

## SWINBURNE'S NEW SOUL: A RESPONSE TO *MIND, BRAIN AND FREE WILL*

JAMES K. DEW, JR.

*Southeastern Baptist Theological Seminary*

Richard Swinburne's recent book *Mind, Brain, & Free Will* is a welcomed addition to the discussion of philosophical anthropology by those of us who are not convinced by physical or material accounts of human persons. And though the book makes an important argument for substance dualism, it has much to say about other important aspects of philosophy as well. As Swinburne notes, *Mind, Brain, & Free Will* advances, develops, and occasionally re-directs many of the arguments he first gave in *The Evolution of the Soul*.

In *Mind, Brain, & Free Will*, Swinburne's primary objective is to argue for a substance dualist perspective of human persons. On this view, persons are pure mental substances (or souls), which are distinct from any physical parts or properties and also distinct from any specific mental properties. In his words: 'Each person has a "thisness", a uniqueness, which makes them the person they are quite apart from the particular mental properties they have and any physical properties (and any thisness) possessed by their body.'<sup>1</sup> In this view, the soul of a person is her essential part, while the body is hers only contingently. He adds, 'My soul therefore carries my "thisness". However – given the normal understanding of a human being on earth as constituted (in part) by a body – it follows that humans, unlike other possible pure mental substances such as ghosts or poltergeists, each have their body as a contingent part.'<sup>2</sup>

---

<sup>1</sup> Richard Swinburne, *Mind Brain, & Free Will* (Oxford: Oxford, 2013), p. 165.

<sup>2</sup> *Ibid.*, p. 170.

As one who comes from a Thomistic perspective, there is much to Swinburne's book with which I agree. For example, I too think that we really do have immaterial souls that are different from our physical bodies and that we are capable of disembodied existence and surviving death. I differ with Professor Swinburne, however, over what constitutes a person and the contingency of our bodies. In my view, substance dualism seems to diminish the important role of the human body. But, my motives for that position are largely theological in nature, and as such I will focus on other things in this review.

For the remainder of this review, I will focus my comments on specific places in Swinburne's *Mind, Brain & Free Will* where I find myself unconvinced or having significant questions. Two such places come to mind: (1) Swinburne's gradual brain replacement thought experiment – what I take to be an important part of his overall argument in chapter six for substance dualism, and (2) the place and importance of mental properties in Swinburne's view of what it means to be a particular person.

#### SWINBURNE AND GRADUAL BRAIN REPLACEMENT

An important piece of Swinburne's case for substance dualism is his argument from the possibility of gradual brain replacement over time. He asks us to consider the following possibility:

Suppose that  $P_1$  undergoes an operation in which a small diseased part of his or her brain (a tenth of the whole brain) is replaced by a similar part from another brain (perhaps that of a clone of  $P_1$ ) .... But now suppose that each year a different tenth of  $P_1$ 's brain is removed and replaced by similar parts from another brain (perhaps that of a different clone of  $P_1$  on each occasion). At the end of ten years there is a person whose brain is made of entirely different matter. It seems at least logically possible that – because the process has been gradual and each new part has become integrated into the brain before a new operation is done – the resulting person is still  $P_1$ .<sup>3</sup>

Swinburne then adds a further condition: 'Now suppose that during each of the ten operations in which brain parts are replaced, the patient remained conscious and has a series of overlapping conscious experiences lasting for the whole operation.' To clarify, we could summarize it this way:

---

<sup>3</sup> Ibid., p. 155.

At  $t_1$ , Bob is conscious and his brain is composed of 1, 2, 3, 4, 5, 6, 7, 8, 9, & 10.

Then,

At  $t_2$ , Bob is still conscious, but his brain has part 1 replaced by part 11, such that his brain is now composed of 2, 3, 4, 5, 6, 7, 8, 9, 10, & 11.

Supposing this process repeated itself once every year for ten years as Swinburne suggests, we would eventually end up with:

At  $t_{11}$ , Bob is still conscious but his brain has part 10 replaced by part 20, such that his brain is now composed of parts 11, 12, 13, 14, 15, 16, 17, 18, 19, & 20.

It is important to note that Swinburne's thought experiment postulates a gradual replacement of a brain over the course of ten years. During any given surgical procedure, no more than a tenth of the brain is removed and replaced. If this is really possible, then Swinburne thinks that each of us 'can continue to exist without any continuity of brain, memory, or character. It follows that the simple theory of personal identity is true.'<sup>4</sup>

This is an intriguing thought experiment that – according to Swinburne – leads us to substance dualism. Yet, physicalists are unlikely to be persuaded by this argument. I for one – though not a physicalist myself – suspect that this argument is built upon an unstated assumption about consciousness itself. That is, it seems like this would only favour substance dualism if it were impossible – in the first place – for purely physical organisms to have conscious experiences. If they cannot have consciousness, then an organism going through the surgical process Swinburne describes is not conscious and does not have overlapping conscious experiences. As such, there is no person and no particular human identity. But if purely physical organisms can have conscious experience – which is something he does not seem to say much about – then they may also be able to endure the kind of gradual brain replacement that Swinburne describes while also remaining conscious and having 'overlapping conscious experiences lasting for the whole operation.' I suspect that a physicalist could account for these things for two reasons: First, this scenario is significantly analogous to the metabolic process our bodies use to gradually replace parts over time;

---

<sup>4</sup> *Ibid.*, p. 158.

Second, some physicalist philosophers have already developed theories that seem to allow for this very thing.

Consider the similarity between Swinburne's surgical process and the metabolic process our body employs to replace old parts with new parts over time. Physical organisms replace parts – through metabolic processes – all the time without ceasing to be the beings that they are. And yet, this happens even though the vast amount of the parts that compose the body at an earlier time are now gone and have been replaced by new parts. And yet, this happens while we are conscious and have overlapping conscious experiences. In short, we have our parts gradually replaced over time without losing consciousness or without ceasing to be the people that we are. Now if, through metabolic processes, physical organisms are able gradually to replace old material with new material over the course of time without ceasing to be the organisms that they are, then why wouldn't this be possible in the case that Swinburne has sketched out in *Mind, Brain, & Free Will*? What if physical organisms have the ability to gradually 'pass on' their identity to the new parts as they come into the organism? What this would require us to say is that a particular stream of consciousness – or first person perspective – is maintained by an ever-changing physical organism that gradually replaces old parts with new parts. This clearly happens in the case of metabolic processes, so why couldn't this also happen in the case a gradual brain replacement through a surgical process? One might say that there is a marked difference between the metabolic and surgical processes since one happens 'naturally' while the other is 'artificial'. But I suspect that this is not a sufficient difference since this is simply a difference in how the material arrives in the organism. Once it is there, the organism incorporates the material in the same way for both cases. All that is important here is that the physical organism be able to gradually change parts without the loss of consciousness or the life of the person in question.

Consider, also, some physicalists and materialists whose accounts of human persons might allow for Swinburne's scenario. Again, in his account a given person (1) has the parts of her brain gradually replaced over ten years until the whole brain has been replaced, and (2) the person maintains consciousness throughout the surgeries and has overlapping conscious experiences. Can physicalists account for these things? I suspect that Lynne Baker's Constitution View just might. In

her approach, the capacity for first-person perspective is the essential criterion for what it means to be a person. To be clear, having first-person perspective is more than just having some mental property or memory. It is the ability that one has to think of herself as herself. Or as she explains, 'A being with first-person perspective not only can have thoughts about herself, but she can also conceive of herself as the subject of such thought.'<sup>5</sup> In other words, first-person perspective requires more than me being able to have desires, intentions and plans, it requires me to realize that *it is me who has such* desires, intentions and plans. In Baker's Constitution View, a person can persist over time even if she does not have an immaterial soul, as long as she maintains first-person perspective. In other words, in light of changes in the human organism that constitutes a specific human person, Baker suggests that all that is necessary for the persistence of the person is continuity of first-person perspective and higher brain functions. She says, 'Suppose that a person slowly had her organs replaced by nonorganic parts, to the point where there was no longer metabolism, circulation, digestion, and so on, but the higher brain function remained and the person's sense of herself was uninterrupted. In this case, the person would persist but the organism would not.'<sup>6</sup> And so, in Baker's view, first-person perspective and higher brain function provide sufficient conditions for a person to persist over time. If her view is correct, then human persons – material persons – could maintain their consciousness throughout the gradual brain replacement scenario that Swinburne describes. As such, Baker's Constitution View might provide everything necessary to account for the persistence – and therefore survival – of a person who undergoes Swinburne's gradual brain replacement.

Peter van Inwagen's materialistic approach might also give materialists everything they need to reject Swinburne's conclusions. Like Baker, his approach allows for the parts of a material organism to change over time and for the organism – in this case a person – to persist throughout the change as long as at each successive stage of the process, the given parts are caught up into the same life as the original organism. He calls this the *Life principle*:

---

<sup>5</sup> Lynne Rudder Baker, 'On Being One's Own Person', in *Reasons of One's Own*, eds. Maureen Sie, Bert van Der Brink, and Marc Slors (Hampshire, UK: Ashgate, 2004), pp. 129-141 (p. 131).

<sup>6</sup> Lynne Rudder Baker, *Persons and Bodies: A Constitution View* (Cambridge: Cambridge University Press, 2000), p. 19.

If an organism exists at a certain moment, then it exists whenever and wherever – and only when and only where – the event that is its life at that moment is occurring; more exactly, if the activity of the *x*s at  $t_1$  constitutes a life, and the activity of the *y*s at  $t_2$  constitutes a life, and the organism that the *x*s compose at  $t_1$  is the organism that the *y*s compose at  $t_2$  if and only if the life constituted by the activity of the *x*s at  $t_1$  is the life constituted by the activity of the *y*s at  $t_2$ .<sup>7</sup>

Then, to clarify what *Life* would mean for the continuation of a particular material being, van Inwagen says:

Suppose that the activity of the *x*s constitutes a life at  $t$ ; suppose that a few of the *x*s cease to be caught up in that life and that the remnant continue to be caught up in a life; suppose that those of the *x*s that have ceased to be caught up in that life are ‘replaced’ – that certain objects, the *y*s, come to be caught up in the life the remnant of the *x*s are caught up in, in such a way that the *y*s and the remnant of the *x*s constitute that life. Suppose that this sort of replacement happens a sufficient number of times that eventually none of the *x*s is caught up in the life that has evolved, by continuous (and ‘insensible’, as Locke calls it) replacement of the *x*s, from the life that was once constituted by the activity of the *x*s. Is this life the life that was constituted by the *x*s?

In many cases, cases of the more usual sort, the answer is undoubtedly yes.<sup>8</sup>

Once again, this materialistic account of human persons allows for the gradual replacement of parts over time and for a person to persist through all of the changes that take place within her body. As such, I suspect that the thought experiment offered by Swinburne does not establish substance dualism and that materialists would reject his overall argument. And, despite my sympathies with his perspective, I am inclined to lean on other kinds of arguments that support the existence of the soul.

#### SWINBURNE AND MENTAL PROPERTIES

I also raise a second concern with Swinburne’s argument for substance dualism that, while not a defeater for it, is at least epistemologically worrisome. As he makes clear throughout chapter 6, there are no physical

---

<sup>7</sup> Peter van Inwagen, *Material Beings* (Ithaca, NY: Cornell, 1990), p. 145.

<sup>8</sup> *Ibid.*, p. 149.

or mental properties that are essential to a particular person's identity. That is, if Daphne at  $t_2$  is the same person named Daphne that was once at  $t_1$ , then Daphne persists because she has the same 'thisness' as the earlier person named Daphne, and that 'thisness' is not tied to any particular mental property (psychological continuity) or physical property. He says, 'I begin my defence of this position by arguing that it is logically possible that some person  $P_2$  at  $t_2$  can be the same person as a person  $P_1$  at  $t_1$ , even if he or she does not apparently remember anything done or experienced by  $P_1$  at  $t_1$  or earlier and has an entirely different character from  $P_1$ , and also has a largely different body (including brain) from  $P_1$ .'<sup>9</sup> He later adds:

Hence, given that an earlier person who had all the same physical parts as me, and all the same physical and mental properties as me, could, it is metaphysically possible, not be me, and could, it is metaphysically possible, be me, it follows that the difference must consist in the presence or absence of some non-physical part. I must now have a non-physical part (i.e. a part which is a pure mental substance) which makes me me, which the earlier person (even if they were in all other respects the same) would not have had if they had not been me. We may call this non-physical part of me my 'soul'.<sup>10</sup>

So in short, Daphne's persistence over time does not require the preservation of any of her mental properties (memories, characteristics, dispositions, etc.) or physical properties (particular body, particular body parts, particular brain, size, shape, etc.) across time. She could lose all her mental properties and physical properties and still be the same person she was at an earlier time. What is more, Daphne's original physical and mental properties could appear later in some other person other than Daphne. As such, we could have the following situation. At  $t_1$ , Daphne – who has the 'thisness' of Daphne – has the mental properties of 1, 2, and 3 and the physical properties X, Y, and Z. But, at  $t_2$ , Daphne – who still has the 'thisness' of Daphne – loses all the mental properties (1, 2, and 3) and physical properties (X, Y, and Z) that she once possessed, only to have them replaced by mental properties 4, 5, and 6 and physical properties A, B, and C. Running alongside Daphne's life, at  $t_1$ , Velma – who has the 'thisness' of Velma – has mental properties 4, 5, and 6 and

<sup>9</sup> Swinburne, *Mind, Brain & Free Will*, p. 151.

<sup>10</sup> *Ibid.*, p. 170.

physical properties A, B, and C. Then at  $t_2$ , Velma – who still has the ‘thisness’ of Velma – loses all the mental properties (4, 5, and 6) and physical properties (A, B, and C) that she once possessed only to have them replaced by mental properties 1, 2, and 3 and physical properties X, Y, and Z. So, just to clarify, we would have the following:

At  $t_1$ :

Daphne has mental properties 1, 2, and 3 and physical properties X, Y, and Z.

Velma has mental properties 4, 5, and 6 and physical properties A, B, and C.

At  $t_2$ :

Daphne has mental properties 4, 5, and 6 and physical properties A, B, and C.

Velma has mental properties 1, 2, and 3 and physical properties X, Y and Z.

Yet,

Daphne at  $t_1$  = Daphne at  $t_2$ .

And,

Velma at  $t_1$  = to Velma at  $t_2$ .

This is far more drastic than what would happen in Locke’s Prince and the Cobbler thought experiment. For on this account, persons may switch bodies (and thus physical properties), but they maintain all their mental properties (memories, dispositions, desires, etc.). And so while the Prince may be terribly confused about how he got into the Cobbler’s body (and vice versa), he is not confused about who he actually is. He remembers his life and possesses all the former mental properties that he once had. By contrast, Swinburne’s approach makes it possible to lose both the physical and the mental properties. I can imagine that both Daphne and Velma at  $t_2$  are horribly confused about who they really are. At  $t_2$ , Daphne is sure to think that she is Velma since she now has all of Velma’s old mental and physical properties, and Velma is sure to think that she is Daphne for the same reason. Both have become lost but are not even aware that it has happened. They are deceived into thinking that they are someone that they are not.



My concern here is not with Swinburne's contention that we have a 'thisness' unique to each of us. My own Thomistic leanings incline me to think that we do have an immaterial soul, even if Swinburne and I might differ over the nature of that soul. My concern is with, on his model, our complete inability to identify ourselves across time given the contingency of our mental and physical properties. If Daphne has a complete loss and replacement of her mental and physical properties between  $t_1$  and  $t_2$ , such that she now has all of Velma's old mental and physical properties, then even Daphne will not be able to properly identify herself in the future. She will think she is Velma when in fact she is not. The same is true for Velma who would now have all Daphne's old physical and mental properties and be equally confused and deceived. As I mentioned earlier, this is certainly not a logical defeater for Swinburne's position, but it does seem to be a troubling and unfortunate consequence of it.

#### CONCLUSION

On the whole, Swinburne's book is helpful and constructive. He has much to say about a number of important philosophical issues that need to be taken seriously. His critique of Thomas' view of the soul has caused me to question my own approach. For this I am grateful and I look forward to working through this difficult issue. And his thought experiment for gradual brain replacement is intriguing and enlightening. In the end, however, I find myself unconvinced by some key parts of his argument and concerned about the way his approach dismisses mental and physical properties of human beings.