ROCZNIKI FILOZOFICZNE
Volume LXIX, issue 1 – 2021
DOI: https://doi.org/10.18290/rf21691-2

PAUL SNOWDON

SWINBURNE ON PHYSICALISM AND PERSONAL IDENTITY

Professor Swinburne's book *Are We Bodies or Souls?* is a very rich discussion of some fundamental issues in the philosophy of mind. His answer to his own question is that we are souls, and the complex, many-staged, and carefully formulated argument presented here leading to that answer is the improved descendent of earlier discussions by him supporting in a similar way the same conclusion. My own sympathies are, rather, with the answer that we are bodies. Indeed, we once had a seminar in Oxford considering these issues, and more recently we would sometimes encounter each other walking in Oxford and continue our discussion. Not surprisingly, agreement was never the outcome. However, my aim here is not to argue in favor of my own earlier metaphysical prejudices. I wish, rather, to voice simple skepticism about two stages in his argument. The two stages are his renunciation in chapter 2 of physicalism, and his endorsement in chapter 3 of the Simple View of Personal identity. By "simple skepticism" I mean the claim that neither thesis is convincingly supported. I leave unexplored here what consequence for his pro-dualist argument there would be if my skepticism about these stages is reasonable.

SWINBURNE ON PHYSICALISM (CHAP. 2)

Although this is not the beginning of Swinburne's discussion of physicalism, he has, by page 26, arrived at the version of physicalism that he analyses in greatest

Paul Snowdon, Emeritus Professor at University College London; address for correspondence: Department of Philosophy, University College London, Gower St., London WC1E 6BT, UK; e-mail: p.snowdon@ucl.ac.uk; ORCID: https://orcid.org/0000-0003-0057-7849.

detail. This attention is appropriate because the version is the one, or at least close to the one, that many philosophers sympathetic to physicalism would affirm today. He calls the view the "mind–brain supervenience theory" which I shall abbreviate to MBST. As he puts it, "the theory holds that the occurrence of every mental event is involved in the occurrence of some brain event, and so when the brain event occurs the mental event occurs, and so the mental event is nothing extra beyond the occurrence of the brain event."[1] Later he says: "If this theory were true there would be no more to the history of the world than physical events" (27).

Now, why does Professor Swinburne think MBST is not true? He has two main reasons. The first reason is presented on pages 27 to 28, and in effect claims that the MBST is committed to an entailment that does not obtain. The second reason is thought of as flowing from what Swinburne calls the "source of the distinction between mental and physical events" (28).

I want initially to engage with the first reason. I am though a little uncertain how to formulate the argument, but I shall represent it as resting on two premises. The first is that if MBST is true, then there are entailments from neural physical claims to claims about mental features, and the second premise is that there are no such entailments. Now, this sort of argument as a general structure is a very familiar thing. In recent philosophy of mind one well known and extremely influential example is Kripke's argument in chapter 3 of *Naming and Necessity*. Indeed, Kripke's version was influential because prior to his discussion physicalists formulated their views employing the notion of identity and said or implied that their proposal should be seen as a contingently true identity. Kripke's novelty, or one of them, was to argue that this sort of identity has modal implications. He then argued that the modal implications were false. Another variant of the same sort of argument is the so-called zombie argument. As we might put it, it seems to Swinburne perfectly possible for there to be a creature with the physical makeup of himself when he, Swinburne, is in pain but who does not experience anything at all, something which might be called "a zombie" resembling Swinburne in every physical respect. These variants differ in what they bring forward to support the second premise in the general argument.

Why should we accept Swinburne's rather streamlined version of this type of argument? What does he offer in support of the second premise? His support becomes clear in chapter 5. There he writes: "I defined a sentence as logically possible iff it is not a contradiction and does not entail a contradiction. Whether or not a sentence entails a contradiction depends only on the rules of the language and can be deter-

[1] Richard Swinburne, *Are We Bodies or Souls?* (Oxford: OUP, 2020), 26. All subsequent parenthetical references are to this volume.

mined a priori; that is determined by a competent speaker of the language" (113). So, in the language employed in chapter 2 Swinburne's reason for denying any entailment is the conviction that a competent speaker cannot work it out a priori. Why is Swinburne sure a competent speaker could not work it out a priori? One remark he makes is that if a traveler from another planet had "learnt everything about human bodies and brains" that traveler could still "rationally" wonder whether a knife wound causes pain in a human or just causes the human to scream (28). Now, it seems to me that this remark about the creature from outer space is reasonable, but does the ability of the creature to *rationally* wonder about humans show there is no entailment that armchair reasoning can detect? Surely there can be cases where, say, P entails Q but working out the entailment is very difficult and so someone can reasonably wonder. Another thing to say is that Swinburne has not scrutinized the full range of potential initial premises to determine what armchair reasoning can extract from them. No philosopher could supply any such premises. But even if Swinburne were acquainted with such candidate premises, how could he determine that no armchair reasoning could get from them to a conclusion about consciousness? It is clear that if Swinburne wants to defend the argument as interpreted, he must have a theory of armchair inference. Now, if we go back in the text to his account of what might be called deductive reasoning (14–16), Swinburne remarks evidently sensibly "that some sentence s1 entails another sentence is sometimes very obvious, but when this is not obvious you can often show it by producing a chain of entailments; that is by showing that s1 obviously entails s1a, and s1a obviously entails s1b, and so on until you reach s2" (14). Taking our lead from this conventional description, the person in the armchair is faced with the following task. He or she has been given (somehow) a very long physical description of Swinburne's when he is in pain. Let us call that long claim SBP (short for Swinburne's Body in Pain). The hoped-for conclusion is: a human is in pain. Let us call that P. Looking at these two things the person in the armchair is stymied. What they need is a series of intermediate claims linking the starting and end point. Swinburne is sure that no one can solve this problem. But maybe it is like some mathematical proofs which are sought for centuries and finally located by a genius. Either Swinburne at this point proposes a theory of deductive reasoning which reveals that no such deduction can ever be discovered or all we can say is that we shall have to wait and see.

A more significant problem, I think, faces the first premise. Why does Swinburne think that MBST commits itself to there being an entailment in his sense from SBP to P? The significance of this question is that the premise must mean by "entailment" a link that is a priori recognizable in virtue of understanding the language. Swinburne might reply that that simply is the theory he is discussing. The issue then is transformed into the question: Why formulate the physicalist proposal

in this way? At this point the physicalist as I interpret the view might make a speech along the following lines. Mental processes such as Swinburne's feeling a pain at t are occurrences in the world; similarly, we can ask: What is the real nature of these occurrences? What are the constituents which make up this occurrence? In their purport such questions are similar to: What is the nature of water? What is the nature of life? We, the speech makers continue, have a highly general proposal: the actual nature of the occurrence of such a pain is that it is a complex and special physical happening. An implication of this proposal is that since the physical happening amounts to the feeling of pain, it was, on its own, enough for the pain, and so it could not have occurred without pain. So the proposal has a modal implication. However, it is not meant to be a modal truth that anyone sitting in an armchair could recognize. It is recognizable only if the theory itself seems to be true. So, let us now look for substantial evidence for and against the theory. This clumsy short speech is offered as elucidating the theory, justifying the modal implication, and explaining why armchair cogitation is irrelevant.[2] Now, in chapter 5 Swinburne is prepared to follow Kripke in allowing that it is impossible for water not to be H_2O (113). He seems to hold that there was no way for someone who merely understands the language to recognize this impossibility despite its being an impossibility. If he allows that then he is allowing that the proponent of physicalism is entitled to claim to be read in that way. The necessity that P is true given SBP is an a posteriori necessity.

I turn now to the second reason that Swinburne offers to reject MBST. He begins section D in chapter 2 by saying: "While I believe that the conclusion in the last sentence follows from the simple argument given in the previous paragraph, that conclusion becomes even more obvious when we examine the source of the distinction between mental and physical events" (28). I do not have the space to follow the very detailed development of the distinction he is drawing but I think that the central point is well conveyed in the following sentences.

> On my definitions a "mental event" is an event to which the substance involved in the event has privileged access by experiencing it, and a "physical event" is one to which no substance has privileged access by experiencing it.… No physical event involves a mental event, for no event to which there is no privileged access could involve an event to which there is privileged access. (31)

[2] Longer expositions of the points in the speech can be found in Paul F. Snowdon, "On Formulating Materialism and Dualism," in *Cause, Mind, and Reality*, ed. John Heil (Dordrecht: Kluwer Academic, 1989), and in Paul F. Snowdon, "Philosophy and the Mind/Body Problem," in *Mind, Self and Person*, ed. Anthony O'Hear (Cambridge: CUP, 2015).

I want to respond by querying some of the things that Swinburne says when defending his characterization of the mental and the physical, and then to raise a question about the logic of the argument I have just quoted.

Swinburne claims that "what makes an event a physical event is that it is an event to which no one substance can have privileged access by experience" (30). A problem case is that we seem to have privileged access to physical facts about our bodies. I have a way of knowing that my leg is bent by *feelin*g that it is bent which no one else possesses as a way knowing that physical fact. We have internal routes of access to such facts about our own bodies which another person does not have. For a second case consider the physical fact that there is a tree in front of me. Now, that does seem to be a physical fact. Others can tell that it obtains. However, their way is to see me and the tree in front of me. My way can be different. All I need to do is to see the tree, and I thereby know it is in front of me. I do not need to see myself as well. Perception then seems to provide examples of privileged access to lots of physical facts. A further type of case might be called knowledge-dependent privileged access. Suppose I know that if I feel pain then chemical C is in my body. Then I am able to know that C is in my body, a physical fact surely, in a way that others cannot know it about me. A less recherche case is that I can tell my arm has touched a nettle, surely a physical fact, in virtue of feeling the distinctive pain of a nettle sting. Others cannot, in that way, know that physical fact about me.

How do things stand with the characterization of the mental? In discussing this it is not easy how to apply it. Officially it is this: "What makes an event a 'mental event' is that it is an event to which the substance involved in the event has privileged access by experiencing it" (29). Swinburne illustrates this property in the following quotation: "But I have an additional way of finding out whether I am seeing the tree or have the thought that today is Thursday or am intending to shut the door, by actually having the experience of seeing the tree or having the thought or intention" (29). There are, it seems to me, two puzzles here. Take the case of seeing the tree. Now, as Swinburne says, when someone sees a tree they have the experience of seeing the tree. That is because seeing a tree is an experience which the subject is having. The question is: Does the subject know about the experience *by* experiencing *it*? Saying "yes" seems to imply that one knows about the experience which is seeing by having a further experience of that experience. That does not seem to be true. It seems more accurate to say that we have a way of knowing about the experience, which others lack in relation to that experience, but it is not by experiencing that occurrence.

The second puzzle emerges, I think, if we consider the case of awareness of belief. Swinburne says that our awareness of our beliefs fits his characterization. He says: "But I can bring these beliefs, desires, and intentions to consciousness by

asking myself what do I believe about so-and-so,… that some person has privileged access to some belief by being able to experience having it is what makes that belief his or hers" (37). What these remarks seem to imply is that by asking what I believe about so-and-so I bring it about that I actually experience having the belief about so-and-so. This is hard to accept. What might happen if I ask myself: What do I believe about Einstein? What might come before my consciousness is something like the words "Einstein is a genius." I may be allowed thereby to know that I believe that Einstein is a genius. But is the occurrence of those words in my experience there and then a matter of *experiencing having the belief*? Would I think: I have just had a glimpse of my holding a belief? Clearly having that sentence before my consciousness cannot be what having the belief is, since I have the belief permanently, whereas the sentence is not there. Further, if that sentence should come before my consciousness sometime that would not mean it was a belief. And yet what is presented to consciousness is no different between the two cases. It seems to me that it may be that Swinburne has described a way that a subject has for becoming aware of what he believes, but the method cannot really be described as experiencing having the belief.

These rather quick remarks are intended to generate some doubt about the application of Swinburne's talk of having a privileged access by experiencing the mental event.

However, things stand with these remarks, I want to make a logical point about the general argument here. If we ask why Swinburne claims that physical events are things to which there cannot be privileged access, the answer would be that he checks this against a comprehensive list of physical events and finds that none permit privileged access. I have queried the outcome of that survey but let us leave that doubt aside. Now, Swinburne is arguing on the basis of this premise that mental states of affairs cannot be physical, so when compiling his list of physical events, he does not yet know whether there are mental events with a physical nature. He cannot therefore suppose he is entitled to the completely general claim that no physical state of affairs is available for privileged access. Until he already has his conclusion, he cannot have his premise.

We can add a more constructive suggestion. If we think of human experience as having a physical nature, we shall also think of human cognition, for example, knowledge and belief, as also having a physical nature, so privileged access can be explained by postulating physical links between the former physical events and the latter within a single human being, which are not available between that experience and another human being. Of course, I cannot fill in the details of these links, but physicalism and channels of privileged access are not irreconcilable and alien. They fit rather nicely together.

My response then is to query both of Swinburne's arguments against physicalism.

THE SIMPLE VIEW OF PERSONAL IDENTITY

I want in this final section to try to provide some resistance to accepting Swinburne's conclusion at the end of chapter 3, which is that the simple theory of personal identity is correct. Now, this plainly is a difficult topic, but are we obliged to end up where Professor Swinburne does?

At the beginning of chapter 2 Swinburne says something relevant to his arguments about personal identity which we perhaps should not accept. He says: "This book is about human beings." That is, of course, undeniable, but he then says what human beings are: "[They] are persons belonging to the same biological group as ourselves" (12). This proposal is that human beings are a sort of person tied to a biological group. What though is a person? Swinburne's definition is: "A being who has the capacity (or will have the capacity as a result of normal developmental processes) to have events…, including occurrent thoughts,… about past and future … moral beliefs … and the ability to do some very simple logical reasoning" (12). Swinburne provides no reason to define 'person' in this way, and his definition includes a feature that is not normally inserted—namely having moral beliefs—to which one might say: Cannot there be amoral persons? And it leaves out features which are often included—namely self-consciousness and memory—even though Swinburne intends his account to fit "normal usage" (13). More important though is why 'human being' should be regarded as a restriction on the more basic category of 'person'? Human beings seem to be a type of natural creature with a myriad of distinctive features, a natural kind to which we belong. It is as odd to define human beings as a type of person as it would be to define orangutans as a type of 'zerson', where 'zerson' is defined as a "creature with a certain restricted range of mental capacities (to be specified), which can be shared by other creatures." Further, defining human beings that way cannot be taken to reveal anything about human beings, that is to say, us. Thus, I can ask: Might some human beings not be persons? Might some human beings cease to be persons? These questions should not be regarded as settled by straight off defining 'humans' as a type of person and defining 'person' in some strict way. Swinburne's definition of person cannot be allowed to have any influence on what a human being is, and so not on what we are. But later we find that it does. Swinburne says: "The obvious objection to the view humans are physical substances … is that if we ceased to have the capacity to be conscious, we cease to exist. A dead person, no longer capable of being conscious, is not a person. So having this capacity is an essential property of humans" (39). The problem is that there is no reason to regard humans as types of persons as defined by Swinburne.

The Simple View that Swinburne supports says: "There are no necessary or sufficient conditions for personal identity in terms of degrees of any features of which there can be degrees" (65). The chief obscurity in it as far as I am concerned is quite what features are ones for which there can be degrees. Thus, as some propose, if the continued existence of a human being requires the continuous presence of life, which is standardly regarded as a purely physical process, is that a feature that comes in degrees, or is it a feature that escapes Swinburne's claim? I am unable to settle that issue and so shall leave it obscure. I want, rather, to bring forward the following highly commonsensical proposal. If we take a typical human being with a typical life-history, who develops physically over time in the way that countless human beings have, eventually and sadly ceasing at some point to exist, and we label that physical trajectory HPT, then if a human being lives through that physical trajectory it is a single human individual existing over time. That physical trajectory is a constitutively sufficient condition for what we can call personal identity. If, on the other hand, a human being steps on a mine and is physically obliterated with his or her parts scattered, then that physical sequence constitutes that human being ceasing to exist. So, a necessary condition for a person to remain in existence is that it does not suffer that physical sequence. Swinburne thinks, if I have understood, that these proposals, unambitious though they both are, can be ruled out. On the face of it that is somewhat surprising. It is not surprising that these proposals should be rejected since many people hold a conception of themselves which reject them. What is surprising is that there are some fairly easily presented reasons to rule them out prior to a proper empirical investigation of our actual nature.

How does Swinburne make a case for rejecting them as possible constitutive truths? Putting it vey generally, in the first half of the chapter (42–52) he argues that if constitutive proposals are to be plausible, they must focus on brain continuity and mental continuity, and then argues (52–64) that because both of these features or elements can come in degrees that generates some unacceptable consequences, even if the theories are made somewhat weaker than they normally are.

Swinburne highlights the importance of the brain in the attempted analysis of personal identity by endorsing what is normally called a brain transplant argument. This type of argument is certainly powerful but also hotly debated. Swinburne's own formulation of it is question begging. He says: "Yet if surgeons take your brain and put it into another body, would the resulting body now be your body, or would it still be the body of the person whose brain has been removed? It would surely be your body, since you could not move its parts in the same way as you could previously move the parts of your earlier body … and so the resulting person would surely be you" (46). The question raised by a transplant story is whether the person who consists of the transplanted brain in a new body is the person who donated the brain

or not. That question cannot be settled by *describing* it as an outcome in which the brain donor controls that body. The issue precisely is who is controlling the body. But this brings out an interesting aspect of Swinburne's discussion. When brain transplants are discussed normally the verdict that is elicited is elicited in response to some description of the case. The point of the example is to support a conception of the features which constitute personal identity. Thus, in Shoemaker's original presentation of it what is stressed in describing the case is the memories in the new person will be those of the person who donated the brain. Shoemaker therefore took the verdict that is usually elicited as supporting a psychological continuity theory, and not a brain-based view, even though the example is a brain transplant. The problem facing Swinburne here is that he wants ultimately to conclude that neither brain continuity nor psychological continuity constitutes personal identity, and so he cannot stress a feature to motivate the verdict the described case is meant to elicit. In effect his presentation simply begs the question. He does stress something else. He argues that the body left behind without a brain cannot be a person because that entity cannot have consciousness and so cannot be a person as he has defined 'person'. This does not carry any weight. I have already argued that Swinburne's definition of 'person' is irrelevant. Further, if that object is no longer the original person, it does not mean that the new person is the original person. Do we actually know who it is?

Swinburne then introduces another transplant case (48). The example, if I have got it right, is that you have an identical twin. Scientists remove the left hemisphere of your brain and place in your skull the left hemisphere of your twin which they have also removed. So, there is an equal amount of brain from you and from your twin. Sadly, the brain theory is now stymied. Swinburne then simply says: if the result should be that the new complex structure remembers more of your twin's life than yours, we should judge it was the twin. This seems rather quick. Is not the person who has lost half a brain simply your twin with less brain power? Would no weight be given to the fact that the half brain from your twin is put into your body? Is it not time to rethink the earlier conclusion that the physical continuity only concerns the brain? Basically, the comment I think that is warranted is that Swinburne's verdict here will not strike many as right, and we have wandered into a restricted conception which is already mistaken.

Within a few pages Swinburne's central points are presented. The first, which is the one I shall concentrate on, is that specifying precise values for the brain and mental continuities to constitute personal identity will be totally arbitrary (52–53). There is, that is, no non-arbitrary way to precisely specify conditions for personal identity. Now, this is an interesting point. Let us suppose, for the moment, as

I suspect Swinburne thinks, that with personal identity there cannot be combinations of constituting circumstances where, as one might say, there is no determinate answer. If we add that assumption to the restricted framework that has been adopted where the analysis has to be in terms of brain continuities and psychological continuities it would follow that there is some fact of the matter about identity for each combination of circumstances. Now, what would be arbitrary is a theorist specifying what is true in each case. How can we decide what the answer is in each case? However, that that would be arbitrary does not mean that there is no fact of the matter. If someone then asked why whatever is the answer is the answer, the response would be that it is simply a basic fact about personal identity, for which there is no more reason than there is as to why the basic causal laws of the universe are the ones that obtain. Explanations stop at some point. As a theory there is nothing evidently more absurd to this than there is to abandoning the idea that personal identity is constituted by more basic continuities. If we were to ask Swinburne what reason there is why the simple theory is correct, meaning "why is it true?", the answer would be that it just is!

There is a second point to make. Swinburne says that he is happy to abandon the idea that questions of identity always have answers. He says this is quite plausible for inanimate physical objects (56). The problem then for him is to show that this type of response is not available in the case of human beings. Swinburne responds to this on page 57, but one point he makes is that in a difficult case involving someone called Alexandra she will think that there must be a determinate answer. No doubt people tend to think this, but why should their conviction about determinacy be correct? The second point he makes is that it is logically possible that there is an answer. However, that there is no contradiction in supposing there is an answer does not mean that there is an answer. If the determinacy demand is rejected, then the arbitrariness objection is weakened.

There is considerably more to Swinburne's discussion, especially about personal identity, than the elements I have singled out and tried to respond to. But I hope that I have managed to put a question mark against some of Swinburne's claims.

BIBLIOGRAPHY

Snowdon, Paul F. "On Formulating Materialism and Dualism." In *Cause, Mind, and Reality*, edited by John Heil, 137–58. Dordrecht: Kluwer Academic, 1989.

Snowdon, Paul F. "Philosophy and the Mind/Body Problem." In *Mind, Self and Person*, edited by Anthony O'Hear, 21–38. Cambridge: CUP, 2015.

Swinburne, Richard. *Are We Bodies or Souls?* Oxford: OUP, 2019.

SWINBURNE ON PHYSICALISM AND PERSONAL IDENTITY

Summary

In chapter 2 Swinburne rejects physicalism for two reason. The first is that it is committed to entailments that do not exist. It is suggested that this reason is questionable both because there is no persuasive reason to deny there are such entailments, and also no reason to think that physicalism has such entailments. The second reason is that the mental involves privileged access by the subject and physical features do not allow privileged access. It is proposed that the physical does in fact permit privileged access. In chapter 3 Swinburne defends the Simple View of personal identity. The reasoning is very complex and rich, but it is proposed that Swinburne has not really shown that a reductionist account cannot be correct.

Keywords: Swinburne; physicalism; entailments; privileged access; personal identity; Simple View; Reductionist View.