

WILLIAM HASKER

### SWINBURNE'S *ARE WE BODIES OR SOULS?*

Richard Swinburne's *Are We Bodies or Souls?*<sup>1</sup> is a remarkable book. This comparatively brief volume restates and re-argues the main contentions of earlier writings, including two earlier books,<sup>2</sup> as well as some conclusions that are new in this volume. It does this in a way that Swinburne rightly considers more widely accessible than his earlier work on the topic. And this is done in an unusually systematic way, even for Swinburne. All necessary concepts and principles are built, as it were, from the ground up; we do not need to go back and consult other writings of his, or to rely on a supposed consensus among contemporary philosophers. What we need to understand and evaluate Swinburne's contentions is right here waiting to be understood and accepted—or not.

As it happens, I do agree with Swinburne on many of his central conclusions. I agree that there are pure mental properties, distinct from, and not entailed by, the physical properties of things. Even more controversially, I agree that there are mental substances, distinct from physical substances and causally interacting with them. Agreement on these central points, however, does not imply agreement on all the details, much less agreement on all of the arguments by which the conclusions are reached. In this essay I shall discuss, and criticize, two important arguments Swinburne uses in reaching his conclusions. I shall then discuss his views on a topic of central importance on which we partially agree—the origin of souls.

---

WILLIAM HASKER, Professor Emeritus of Philosophy at Huntington University; address for correspondence: 2753 W 450 S, Huntington, IN 46750, USA; e-mail: [whasker@huntington.edu](mailto:whasker@huntington.edu); ORCID: <https://orcid.org/0000-0002-9892-3633>.

<sup>1</sup> Richard SWINBURNE, *Are We Bodies or Souls?* (Oxford: OUP, 2019); page references in brackets are to this volume.

<sup>2</sup> Richard SWINBURNE, *The Evolution of the Soul* (Oxford: OUP, 1986); SWINBURNE, *Mind, Brain, and Free Will* (Oxford: OUP, 2013).

---

FROM PERSONAL IDENTITY  
TO SIMPLE IMMATERIAL SOULS

One important argument occupies most of chapter 3, on personal identity. The upshot of this argument is that a satisfactory theory of personal identity is only possible on the assumption that persons are pure mental substances—substances whose only essential properties are pure mental properties, properties that do not involve any physical properties. This is actually a rather dramatic conclusion. If there are persons, there must be personal identity, and so the argument entails that if there are persons, they must be pure mental substances; they must be souls rather than bodies. Swinburne goes on in chapter 4 to endorse an argument, revised from Descartes, for the affirmative conclusion that persons are in fact pure mental substances. But even without this additional argument, the argument based on personal identity gives him a good bit of what he is arguing for in the book as a whole. Of course, in order to establish this, he needs to show that no possible account of persons as being physical in any essential respect can produce a satisfactory account of personal identity. As we would expect, Swinburne proceeds to argue systematically that this is the case.

Swinburne aims to arrive at a “simple theory” of personal identity; correspondingly, the theories he opposes are “complex theories.” A complex theory locates personal identity in some combination of physical and/or mental properties of the organism. Swinburne assumes that a theory of personal identity aims to provide necessary and sufficient conditions for the persistence of a person over time. Unfortunately, however, complex theories are unable to accomplish this. The main reason for this is that continuity, both physical and mental, comes in degrees. This immediately gives rise to the problem of arbitrariness: if 55% physical continuity between  $P_1$  and  $P_2$  is necessary and sufficient for identity (assuming whatever mental properties may be required), why not 56%? Or 54%? There will also, for many if not all such theories, be the possibility of multiple candidates. If there are two later persons, each containing roughly half the matter of the original individual, and each has sufficient psychological continuity with that individual, we may have two resulting persons, each sufficiently similar to the original person to qualify as identical with that person. But both  $P_1$  and  $P_2$  cannot be identical with the same individual—and if not both, then neither, since their claims are equal. Swinburne states:

I conclude that no strong complex theory can give a plausible answer to whether some later person is or is not the same person as some earlier person in abnormal cases, where the later person is on the border between satisfying and not satisfying the requirements of the theory; and many such theories will have the additional problem that there can be two equally plausible candidates for being a certain earlier

person. So no strong complex theory can provide a universally applicable general answer to what constitutes being the same person. (55)

He goes on to say: "The only way for a strong complex theory to meet [these] objections ... is by denying an assumption made in the previous discussion that it is always either true or false that some later person  $P_2$  is identical with an earlier person  $P_1$ " (56). Swinburne terms this the "partial identity" theory, and discusses it at some length, considering versions proposed by Harold Noonan, Robert Nozick, David Lewis, and Derek Parfit. For our purposes, it will be most useful to quote an extended example proffered by Swinburne:

Suppose that Alexandra and two other unfortunate humans, Alex and Sandra, have been captured by a mad surgeon. The surgeon tells Alexandra that he is going to remove her cerebrum from her brain, divide it into the two hemispheres, put Alexandra's left hemisphere into Alex's skull from which her left hemisphere has been removed, and put Alexandra's right hemisphere into Sandra's skull from which her right hemisphere has been removed; the two hemispheres will then be connected to the other parts of the two brains of what were Alex and Sandra respectively. There will then at the end of the process, both the surgeon and Alexandra reasonably believe, be two conscious persons. The two resulting persons, they have reason to believe, will have equal degrees of physical and mental continuity (and connectedness) with the earlier Alexandra. The surgeon tells Alexandra that after the operation, he will kill one of these resulting persons; but he offers Alexandra the choice now of whether it is the person with her right hemisphere or the person with her left hemisphere who will be killed. Alexandra believes that the surgeon will do what she chooses, and she wishes to survive. How should she choose so as to have the best chance of surviving? On a 'semantic indecision' view, it won't matter which choice Alexandra makes, since whatever she chooses, it will be neither true nor false that she will survive. But surely there is a truth about whether or not Alexandra will survive. Or at least there may be a truth that Alexandra will survive or that Alexandra will not survive the operation if she makes a certain choice; and yet the 'semantic indecision' view has the consequence that whichever person is killed, that could not make any difference to the outcome of the operation. Yet it is surely (logically) possible, there is no contradiction in supposing, that, as with kidney or heart transplants, there is a truth about whether someone will survive an operation. It cannot be ruled out by some a priori philosophical theory that someone could survive an operation to remove some of their brain. When someone is about to undergo a brain operation of a more familiar kind, they normally hope to survive, and there seems to be nothing irrational in having that hope. How can this case be different, except in the respect that the doubt arises not from a doubt about whether the brain on which

the operation is performed will ever again be the brain of a conscious person, but from a doubt about whether that brain will be the brain of the same person as the person whose it was before the operation? (56–57)

The first thing to notice about this example is that the comparison with an ordinary surgical operation is a red herring. For an ordinary operation, what is in question is a biological process: will the heart, brain, and other organs continue to function in their usual, life-enabling, fashion after the operation has been completed? And it is clear that no a priori philosophical theory can have any bearing on the biological process in question. But with Alexandra, what is in question is a conceptual matter, and the truth of a philosophical theory can have a lot to do with who the surviving individual may be. Not, of course, that the theory is what makes it the case that the individual is or is not Alexandra, but a correct theory may be crucial for enabling us to find the correct answer to that question.

But what exactly is the argument that is being made here? Swinburne writes, “surely there is a truth about whether or not Alexandra will survive.” This, however, is persuasion, not argument. That “Alexandra will survive” is straightforwardly either true or false in the situation as specified is precisely what the partial-identity theory denies, so in saying that “surely” there is a truth about this Swinburne is merely repeating his own position and inviting the reader to find it plausible. He may seem to be providing more in the way of support when he says: “Yet it is surely (logically) possible, there is no contradiction in supposing, that, as with kidney or heart transplants, there is a truth about whether someone will survive an operation.” The appearance of support, however, is illusory. True, the bare proposition “Alexandra will survive the operation” is not, in itself, logically contradictory. But the partial-identity theory specifically asserts that, once the nature and apparent result of the operation have been specified, it is impossible that (there is no possible world in which) that assertion is either straightforwardly true or straightforwardly false. For Swinburne to ignore this assertion by his opponents would amount to begging the question. The most he can hope to accomplish with his example is to put forward a theory of his own that contradicts the partial-identity theory, and endeavor to persuade the reader that of the two theories his is the more plausible.

At this point I offer an example of my own. The person featured in my example is Tubby. Tubby is not a human; rather, his size and shape are approximately those of a rugby ball, or an American football. Tubby, however, has sufficient brain matter (or the analog thereof) to support all of the sorts of mental experiences persons normally have. And while his physical shape in a resting state is quite simple, he is able on short notice to extrude temporary appendages that enable him to perform all of the manipulations required to put his volitions into effect. There is, however, one further, very interesting feature of Tubby: he expects, in the near future, to un-

dergo fission around his midsection. The consequence will be two individuals, call them TubbyOne and TubbyTwo, each initially possessing approximately one-half of Tubby's matter, and each with a high degree of psychological continuity with Tubby. Each of TubbyOne and TubbyTwo will think of himself as carrying on Tubby's life, though of course neither of them can be identical with Tubby. In fact, each will have enjoyed continuous consciousness during the fission event, experiencing a notable but difficult-to-describe sensation of loss as the separation between the two "Tubby-successors" occurs.

It is noteworthy that Tubby does not experience the anxiety we have seen in the case of Alexandra. If he were to learn that one of the two "successor Tubbys" was going to perish, he would certainly regret this, but would be comforted by the thought that "he" would still live on, in the person of his other successor. But so far as his own persistence is concerned, it would make no real difference which of the two were to survive.

I submit that if Alexandra were to adopt a viewpoint analogous to Tubby's, her own anxiety would be considerably alleviated. She would, I believe, have achieved a satisfactory materialist answer to the multiple-candidate problem. And the arbitrariness problem also seems susceptible of a solution. There need be no arbitrary dividing line between the degree of continuity that is and the degree that is not sufficient for her survival. Rather, at some point it will be objectively indeterminate whether the remaining continuity is or is not sufficient to make it the case that Alexandra is still among the world's inhabitants. No doubt this would be an uncomfortable situation, for Alexandra and also for us. But there is nothing logically incoherent about it.

It now becomes Swinburne's obligation to show that this view proposed for Alexandra is unacceptable. Notice that it is not sufficient for him to claim that his theory is more plausible than Alexandra's. Plausibility in many cases is person-relative; furthermore, it is Swinburne who is making the argument, so it is up to him to make the case. There is however one further argument he puts forward that might seem to be relevant here. Suppose Alexandra learns that, subsequent to the operation, Alex has a very enjoyable experience to look forward to. Then since "the resulting Alex is partly identical to Alexandra ... then presumably Alexandra will have some good experience to which to look forward, but—since Alex is only partly identical to Alexandra—an experience which only a part of Alexandra will enjoy or which she will only partly enjoy or of which she will enjoy only a part" (60). Swinburne goes on to argue that none of the three alternatives is acceptable. However, Alexandra's position as I have elaborated it does not embrace any of those alternatives. Rather, it asserts that Alexandra can look forward to the experience because it will be experienced in her continuing life in her successor Alex.

In view of this, I do not believe that Swinburne has so far fulfilled his obligation to show that Alexandra's view of personal identity is unsatisfactory. Notice,

however, that this is *all* I have claimed to establish with this discussion. I have not proved, or claimed to prove, that Swinburne's "simple view" of personal identity is mistaken, or that the view of the soul as a simple immaterial substance that goes with it is wrong. I do think I have shown that one cannot establish either of these views by claiming that there is no acceptable materialist theory of personal identity.

It should also be noted that, even were Alexandra to stick with her original, Swinburnian theory of personal identity, she would still have difficult questions to answer. Faced with the neurosurgeon's ultimatum, she would still need to decide, without evidence, which of the resulting persons she should expect to be after the surgery. Now, suppose she guesses that she will be Sandra—that is, that her soul will come to inhabit Sandra's body—and suppose she is correct in her guess. At this point she cannot help but wonder what will be the case with Alex, if for some reason she is not killed as was planned? Both Alexandra and the neurosurgeon have supposed that, immediately after the surgery, there will be "two conscious persons [viz., Alex and Sandra ... with] equal degrees of physical and mental continuity (and connect-edness) with the earlier Alexandra." But if they are correct in this supposition, some explanation for these facts is required. Why might Alex's body not be a mere automaton, with no conscious experience at all? Or why might Alex not have conscious experiences completely unconnected with those experienced earlier by Alexandra? After all, Alexandra herself—that is, Alexandra's soul—will exist only in the body of Sandra. If, however, Sandra and Alex do have "equal degrees of physical and mental continuity" with Alexandra, this will presumably include their having identical apparent memories. But for Alex, unlike for Sandra, these will be *false* memories since Alex (the Alex-soul) will never have had the experiences she thus "remembers." There is, I suppose, nothing logically incoherent in all of this; logically speaking, it could very well all be true. But Alexandra is clearly left with some puzzling questions she would like to have answered. (Tubby, in contrast, has no such problems.)

#### FROM INFORMATIVE DESIGNATORS TO THISNESSES

Another important argument employs the concept of an "informative designator." This concept is distinctive to Swinburne, and he goes to considerable trouble to explain it. He introduces the term and its antonym in the following words:

I define a word as an 'informative designator' iff it is such that if we know what the word means, necessarily we know what is the object (substance, property, or whatever) to which it refers (its 'referent'), in the sense that we know what it is for an object to be that object and so we know what I shall call the 'essence' of that object.

I define a word as an 'uninformative designator' iff it is such that even if we know what the word means, we do not necessarily know the essence of the object to which it is referring; that is, know what it is to be the object to which it is referring. (86–87)

Swinburne continues by discussing further both informative and uninformative designators. Sometimes we know what a word means because it can be explicitly defined in terms of other words whose meaning we know. "But in the end, if we are to understand what some defined word means, the words by which it is defined need to be understood in terms of words whose meaning we know in some other way than by knowing a definition of them, and that other way involves being able to recognize straight-off objects to which they apply" (87–88). For example: "Being 'a molecule of H<sub>2</sub>O' is being 'a molecule consisting of two atoms of hydrogen and one atom of oxygen'" (87). Swinburne goes on to explain how each of the terms in that definition can themselves be defined using only terms such that we can immediately recognize that they apply or fail to apply.

It is evident that a lot depends on understanding our ability to recognize immediately that a term does or does not apply. Swinburne states:

In the case of words whose meaning we know straight-off and so are able to recognize under ideal conditions whether or not they apply, we know—simply in virtue of knowing the meaning of the word—what it is for the object to which they apply to be that object; we know the logically necessary and sufficient conditions for something to be that object. For an object to be 'a door' just is for it to look, feel like, and behave like (e.g. open when pushed) paradigm instances of doors.... Hence these words whose meaning we know straight-off are all informative designators. (90)

And on the other hand, "by contrast, an 'uninformative designator' is a word (or longer expression) which is such that if we know what the word means (that is, the meaning which is common to its use in different contexts), that is not by itself enough to know to what it refers on a particular occasion of its use; that is, to know what it is for an object to be that object" (91). Most definite descriptions and indexicals are uninformative designators; in such cases, one needs information about the circumstances, over and above what is conveyed by the meanings of the words, to know to what these expressions refer on a given occasion of use. And "there are other uninformative designators, such that the essence of the object to which they are referring depends on some underlying fact which may be totally unknown to anyone" (92). For example, we now know that the essence of water is to be H<sub>2</sub>O, so for us, 'water' is an informative designator. Earlier, however, the chemical composition of water was not known, and "the word 'water' was used in the early nineteenth century as a designator of the actual transparent drinkable liquid prevalent in our rivers and seas, and of whatever has the same chemical essence as that liquid" (92).



But contemporary science has not by any means eliminated the use of uninformative designators. “Many of the terms of astronomy or particle physics apply to objects solely in virtue of their causes or effects, while physicists remain ignorant of what being such an object consists in” (94). Summing up:

Whether a word (or longer expression) is an informative designator depends on the rules for its current use in the language. A word is an informative designator of some object iff the rules for its application to the object are such that speakers who know what the word means (= know what is common to its meaning in different contexts) thereby know what is the essence of any object to which it applies (that is, a set of necessary and sufficient conditions for being that object). It is an uninformative designator iff the criteria for its application to an object are such that it needs either further generally available knowledge including knowledge about the context of utterance (as with many definite descriptions and indexicals), or future scientific or other empirical investigation (which might or might not be successful) to determine what are the necessary and sufficient conditions for being that object. (94)

All this prepares for the main use Swinburne proposes to make of the concept of an informative designator. He argues that, for each person, the pronoun ‘I’, used to refer to oneself, is an informative designator. The same is true, furthermore, for one’s use of one’s own name to refer to oneself. Thus, Swinburne’s use of ‘I’ or of ‘Richard Swinburne’ to refer to himself is for him an informative designator, though it is an uninformative designator for anyone else who becomes aware of his use of these expressions. He states:

I suggest that, as used by each person, ‘I’ is an informative designator of themself. We mean by it this person who is currently aware of some conscious experience which he or she can informatively designate. For we are always able in ideal circumstances—that is, when (1) our faculties are working properly, (2) we are in the best possible position for recognizing ourself, and (3) we are not subject to an illusion—to recognize when some substance is ‘I’ and when it is not. Each of us is in the best possible position for recognizing themselves when they pick out themself as the subject of a current conscious event, as the person who is now having this pain or that thought. (105)

This technical conclusion has some very important implications. To know an informative designator for some object is to know the essence of that object; that is, to know a set of necessary and sufficient conditions for being that object. But such a set of necessary and sufficient conditions will entail the possession by the object of “innumerable other sets of logically necessary and sufficient conditions, most of which may be such that most of us are not clever enough or have a sophisticated



enough set of concepts to recognize the entailment.”<sup>3</sup> That is to say, any given set of logically necessary and sufficient conditions for being a particular object will entail all other logically necessary and sufficient conditions for being that object; one might say that all other such conditions are *in principle deducible* from it. However, given the limitations noted, we may in a particular case be unable to carry out the deduction that would show that some set of conditions is indeed a logically necessary and sufficient condition for being the object in question.

In view of this, we can set up an inference-schema; this schema will be designated Exclusion, because it can be used to exclude certain putative essential properties of persons and, *mutatis mutandis*, of various other sorts of objects.

Exclusion:

1. Anyone using the informative designator ‘I’ to refer to herself knows thereby her own essence, which constitutes a set of logically necessary and sufficient conditions for being that person.
2. This set of logically necessary and sufficient conditions for being that person entails all other sets of logically necessary and sufficient conditions for being that person.
3. If some property F is not entailed by the person’s use of ‘I’ to refer to herself, then F is not logically required in order for the person to be the person she is.
4. F is not entailed by the person’s use of ‘I’ to refer to herself. (premise)
5. F is not logically required for her to be the person she is (so, is not an “essential property” of that person).

Exclusion enables us to recognize that Descartes, for example, does not have any essential property that involves his embodiment, since no such property is entailed by his use of ‘I’ to refer to himself. The application of Exclusion in particular cases, however, is complicated by the epistemic limitations of us humans, as referred to above. It can sometimes happen that a particular property F may be entailed by the set of necessary and sufficient conditions for being a given object, yet due to those limitations we may be unable to recognize the entailment. Swinburne holds, however, that sometimes we are entitled to be certain that our inability to recognize an entailment is not merely a result of these limitations, but that in fact the entailment does not obtain. According to Swinburne, Descartes can clearly and distinctly conceive of his body’s non-existence (or, in Swinburne’s variant, of his body’s ceasing

---

<sup>3</sup> Personal communication from Richard Swinburne, in which he helped me to understand accurately this aspect of his views.

to exist while Descartes himself continues to exist (76ff.); thus, nothing involved in Descartes's body can be essential to Descartes's own existence.

Swinburne's conclusion on this point enables him to reach a further, quite significant, result, as follows:

The simple theory of personal identity has the consequence that two different persons (for example, Alexandra and the person who was not Alexandra) ... could have had (at each time in their lives) all the same physical matter of their bodies and the same physical and mental properties. So the two different souls which make the different persons the persons they are could have had (at each time) all the same mental properties. So it cannot be sufficient to make a soul the particular soul it is that it has had a certain past mental life.... It follows that persons, and so their essential parts, their souls, have what philosophers call 'thisness' (= haecceity), which makes them and so their soul the particular person and so their soul the particular ones they are. (108–9)

This in turn leads to a further, and quite dramatic, conclusion:

It follows that instead of me living the life I have lived, (it is logically possible that) there could have been a different person, different from me in virtue of having a different soul from my soul, who lived a life which was the same as my life in all qualitative respects. Indeed, there could have been any of an infinite number of different souls and so persons from myself who lived a life which was the same as mine in all detailed respects. (111)

As we shall see, this result leads Swinburne to take a very distinctive position concerning the origin of each person, that is, of each soul. But are (possible) souls being multiplied beyond necessity?

There can be no doubt that the entailment enshrined in Exclusion is valid. However, I see no good reason to accept the assumption made in step (2) of that argument, that a "set of logically necessary and sufficient conditions for being that person entails all other sets of logically necessary and sufficient conditions for being that person." Why should it be the case that essences and essential properties all entail one another in this way, so that, given an essence of some object, we have entailments to all of the object's essential properties? I maintain, on the contrary, that *A set of necessary and sufficient conditions for being an object need not include or entail all of the object's essential properties.* I believe there are convincing counterexamples to Swinburne's assumption here. Take his own example, which affirms that the essence of water is to be H<sub>2</sub>O. Without doubt it is true that being composed of molecules of H<sub>2</sub>O is both necessary and sufficient for a substance to be water, and thus qualifies as an essence in Swinburne's sense. There are, however, compli-

cations that are hidden by that simple formula. For one thing, "water" includes not only ordinary water but so-called "heavy water"—water in which the nucleus of the hydrogen atoms is not a proton, but rather a proton joined to a neutron, the result being known as deuterium. Heavy water has the same chemical properties as ordinary water, but in certain applications involving nuclear reactions the difference is important. Any accounting of the essential properties of water would need to include the fact that water can involve two (actually three) different isotopes of hydrogen, and the different behavior of the different materials in certain circumstances. Yet none of this is even hinted at in the simple formula "Water is  $H_2O$ ," and there is no reason to suppose that any degree of logical skill and sophistication would enable one to deduce from that formula the facts about deuterium, tritium, and heavy water.

Another example is provided by the previous section of this paper. Consider again Alexandra, as she is wondering about the results of the experiment in which the two halves of her cerebral cortex are to be transplanted into two different bodies, resulting in two living persons. Alexandra, we may assume, is well able to conceive the outcome postulated by Swinburne, in which she, that is her soul, would be the soul of one of the resulting persons, whereas the soul of the other person would not be Alexandra but a different soul, one which shares all of Alexandra's memories in spite of the fact that she has never actually had the experiences thus "remembered." But Alexandra is also able to conceive the outcome that parallels our story of Tubby, in which Alexandra's life is in effect divided and continues in the lives of each of the surviving persons, neither of which is strictly identical with Alexandra. Both of these two incompatible futures are clearly conceivable to Alexandra, and there is no reason to suppose that either of them is entailed by the essence which Alexandra possesses in virtue of her ability to refer to herself using 'I'. Nevertheless, Alexandra is essentially such that one of these alternatives, and not the other, describes what may possibly be the case as a result of the brain transplant experiment. If so, we have a clear counterexample to premise (2) in the argument-form Exclusion.

Let us, however, return to Swinburne's claim that "persons, and so their essential parts, their souls, have what philosophers call 'thisness' (= haecceity), which makes them and so their soul the particular person and so their soul the particular ones they are." The concept of a thisness is a subtle and difficult philosophical concept, one which it would be a major task to discuss in detail. I believe, however, that it is possible to address the concerns Swinburne has raised in this section without invoking that concept, and if this is indeed possible that may lessen our inclination to believe that thisnesses need to be included in our account of the world. (I myself think there is no need for them, but arguing for that point in its full generality would be a large project.) The starting point for this alternative account is the point already made: *A set of necessary and sufficient conditions for being an object need not include or entail*

*all of the object's essential properties.* Now, let us grant to Swinburne that each of us can informatively designate himself or herself as "this person who is currently aware of some conscious experience which he or she can informatively designate." That designation, then, serves as a necessary and sufficient condition for "being me." It does not follow, I maintain, that one's essential properties are limited to those that are logically entailed by the designation in question. Now suppose that (as we emergent dualists<sup>4</sup> believe) my soul has emerged from a particular brain and nervous system, under a specific set of circumstances which include that brain and nervous system's having reached the stage of development at which such emergence is possible. Suppose, furthermore, that having emerged at that time in that particular way from that particular brain and nervous system is both necessary and sufficient for the soul's being the particular soul that it is. The emergence of a soul, that is, determines its existence as the particular soul that it is. Swinburne, of course, would deny this, but his reason for denying it seems to depend essentially on Exclusion, an inference-form which I maintain to be invalid. *If* this emergent dualist hypothesis is correct, then it is not true, as Swinburne has claimed, "that instead of me living the life I have lived, (it is logically possible that) there could have been a different person, different from me in virtue of having a different soul from my soul, who lived a life which was the same as my life in all qualitative respects." On the contrary, any soul that originated from the same brain and nervous system at that particular time and in that particular way just *would be* my soul, the soul that I now in fact have. And since this is the case, then it is not the case that there is need for a thisness to make my soul the soul that it is.

I need to make it clear that I have not, in the previous paragraph, claimed to prove that Swinburne's view is wrong and that mine is correct. What I have done is to present an alternative view, one that I believe has considerable merit, and one that some readers may, on reflection, find to be preferable to Swinburne's view. No doubt Swinburne is right when he titles this chapter, "We Know Who We Are." But it may not be equally clear that we know *what* we are.

## THE ORIGIN OF SOULS

The two previous sections have addressed major arguments proposed by Swinburne, arguments I have disagreed with. This final section concerns a topic on which he and I are in general, though not complete, agreement. But searching out the ways

---

<sup>4</sup> The primary text for my version of emergent dualism is *The Emergent Self* (Ithaca, NY: Cornell University Press, 1999).

in which we are and are not in agreement, and the significance thereof, requires careful attention.

Swinburne has a great deal to say about the nature and function of the soul; he says comparatively less about its origin. In the present book, what he does say about this is found in the final chapter, "Could Science Explain Souls?" After a summary of the results reached in the book so far, Swinburne asks when a human soul begins to exist. Noting that there is as yet no agreed theory concerning the neural correlates of consciousness, he states:

It would, I think, be fairly generally agreed that consciousness is caused by brain events interacting in a network of neurons constituting the cerebral cortex and thalamus. That network is in place in the brain of the foetus only by the seventh month of pregnancy. So a reasonable guess as to when humans are first conscious is that it is around the seventh month of pregnancy. If souls exist before then, they are not conscious. If a human soul exists before the seventh month in connection with the body of a foetus, it does not manifest its presence and so we have no good reason to believe that it exists before then. (145)

Swinburne now asks, "Could There Be a Scientific Explanation of the Existence and Life of the Soul?" (156). After discussing what is needed in general for a scientific explanation of some phenomenon, he states:

If there is to be a scientific explanation for the coming into existence of human souls, there must be some law of nature of the form that 'all human foetuses at a certain stage of development cause the existence of a human soul' (at that or an immediately subsequent time)... And if there is to be a scientific explanation of the continuing causal interactions between the brain and the soul, there must be innumerable laws about which brain events of which kind cause mental events of which kind (in the soul caused to exist by that brain), and which mental events of which kind cause brain events of which kind (in the brain which caused that soul to exist). (157–58)

Note that these claims are hypothetical: "*if* there is to be a scientific explanation...." Does Swinburne mean to imply that there *actually are* scientific explanations of these phenomena, and therefore that there are laws of the sort described? The text would lead one to assume this, but there are other considerations that might lead us to question it. One such consideration is a feature of the soul, as Swinburne understands it, that emphatically *cannot* be given a scientific explanation along these lines. Remember that Swinburne has said that "two different persons (for example, Alexandra and the person who was not Alexandra) ... could have had (at each time in their lives) all the same physical matter of their bodies and the same physical and mental properties." But this immediately entails that there could not be any law

stating that, under given conditions, there would be such a person as Alexandra. For in any such law, the antecedent of the conditional, the circumstances under which the person would exist, would be exactly the same for Alexandra and for the person who is not Alexandra—indeed, for any of the infinitely many other persons who are not Alexandra. There could be a law stating that *some person or other* would exist under those circumstances, but not a law that *Alexandra* would exist.

This problem is not one Swinburne has only recently discovered. On the contrary, already in *The Evolution of the Soul* he had asserted: “It needs either God or chance to allocate bodies to persons.”<sup>5</sup> (Chance is included here for completeness; I know of no evidence that Swinburne ever seriously considered attributing the allocation of bodies to persons to chance.) Would it not be reasonable, then, to infer as follows: “In order for a soul to come into existence, it must have a thisness, to distinguish it from all the other souls that could have existed in the exact same circumstances. But only God can be the source of such a thisness. Therefore, only God can bring a soul into existence”?<sup>6</sup> This interpretation is supported by another assertion in *The Evolution of the Soul*. After discussing the opposition between Creationist and Traducian views of the soul in the ancient church, Swinburne states: “However, I have argued against Traducianism and in favour of Creationism, the creation of each human soul anew by God who gives one to each embryo able to receive it. Indeed I have passed beyond the orthodoxy of Creationism to affirm the same in respect of animal souls.”<sup>6</sup>

It turns out, however, that this is one of the comparatively rare points on which Swinburne has undergone a change in his views, or at least a marked shift of emphasis. Commenting on a passage in his more recent writings,<sup>7</sup> where he affirms that “the development of a foetus ... causes the existence of the soul,” he states “my ‘creationism’ amounts to the fact that God has to intervene in this operation whereby the development of a foetus causes the existence of a soul (in a way codified by laws of nature) in order to determine which soul emerges.”<sup>8</sup> He goes on to state, “I agree that this is not the way which I expounded my ‘creationism’ in *The Evolution of the Soul*, p. 199; but it does make that view more precise and indeed justifies it.”<sup>9</sup>

<sup>5</sup> SWINBURNE, *The Evolution of the Soul*, 153.

<sup>6</sup> *Ibid.*, 199.

<sup>7</sup> The writing in question is Richard SWINBURNE, “Cartesian Substance Dualism,” in *The Blackwell Companion to Substance Dualism*, ed. Jonathan Loose, Angus Menuge, and J. P. Moreland (Hoboken, NJ: John Wiley and Sons, 2018), 133–51.

<sup>8</sup> From an e-mail dated November 22, 2016.

<sup>9</sup> Swinburne here seems to affirm that his present view falls within the scope of Creationism as referred to in his earlier writing, and thus does not amount to a reversal of that view. It seems fair to point out, however, that few readers would have understood the passage in *The Evolution of the Soul* in this way.

So we have here a clear and unambiguous statement concerning the senses in which Swinburne is, and is not, a Creationist. His Creationism, in fact, seems to be a variety of emergent dualism, in that he holds that the coming-into-existence of the substantial soul is caused by the development of the foetus. It is, to be sure, a distinctive variety of emergent dualism, in its insistence on the need for divine intervention to determine which particular soul emerges.

But the fact that we now have a clear statement of Swinburne's view does not mean that all problems have been resolved. How are we to conceive of the intervention by which God brings it about that Alexandra's soul, in particular, is created? An image that may come to mind is that of a manufacturing process, in which a machine turns out a part of some kind, and then stamps a serial number onto the part to identify it. But this can't be the right analogy: the part in question must already exist, in order to receive the serial number, whereas the soul must already have its thisness in order to exist at all; the notion of a soul that is no particular soul is nonsensical. Nor can the thisness be somehow provided by God at the beginning of the manufacturing process, so that the soul is then "built around" the thisness. This would imply that the thisness initially exists without being the thisness of any substance, which again is nonsensical. What we must conceive is that God is somehow *active within the process of soul-production*, so that this process, which in itself is incapable of producing any particular soul, will in fact produce the exact soul God has chosen for it to produce. Furthermore, God's being so active is a matter of *metaphysical necessity*; without this, the natural process would be incapable of producing any particular soul, and thus incapable of producing any soul at all.

Perhaps some readers are beginning to share my discomfort over this. The natural processes leading to the emergence of souls are of course subject to multiple contingencies, both those due to the free will of creatures and those that arise in the course of nature, especially quantum indeterminacy. The supposition that, on each and every such occasion, God is *metaphysically constrained to intervene* in order to permit the emergence of some particular soul looks very much like it might be some kind of infringement on divine freedom and sovereignty.<sup>10</sup> This awkwardness is partially masked by Swinburne's habit of speaking almost always of human souls, human foetuses, and the like. He acknowledges the existence of animal souls but refrains from saying anything definite about their prevalence. But a common-sense

---

<sup>10</sup> An alternative approach might be to state that, were God to refrain from intervening in this way, the soul-generation project would simply not be completed; no new soul would be produced. This of course would amount to "intervention by non-intervention"; by failing to act, God would cause an exception to the natural law concerning the emergence of souls. Readers who find this option preferable are welcome to adopt it.



judgment about this, based on everyday experience, will not tend to place the bar for animal consciousness, and therefore animal souls, especially high. It may be that bees and wasps experience nothing analogous to anger—indeed, experience nothing at all—when their nest is disturbed, but a person who has been pursued by them is unlikely to view the situation that way. But if insects have consciousness, and therefore souls, the occasions on which God is called upon to provide the needed thisnesses will be very numerous indeed.

Emergent dualists will view this situation with considerable interest. We are naturally delighted to have Swinburne as a fellow emergent dualist, agreeing with us that souls are produced by brains and nervous systems (or their analogues in simpler organisms) that have reached the appropriate stage of functional organization. Most of us will not follow, him, however, in requiring divine intervention in order that particular souls will be produced. We will agree with Swinburne that there is a law of nature (one whose detailed formulation still escapes us) to the effect that the development of the brain and nervous system produces at a certain point a soul. This soul, in virtue of being thus produced, is a particular, concrete, individual, and retains that individuality throughout all the later vicissitudes of its existence. Its so existing has no need for a special, God-enabled, thisness, though that soul, like everything else in the universe, is wholly indebted to God for its life and being.

## BIBLIOGRAPHY

- HASKER, William. *The Emergent Self*. Ithaca, NY: Cornell University Press, 1999.
- SWINBURNE, Richard. "Cartesian Substance Dualism." In *The Blackwell Companion to Substance Dualism*, edited by Jonathan Loose, Angus Menuge, and J. P. Moreland, 133–51. Hoboken, NJ: John Wiley and Sons, 2018.
- SWINBURNE, Richard. *Are We Bodies or Souls?* Oxford: OUP, 2019.
- SWINBURNE, Richard. *Mind, Brain, and Free Will*. Oxford: OUP, 2013.
- SWINBURNE, Richard. *The Evolution of the Soul*. Oxford: OUP, 1986.

### SWINBURNE'S *ARE WE BODIES OR SOULS?*

#### Summary

Richard Swinburne's *Are We Bodies or Souls?* presents a sustained case for a view concerning the nature of persons that can be classified as a form of either Cartesian dualism or emergent dualism. This paper comments on two important arguments developed in the book and concludes by considering the problem of the origin of souls.

**Keywords:** Swinburne; dualism; emergence; soul; thisness.