficulties we might permit events before and after the given time to be taken into account. But it would then be possible, I think, to describe situations that would make it reasonable to say—not at the time, but later—that one had been confronted by a counter-example to the Entailment Thesis (a Zombie). For example, a man's patterns of behaviour might go through a sequence of modifications which would be utterly baffling except on the hypothesis (still baffling, but with a good deal of explanatory power in the circumstances) that he had been by slow degrees deprived of all varieties of sensations and other experiences until he reached a state of Zombiehood, which was maintained for some time, and had then, again by stages, regained his senses and returned to normality. A story to this effect (such as I have developed elsewhere⁶) might perhaps satisfy those who require only the possibility of *evidence* for and against the truth of a proposition if it is to be admitted as the description of a logically possible state of affairs. However, some verificationists would not be satisfied with this.

According to a modified version of 'anti-realism' discussed by Bede Rundle, for example, it must be possible to reach a point at which there is no longer any reliance on evidence—although, since the demand for verifiability as such is admitted to be mistaken, "it is required only that *someone* should be able to be in that position, not everyone".⁷ Let us consider how this requirement would apply to statements of the form 'x is now in a dreamless sleep'. Nobody (*pace* Descartes) denies that such statements are sometimes true, and therefore perfectly intelligible. So who, in connexion with such statements, can reach a point at which there is "no longer any reliance on evidence"?

Not observers other than the sleeper: they depend wholly on evidence. But not the sleeper himself either, until after the time in question. Only later could he say that, though asleep, he had not been dreaming. But if this is how we must understand the anti-realist requirement, then, granted the possibility of the sort of story described in the preceding paragraph, the requirement will not save the Entailment Thesis. For the story can easily mention the victim's subsequent descriptions of his experiences and the gaps in his experiences. (One does not rely on evidence in order to be able to describe gaps in one's experience.) Of course his account would not do all the work by itself, but in that respect it would be no worse off than some accounts of cases of sleep-walking and sleep-talking. I conclude that this modified anti-realism is quite consistent with denying the Entailment Thesis.

Verificationistic support for the Entailment Thesis is also vulnerable to the Zulliver example. When the Lilliputians, or for that matter Zulliver's normal-sized contemporaries, described and treated him as having thoughts and feelings, they did so because he was to them indistinguishable from a normal man. They did not have to decide whether or not to treat him as sentient: the question simply did not arise. And of course they were fully justified in treating him so. Nevertheless, for the reason given earlier, the fact (assuming it is a fact) that he resembled a normal man in all behavioural and other externally observable respects does not entail that he was sentient. Indeed, for the same reason—that he was, in effect a puppet—I think most people who knew the facts would be inclined to deny that he was sentient. The same reason, again, warrants the statement that it is perfectly intelligible to describe Zulliver as insentient. Yet although his physical description does not entail that he was sentient, neither does it entail that he was insentient. We all know *we* are sentient; and no doubt we also know there are no little men inside *our* heads. Yet the first item of knowledge does not entitle us to exclude *a priori* the possibility that we might have been mistaken about the second. My knowledge that my head aches is not, so far as I can see, logically or conceptually incompatible with the discovery that the inside of my head is like the inside of Zulliver's. If this is correct, the intelligibility of describing

Zulliver as insentient cannot be explained as a mere logical consequence of publicly verifiable facts about him. It seems that any verificationistic principle strong enough to establish the Entailment Thesis would be too strong to be valid.

Indirect support for the Entailment Thesis might perhaps be sought from arguments against scepticism about other minds. But opposition to the Entailment Thesis does not commit one to scepticism, so far as I can see. To assert that something is logically possible is not to deny that we know it is not the case. I have consistently argued on the assumption that we all know we have thoughts and feelings; indeed I used that assumption in the preceding paragraph. Nor do I treat sensations as a kind of private object (or not in a way that leaves me open to the usual objections). For if, as I should maintain, it is logically possible that a man having an after-image should stand side by side with a Zombie replica of himself which was not having an after-image, it does not follow that 'after-image' is the name of something which the man has and the Zombie lacks. It does not follow, in other words, that the difference between them is *the after-image*—though certainly the Zombie would be lacking in *some* respect that was a necessary condition of having an after-image. Thus my position seems to be compatible with, although it does not bind me to, acceptance of Wittgensteinian arguments against the idea of private identifications of mental states, processes and events. I therefore see no help for the Entailment Thesis from this quarter.

IV

My arguments from the example of Zulliver are of course stated in ordinary language. It might be suggested that they seem to present a difficulty for materialism only if we assume, no doubt wrongly, that our existing conceptual scheme is immutable. But I think this suggestion would be mistaken. What my arguments show, if they are sound, is that there is something about human beings which materialism denies itself the resources to recognise, no matter what linguistic conventions we might come to adopt or abandon.

Materialists contend that the purely physical vocabulary

is adequate to describe everything that exists. They therefore need to explain how it is that certain further sorts of description apply, or seem to apply, to human beings. To show that the Entailment Thesis was true would be a way of providing the required explanation. But if my arguments from the Zulliver example are sound, the Entailment Thesis is false; and for the reasons given in the first part of this paper I do not see how else materialism could be vindicated. However, it seems clear that this difficulty is intimately bound up with the question of 'private access'. Descriptions in the purely physical vocabulary are not generally such that, for their correct application on any given occasion, one individual is in an epistemologically unique position: a position where he, but no one else, is able to apply the description correctly without needing access to what is publicly observable. But the psychological descriptions which bother materialists are all of this kind. (Private access must not of course be confused with 'incorrigibility' or 'indubitability'. One can have private access to matters about which one is very much open to correction, and which are describable in purely physical terms; for example, the presence of a speck of dust in the eye.) And the materialist might argue that in view of these facts the failure of the Entailment Thesis is no embarrassment to him. He might urge that the thesis fails, not because physics leaves something out, not because the physical vocabulary is inadequate to describe what ultimately exists, but simply because psychological descriptions depend heavily on the exploitation of private access, while purely physical descriptions do not. On this view the failure of the Entailment Thesis would reveal (so to speak) a mere difference in approach by the two vocabularies, not a difference in subject-matter.

But this suggestion will not work. If my arguments are sound, they show that if materialism were true there would *be* no private access. For if the Entailment Thesis is false, as I have argued, what I call Zombies are logically possible. If the psychological vocabulary were just a different way of dealing with the same subject-matter as is dealt with by the physical vocabulary, however, Zombies would not be logically possible: by logical or conceptual necessity, whatever sorts of psychological descriptions apply to us would apply to Zombies. Thus my arguments tend to show that private access is a fact

about human beings which is logically independent of the physical facts, and logically independent of whether or not we happen to employ a particular vocabulary. It is a matter of convention that we describe pains, for example, as we do, and distinguish them as we do from other sensations. But if my arguments are sound, it is not a matter of convention that we *feel* pains. Even if men came to abandon all means of talking about sensations, and every other variety of psychological language, confining themselves to what could be said by means of the purely physical vocabulary, they would simply be ignoring their differences from Zombies, not annihilating it. By a consideration of how unmysterious Zombies would be, I think we are forced to recognise that there is more to us than materialists are able to admit.

REFERENCES

- ¹ D. K. Lewis, 'An Argument for the Identity Theory', *Journal of Philosophy*, 63 (1966), 17-25; 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy*, 50 (1972), 249-258.
- ² G. F. Stout, in *Mind and Matter* (Cambridge, 1931), pp. 138-139, describes this possibility as an objection to materialism, though in connexion with the causation of events. The idea of Zombies is of course an old one. See, e.g. T. H. Huxley, 'On the Hypothesis that Animals are Automata, and its History', in his *Methods and Results* (London, 1894).
- ³ D. M. Armstrong, *A Materialist Theory of the Mind* (London and New York, 1968), p. 82, and *passim*.
- ⁴ See, e.g. his 'Minds and Machines', in *Dimensions of Mind*, ed. S. Hook (New York, 1960); or 'Psychological Predicates', in *Art, Mind and Religion*, eds. W. H. Capitan and D. D. Merrill (Pittsburgh, 1967).
- ⁵ Lewis has illuminatingly developed this position in the second article referred to; but the refinements do not affect my main points.
- ⁶ R. Kirk, 'Sentience and Behaviour', *Mind*, 83 (1974).
- ⁷ Bede Rundle, *Perception, Sensation and Verification* (Oxford, 1972), p. 245.

ZOMBIES v. MATERIALISTS

Robert Kirk and Roger Squires

II—Roger Squires

Philosophical materialism still seems far from clear, whether about people or even stones. Mr Kirk supposes that we all subscribe to the view that "stones are nothing but physical objects". If some child or crank thinks that stones are the seeds of trees or the eggs of birds it would be apt to point out that stones are objects of one kind—inorganic, inert—rather than another—organic, ert. Those who subscribe to or oppose materialism often debate as if one side think stones, people etc. are of this kind—physical, material—and the other side think they are also or instead of this other kind—non-physical, spiritual. But if this question is analogous to that whether stones are organic or inert, ought it not to be debated by scientists rather than philosophers?

Can philosophers reasonably debate the constituents of a person? "Flesh, blood, bone and sinew, my son". "But what about the puppy dogs' tails?" "People are nothing but flesh, blood . . ." And father may or may not be right, depending on his knowledge of anatomy and the state of that subject. It can't be like this when philosophers debate whether people are material or spiritual or a mixture of the two.

"Those trees you admired", a friend once said to me, "are simply sticks groping for chemicals". (In just such a tone, "Men are nothing but physical objects" echoes round dusty seminar rooms.) What made possible the clipping of the angel's wings was the plausibility of supposing that trees must be much more than that in order to account for their evident properties. Then better science provides debunking explanations. We are all hardened to such dispersion of magic. "A Will-o'-the-Wisp is just moving marsh gas"; "A Brocken spectre is only the observer's shadow on the cloud". Someone who knows demonstrates that a phenomenon that seemed to need at least this to explain it requires only that or merely the other. It is easy to understand such claims and counterclaims about people. How can what is

over ninety per cent water play football or cross deserts? What is hard to understand is why philosophers think they can take more than an intelligent layman's interest in such questions. They have little enough to say about marsh gas or Brocken spectres or even trees—and with reason.

Mr Kirk begins: "If men were as Descartes describes them, they would of course not be mere physical objects . . . since part of him—his soul—would be by definition non-physical". If this thesis or its materialist denial is to be clear, we will need to know what physical kinds of parts are and what non-physical kinds of parts are. Mr Kirk assumes we can specify "a purely physical vocabulary", for this "might perhaps be explained as 'vocabulary sufficient for describing all the features of non-living things'; or as 'vocabulary of contemporary physics'; or the materialist might simply, if tediously, list his preferred vocabulary . . . the richer the vocabulary is allowed to be, the less interesting the thesis". As if Descartes had been claiming that people have a constituent not now generally admitted by scientists, some chemical untabulated by chemists or particle unrecognized by physicists. How could that itself avoid being a scientific claim?

Similarly, Mr Kirk sometimes describes materialism as if it is a debunking rejoinder to the claim that particular phenomena must be explicable in terms of special ingredients. "Imagine what it would be like if materialists had managed to show convincingly that the applicability to human beings of such descriptions as 'has an after-image', 'feels faint', could, without running into philosophical difficulties, be wholly explained on the basis of contemporary physics and chemistry". What kind of explanation has he in mind? Is it like the explanation of Jack-o'-Lantern in terms of rising marsh gas? What other kind of explanation would proceed "on the basis of contemporary physics and chemistry"? If materialism is to be a thesis on which non-scientists can reasonably take sides, this needs an answer.

Suppose the town council threatens to remove a large stone to make a new roundabout and it is objected that the stone is an ancient landmark. If a councillor replied, "stones are nothing but physical objects", we should not all subscribe to that view. For he would be denying that a stone can have a kind of

importance which it may well have. Whether such 'materialism' is plausible depends on what is being denied.

"You can never pick a dandelion," read the magazine in the dentist's waiting room, "For if you try in summer you get the dazzling flower, but not the fragile seed clock which you would get if you tried in Autumn. There is always much more to a dandelion than what you can pick". And some say that you cannot hold a person in your arms or seal them in a box, can never point to a particular skittle of flesh and claim successfully, "*This* is a human being; weigh, measure, analyse it—here it all is". There is always much more to a person than the mere physical object, bag of bones, which you can point to.

There is little point in the correct denial, "of course, you can pick a dandelion, put a person in a cage, etc". This would miss the truth in the paradox, which is, I suppose, that you cannot properly understand what a dandelion is, what "dandelion" means, by being shown what someone who has picked one holds in his hand. Roughly, this would do for stones or leaves or plastic daffodils, but certainly not for wild plants where a knowledge of their normal life circle is essential. Similarly, to find out what a person is made of, what descriptions are true of a particular specimen at a particular time, is only to have a fractional understanding of what people are and, hence, of what "person" means. The paradox contains no news about plants or people or about the words "plant" or "person".

The purveyors of such sentiments do not intend to take issue with what scientists discover about dandelions or people. Sometimes they unwisely make this point by saying they are concerned with non-physical or immaterial things. This misleadingly suggests the super ingredient, the perennial wooffle dust or immortal spark that survives earthly vicissitudes. The 'spiritualist' reminder about what is involved in being a person now appears after all as a controversial challenge to scientific accounts of human beings. He should have said that there is more to knowing a person than knowing his parts and avoided all talk of non-physical parts as a confusion. Someone who reminds us that there is more to a human being than interests a cannibal is not in dispute with someone who tells his daughter that far from all things nice she consists of so much calcium, so much carbon and so on.

The materialist counterclaim that people are nothing but physical objects is infected by the ambivalence of 'spiritualism'. It has the tone of debunking science when it takes the stand that human beings can be "explained" without postulating extra ingredients unknown to physicists. Yet its air of moral shock comes from the feeling that it denies the truisms on which the spiritualist rightly insists. But it is silly to deny the truisms and, if anyone had challenged scientific accounts of people, it would be unprofitable for non-scientists to debate it. As a philosophical thesis, materialism makes the mistake of denying before it finds out what, if anything, has been said.

Philosophers are neither specially knowledgeable nor specially ignorant of either human beings or the words "human being". They would be wise to avoid anatomical or philological debates. "People are physical objects" may remind us that people are not equations, rôles or institutions. Similarly, we would no doubt all subscribe to a materialism about stones which reminded us that stones are not square roots, distances or pay differentials. At the same time we would equally subscribe to the 'spiritualist' reminder that people are not only physical objects—that they differ from other non-abstractions in various ways, that they have a complexity and importance over and above comparable lumps of matter.

Disputes break out when the spiritualist tries to tangle with the scientists or when the materialist overrides our ordinary understanding of what a person is in the name of science. "For seeing life is but a motion of limbs, the beginning whereof is in some principal part within; why may we not say, that all *Automata* (engines that move themselves by springs and wheels as doth a watch) have an artificial life? For what is the *Heart*, but a *Spring*; and the *Nerves*, but so many *Strings*; and the *Joints*, but so many *Wheels*, giving motion to the whole body . . ." (Hobbes—introduction to *Leviathan*.) Now here is someone blinded by mechanics who had forgotten the differences between clocks and people, between movement and action. It is no criticism of that mechanics to point out that in explaining movements it has not explained actions. Nor, obviously, must we suppose that there are non-mechanical springs and wheels that do explain life, or that life cannot be scientifically explained.

Traditional theorists have claimed that a person has spiritual as well as physical parts. That a person is constituted of a mind or soul and a body. Modern theorists have denied the existence of any spiritual parts, asserting that insofar as the mind is something respectable it is only physical anyway. The problem is to understand what "physical" and "spiritual" mean here. If we catalogue "physical" adjectives, as Mr Kirk suggests, this makes the dispute seem like that between two physicists debating what is the minimum vocabulary for describing the results of their investigations. If we take "physical" to mark the contrast between stones and trees and abstractions, such as the equator and numbers, it is hard to see what sense there can be in either asserting or denying that human beings have abstract parts.

Certainly there can be non-physical parts. Games, battles, arguments, performances, may be made up of various elements. Honesty may be a part of Simon's character and simplicity may be an element in his mental or intellectual make up. Non-physical things have non-physical parts. But the traditional dualist was that minds are mental constituents of persons, that Simon himself has non-physical parts. It is rather as if someone suggested that not only does the knight in a chess set have physical parts—the head unscrews from the base—but it also has non-physical parts, for part of its power is to capture any piece occupying the square on which it lands. Even though the knight is defined by its power and its power has parts, these are not parts of the knight. Similarly, honesty may be part of Simon's character or, if you like, his mind, but neither honesty nor his mind are parts of Simon. A person's states, characteristics and activities are not mental or physical constituents.

Mr Kirk thinks that the materialist is committed to the thesis that "every non-relational description which applies to a given man at a given time is entailed by the conjunction of all the purely physical non-relational descriptions which apply to him at that time". This claim that all true descriptions of people are physical descriptions seems to me to hang in the air for want of an explanation of "physical".

Now for the gulling of the Lilliputians, who mistakenly supposed they were dealing with a human being when they were faced with "nothing but a super-puppet, with resident

puppeteers". For "The Brain Team had . . . taken over those functions of Gulliver's brain which governed his behaviour, and their zealous commitment to their project ensured that the entity of which they formed a part was *disposed* to behave like a man" Mr Kirk draws the conclusion that the entity does not have sensations and that no account of the entity's behaviour and dispositions to behave would entail that it did. Hence, Gulliver is a clear counter example to the philosophical thesis of 'analytical behaviourism'. After all, his behaviour is 'indistinguishable from that of a normal man.'

But a puppet does not behave like a man. It does not kick, struggle, scream and talk. It may look like a man, move like a man who is kicking and struggling, even emit noises like someone screaming and talking. Gulliver may be disposed to move, look and sound like a man, but it is not disposed to behave like one. If a 'behaviourist' holds that what behaves and is disposed to behave like a person is ipso facto sentient, he cannot be refuted by producing something which though insentient does not satisfy the conditions he lays down. When Wittgenstein wrote in the *Investigations* § 283, "only of what behaves like a human being can one say that it *has* pains", he did not mean that we can ascribe pain to puppets, dummies or robots. For these things only *seem* to behave like human beings.

Admittedly, it may be difficult, if not impossible, for the Lilliputians to find out that Zulliver is only an entity that moves in certain ways by virtue of its inhabitants. But the fact that they cannot distinguish between movements with sound track and witty repartee while helping with the dishes does not mean that they are indistinguishable, that there is no difference. Gulliver kicks, struggles and screams. The Brain Team manipulate Zulliver to move and sound as if kicking, struggling and screaming. It is true of Zulliver, but not of Gulliver, that if human beings in general did not go through these motions it would not go through them. Mimicry does not become identical with what it mimics, merely because the hypothetical truths which distinguish it are difficult or even impossible to establish. When we know the facts, it is clear that Zulliver does not behave like a human being. Only the Brain Team do that.

In his 'Sentience and Behaviour', *Mind*, 1973, Mr Kirk tells of someone who winces, exclaims, answers, is astonished (or at

least behaves as if he is) and then comments: "I am using 'behaviour' here and throughout in the sense of mere bodily movement, without commitment as to whether or not it consists of actions".

Zulliver may be intended as a counter example to the view that certain descriptions of actual and hypothetical bodily movements are equivalent to descriptions of actions, which in turn entail the ascription of sentience. When Ludwig's arm rises, an arm moves. When Ludwig raises his arm, a person moves. Descriptions of what arms (or muscle or glands) do cannot entail descriptions of actions, if only because of the reference to persons in the latter. The 'behaviourist' claim presumably ought to be more like this: when a person (who is awake and informed of his situation) moves and is disposed to move in certain ways it follows that he has acted in a certain way. Zulliver, not being a person but "a super puppet with resident puppeteers", cannot be a counter example to this.

Zulliver may be intended to be a counter example to the thesis that certain descriptions of the movements and sounds that an entity makes entail that that entity is a person. Our Lambert Simnelacrum moves as a person, makes the right noises, whilst remaining a fairly obvious pretender. If we take a limited range of movements and so forth, other interpretations—the robot, the man in the chess playing machine—are clearly possible. But the "analytical behaviourist" is allowed to appeal to hypothetical movements. The Brain Team try to block this appeal by arranging that Zulliver makes sounds etc. just as Gulliver would have done. It is also true that Zulliver is disposed to move etc. in the same way so long as the Team are in occupation. At present I'm inclined to agree with Mr Kirk that he has shown a weakness in this version of "analytical behaviourism".

Mr Kirk argues that the contents of a person's head must be logically irrelevant to the question whether he is sentient, whether he is a person at all. "After all, we knew that normal people were sentient long before we knew anything of the causal functions of their brains. Neither the discovery of those causal functions, nor the discovery of certain details of the fine structure and composition of the brain have added anything to the strength of our conviction that others are sentient . . . I conclude that there is no *direct* entailment from the purely physical

descriptions of a man to all the other non-relational descriptions which apply to him". Yet a case has been described, that of invasion by the Brain Team, where discoveries about the contents of a head do lead us to deny sentience. Why could someone not hold that Zulliver shows that the internal constitution of a person is logically relevant?

The argument from ignorance is not compelling. We learned about daisies in ignorance of roots. But, as the groundsman knows to his cost, if someone removes the tops to make daisy chains it is not at all clear that the daisies have been removed. There would be something radically wrong if a nursery tried to sell rootless daisies—probably an infringement of the Trades Descriptions Act. Compare how we show children what daffodils are with the later importance we attach to the bulb. What we discover about something may be relevant to what it is.

Again, to know what a pen is, what "pen" means, we don't have to be able to say whether or not any given object is a pen, even after full investigation. There will be a list of disqualifications which we have no reason to formulate or agree about. For example, a candidate 'pen' may be dry—suppose the marks are made by some cunning pressure mechanism. Do we say, "we knew nothing about such processes when we learned about pens and writing, so their existence is logically irrelevant to the question, though it may be *evidence*"? In the light of the new facts we can decide. We need not step into our rivers before we come to them. Zulliver may indeed be a counter example to certain oversimplified accounts of what some words mean, but we need not suppose that all questions about how to describe it can be settled in ignorance of the inside story.

Mr Kirk's view is that our knowledge of the Brain Team makes it reasonable to suppose, but does not entail, that Zulliver is insentient despite his gullivanting. Similarly, knowledge of how an ordinary person moves, how he looks, sounds and is made, makes it reasonable to suppose, but does not entail, that he is sentient. What does entail it? The presence of non-physical parts would presumably be an unhelpful further description of structure. Mr Kirk's answer lies in what he says about 'privileged access'. He seems, in avoiding ghostly structures, to offer us Descartes without the horse: "We all know *we* are sentient; and no doubt we also know there are no little men

inside *our* heads. Yet the first term of knowledge does not entitle us to exclude *a priori* the possibility that we might have been mistaken about the second. My knowledge that my head *is* is not, so far as I can see, logically or conceptually incompatible with the discovery that the inside of my head is like the inside of Zulliver's".

It is surely absurd to attribute knowledge or headaches to puppets, so we are not being asked to speculate, "Maybe I am a super puppet with resident puppeteers". Mr Kirk's 'cogito' is, rather, "I (know that I) am sentient, even though the inside of my head is exactly like that of a puppet except that someone is inside who is responsible for all those movements and characteristics which it would be reasonable for an outsider to interpret as actions or expressions of mine". The question is whether something with such a constitution can be said to speculate. If speculative noises or marks that issue from it will be by courtesy of the inner man. Certainly there could be slugs and snails and little green men inside Gulliver; but it is not obviously illegible to suppose that the little men produce all Gulliver's personal reactions. It begs the question to insist that at least Gulliver could doubt whether he is sentient. If he doubts, I doubt a person. But if little green men produce all the changes in the entity we took to be Gulliver then it seems inappropriate to attribute doubt here.

The description of Zulliver ("A team of their scientists invaded his head . . . so as to produce behaviour indistinguishable from that of a normal man") cannot, on Mr Kirk's own view, be a *physical* description, since we cannot have scientists, invaders, mimics or chiefs without consciousness. A physical description would have stayed at the level of ganglions and haemoglobin, of this electrical impulse producing that chemical reaction. When we talk of a Brain Team we have already interpreted the physical phenomena; it is this interpretation which makes us suppose Zulliver insentient.

If the movements of an alleged chess-playing machine caused by an internal or external controller we deny that the machine plays chess. It is tempting to suppose that when we say a person *does* play chess (think hard, miscalculate etc.) we must be referring, however vaguely, to some other cause

story which is characteristic of human behaviour. Then it is easy to believe that internal structure and functioning are of the utmost importance to whether a given 'entity' is a person. Thus Hobbes' view that "life is but a motion of limbs, the beginning whereof is in some principal part within" is echoed and elaborated by contemporary theorists.

But it is important to notice that the mere fact of causation by inner men did not lead us to deny that the machine played chess (or that Zulliver was sentient). If the machine needed a man inside to keep the air moving at the right temperature or if the man had to pull levers to move the pieces in accordance with instructions printed out inside it, this would not impute its chess playing capacity. What settled it was that the insider was deciding what to play. The moves were *his* moves not the machine's. Similarly, we interpret the movements of Zulliver as the mimicry of the Brain Team not the antics of Gulliver. These are competing action-descriptions, not alternative causal stories. It may be said that the difference between the actions itself hangs on what caused the movements. But this view now gains no support from the reflections in the last paragraph. Roughly, if it is said that an action is a bodily change caused in a certain way by the brain (not by little men etc.) there seem to be counter examples where the causal role of the brain does not guarantee an action and where the causal intervention of little men etc. does not mean that the movement loses the name of action. The beginning of motion not only has to be within: it has to have a life-giving character. Thus it is that the brain, on many accounts (especially in psychology), is said to think, decide, receive and send messages, operate limbs and so forth. But those are precisely the kind of personal descriptions which a causal account was intended to elucidate. This seems to me a fatal inadequacy in such an account.

It is Mr Kirk's view that sentience depends on "private access", which is logically independent of the physical facts. Psychological descriptions requiring private access are said to be such that "for their correct application on any given occasion, one individual is in an epistemological unique position: a position where he, but no one else, is able to apply the description correctly without needing access to what is publicly observable". This explanation certainly lets in too much. If a man,

but no one else, hears his wife's voice in the next room he is now probably able to apply many non-psychological descriptions that others can't apply. He may also know that in about an hour the door will open (for he has decided to go out then). But "the door will open" is not a psychological description.

In general, talking of private access and contrasting this with public observation points misleadingly to the common idea that sentience is a matter of private observation, whereas it is not a matter of observation or detection at all. What can be observed or detected can also be unobserved and undetected. This is not true of sensations. Nor is having sensations a matter of being able to apply a certain kind of description, or else animals, let alone prawns, would be indubitably insentient. But, having finally arrived at the most sensitive area in the ticklish debate about zombies, I scratch about in vain for fresh observations.