

Christopher Hughes

The Soul: Some Preliminary Considerations

Descartes and various contemporary philosophers (such as Richard Swinburne) have offered modal arguments in favor of the existence of (human) souls. I set out one such argument, and explain why I think it cannot, as is sometimes thought, be dismissed simply by distinguishing conceivability from possibility, and pointing out that (for reasons brought out by Kripke) the former is no guarantee of the latter. Nevertheless, I try to show, the argument is unsuccessful, because, even if some conceivable states of affairs provide us with a defeasible reason to think that human beings have souls, other conceivable states of affairs give us an “undermining defeater” for that reason. I conclude with some speculations about why traditional arguments for human souls are so unsuccessful, even though the view that human beings have souls or something rather like souls (I argue) has considerably more intuitive appeal than is sometimes recognized.

1. Terminology

The word *soul* is often used to translate the Latin term *anima*. For a scholastic philosopher such as Aquinas, it was a more or less definitional truth that (all) (embodied) living beings have souls (see, for example *Summa Theologiae* Ia, 76, 1 *responsio*, where Aquinas says that a soul is the first principle of the life of a body).

If we mean by *soul* (more or less) what Aquinas meant by *anima*, then we will almost certainly regard it as a foregone conclusion that human beings (like all other *animate* beings) have souls.

But nowadays, most of us don't mean by *soul* (more or less) what Aquinas meant by *anima*: we won't regard it as truistic that whatever has a body with a first principle of life (be it a human being, a labrador retriever, or a cycad) has a soul.

The soul, as most of us think of it nowadays, is not unlike Aquinas' *anima intellectiva* (“intellective soul”). For Aquinas, only beings with intellects or minds can have an *anima intellectiva*; and, for Aquinas, the *anima intellectiva* cannot be identified with any material part of a living being. Likewise, as most of us use the term *soul*, only beings with intellects or minds can have souls, and the soul of a living being is not the sort of thing that could be identified with any material part thereof (such as, for example, that being's cerebral cortex). However, for Aquinas,

it is a non-definitional (non-analytic) truth that the *anima intellectiva* of a living being cannot be identified with any material part thereof. By contrast, I take it, nowadays most of us use the word *soul* in such a way that it is true *ex vi terminorum* that if there are souls, they are immaterial. Materialists do not typically say: of course I believe that human beings have souls; but I think a human being's soul, far from being an immaterial I-know-not-what, is a part of that human being's brain. Instead, they typically say something like: of course I think human beings have minds, but I don't think a human being's mind is a soul; rather I think it is a part of that human being's brain. As best I can see, materialists (typically) take this line, because they (typically) suppose that, although minds are not "by definition" immaterial, souls are. (See, for example, the definition of "soul" in the (on-line) Oxford dictionary: "the spiritual or immaterial part of a human being or animal...")

In what follows, I shall understand *soul* in the ordinary contemporary way, rather than in the scholastic way. Thus I will suppose it is not true by definition that (embodied) living beings, or human beings, have souls, and I will suppose it is true by definition that if human beings have souls, then those souls do not have a material constitution. I will not, however, understand the term *soul* in such a way that it is a definitional truth that if there are souls, they are immortal, or capable of existence apart from a human body. (In this connection, it is interesting that the Oxford dictionary defines the soul as "the spiritual or immaterial part of a human being or animal, *regarded as immortal* (my italics).) Nor shall I understand the term *soul* in such a way that it is a definitional truth that souls are essential to the beings they ensoul: I take the meaning of the term *soul* to allow that something already in existence can acquire a soul, and something can go on existing, even after losing its soul.

2. Arguing that Human Beings Have Souls

Suppose we (human beings) had material minds. Could we know that we did? There is no immediately obvious reason to suppose that we could not. After all, it seems natural to suppose, (fancy enough) minds would put the beings that have them in a position to know that they have minds. And how might we know that the minds we have are material? Well, if they are material, they might well have properties that (a) are observable, and (b) imply materiality. So, it seems, we might empirically ascertain that our minds have some materiality-implying property, and are accordingly material.

Suppose on the other hand that we (human beings) had immaterial minds. How might we come to know that we did? How might we ascertain, not just that we have minds, but that the minds we have are immaterial? If our minds are immaterial, it would seem, then they are invisible, inaudible, impalpable, and, it would seem, not empirically detectable in any straightforward way. And how might we ascertain the existence of something that is not empirically detectable in any straightforward way?

Here it might be objected that, even if our minds are immaterial, it remains true that (fancy enough) minds put the beings that have them in a position to know they have them. Fair enough. But again, the question is not how we might ascertain that we might have minds, but how we might ascertain that we have immaterial minds (aka souls). (Even if our minds are in fact souls, knowing that we have minds is not the same thing as knowing that we have souls). So we cannot resolve the epistemological quandary at issue simply by saying that, just as we ascertain the existence of material objects outside us by “extrospection” (sense-perception), we ascertain the existence of our soul by introspection .

Of course, even if our souls are not empirically detectable in a straightforward way (i.e. perceivable, perhaps with the help of special equipment (such as microscopes)) they might nevertheless be empirically detectable in some un-straightforward way. Electrons and photons are not straightforwardly detectable, but are un-straightforwardly detectable from their effects (via processes that involve detection of things that are straightforwardly detectable/perceivable and abduction). According to a venerable tradition in natural theology, God is not straightforwardly detectable, but is un-straightforwardly detectable from His effects (via a process that involves detection of things that are straightforwardly detectable/perceivable and abduction). Mightn't the same go for our souls?

It might; but it's not easy to see how. In the case of electrons, for example, it is clear (to those who know a lot more nuclear physics than I do) why we need to suppose, not just that there are sub-atomic nucleus-orbiting particles, but also that there are sub-atomic nucleus-orbiting particles with the non-straightforwardly-detectible property of negative charge: we get such-and-such straightforwardly observable effects, (only) because there are sub-atomic nucleus-orbiting particles with negative charge (we get such-and-such straightforwardly observable effects, that we wouldn't get, if there weren't sub-atomic nucleus-orbiting particles with negative charge). If the case of human souls is analogous, then we should be able to explain why we need to suppose, not just that there are human minds, but also that there are human minds with the non-straightforwardly detectible property of immateriality, by individuating straightforwardly observable effects that we get (only) because there are human minds that are immaterial (straightforwardly observable effects that we would not get, if there weren't human minds that were immaterial). It is not clear to me what such effects might be.

It is true that some people (and some philosophers) have thought there is “empirical data” that is best explained by the hypothesis that human minds are immaterial (and thus souls). For example, some people have thought that “out of body” experiences, or independently corroborated veridical near death experiences, or independently corroborated memories of past lives in other bodies, provide evidence for the immateriality of human minds. I am not competent to assess the (alleged) evidence that human beings sometimes (veridically) perceive their own body “from the outside”, or sometimes veridically perceive external objects when their brains are not functioning, or sometimes have memories of past lives in other bodies. But it may be worth pointing out that even assuming there is good evidence that, say, human beings can veridically perceive external objects

when their brain is not functioning, it is not immediately evident that this favors the hypothesis that the mind is something immaterial over the hypothesis that the mind is something made of different matter from the matter the brain is made of (perhaps of a different kind of matter from the matter the brain is made of).

In any case, not many philosophers have argued that human beings have soul by appeal to exotic (alleged) phenomena such as out of body experiences, memories of past lives, or the like. To take just one example, (at *Summa Theologiae*, Ia, 75, 2, responsio) Aquinas, drawing on certain Aristotelian ideas, argues that human beings have an immaterial intellective soul or mind roughly as follows:

Human beings have the ability to know the natures of all corporeal substances.

If the human intellective soul were material, they would not have that ability.

Human beings have an immaterial intellective soul.

Few contemporary philosophers find this argument cogent, because few think Aquinas has made out that a being with a material mind would necessarily be more limited in its ability to understand the natures of corporeal substances than human beings actually are. Contemporary philosophers have been far more engaged by modal arguments for the immateriality of the (human) mind/existence of the human soul. Arguments of this kind are broadly Cartesian, and we could formulate a representative one as follows:

- (1) There is nothing more to a human being than its mind and its body.
- (2) A human body is a material being.
- (3) If a human mind is a material being, then a human being (is composed of a material body and a material mind and accordingly) is a material being.
- (4) No material being could survive instantaneous dematerialization.
That is, no being could survive an instantaneous passage from materiality to non-materiality. In other words, it could never be true that (a) a being b exists at an (earlier) time t and at a (later) time t' and at all the times between t and t' ; (b) b is a material being at t ; (c) b is not a material being at t' , or at any time (strictly) between t and t' .
- (5) But human beings could survive instantaneous dematerialization.
- (6) So human beings are not material beings (from (4) and (5)).
- (7) So human minds are not material beings (from (3) and (6)). Human minds are immaterial, and the minds that human beings have are souls.

To forestall a possible misunderstanding: in this argument, *a could survive X* should be understood in terms of metaphysical rather than epistemic possibility. That is, *a could survive X* should be understood as meaning, not that, for all we know, a survives X , but rather as meaning that a has the (genuine) (metaphysical)

possibility of surviving X (that there is some (genuinely) possible world in which a survives X).

We must understand *a could survive X* metaphysically rather than epistemically, or else the argument just formulated will be an obviously bad argument. For that argument moves from *human beings could survive something that no material being could survive* to *no human being is identical to any material being*, and thence to *no human mind is identical to any material being* and thence to *human minds are immaterial minds/souls*. But if *could survive X* is understood epistemically, it can be true that $a = b$, even though a could survive X and b could not survive X . (Suppose I see a celestial body in a certain region of the sky in the evening, and call it “Hesperus”. I see a celestial body in a different region of the sky the next morning and call it “Phosphorus”. If “could survive X ” is understood epistemically, then Hesperus could survive the destruction of Phosphorus (i.e., for all I know, Hesperus will survive the destruction of Phosphorus). And, if “could survive X ” is understood epistemically, Hesperus could not survive the destruction of Hesperus (i.e. it’s *not* the case that, for all I know, Hesperus will survive the destruction of Hesperus (I know that Hesperus won’t survive its own destruction). Clearly, all this is compatible with the hypothesis that as a matter of fact, Hesperus = Venus = Phosphorus.

On the other hand, if *a could survive X* is understood in terms of metaphysical possibility, one can legitimately move from *a could survive X and b could not survive X* to *a is not identical to b* (and, thus, one can legitimately move from *a could survive X and nothing of kind K could survive X* to *a is not identical to anything of kind K*). Think of it this way: if it is metaphysically possible that a could survive X , and it’s not metaphysically possible that b could survive X , then a has a (genuine) possibility (to wit, the possibility of surviving X) that b does not have. In which case, a and b cannot be identical, lest one and the same thing have and not have the (genuine) possibility of surviving X (which is a contradiction).

How good is the argument just set out? Judging from my reading, my colleagues, and my graduate students, quite a few contemporary philosophers would say:

Not very. For we have no reason to suppose that (5) is true. It may be that a human being’s surviving instantaneous dematerialization is conceivable. But conceivability is one thing, and genuine (metaphysical) possibility is another. If human beings have material minds (as well as material bodies), then a human being’s surviving instantaneous dematerialization is not genuinely (metaphysically) possible (even if it is conceivable).

I think there is something right but also something wrong about this objection. To bring out what I think is wrong with the objection, it may help to think about a different modal argument—a (more or less Newtonian) argument for the distinctness of space from matter.

As I understand his view, Newton thought that space, like time, was a necessary being. (For Newton, space and time are necessary beings, because their existence is implied by the necessary fact that God exists and is omnipresent and eternal). On the other hand, matter is not a necessary being (it owes its existence to the

contingent fact that God decided to create a material world). If space is a necessary being, and matter is contingent, then matter depends on space, but not vice versa. And this is indeed what Newton appears to think.

So against philosophers (such as, perhaps, Descartes) who want to identify space with matter, Newton could have offered the following argument:

Space could have existed, even though matter did not exist.

It is (trivially) not the case that matter could have existed, even though matter did not exist.

Space is distinct from matter.

Suppose that a defender of the identity of space with matter objected to this argument as follows:

The argument is not very good. For we have no reason to suppose that its first premiss is true. It may be that space without matter is conceivable. But conceivability is one thing, and genuine (metaphysical) possibility is another. If space is matter, then space without matter is not genuinely (metaphysically) possible (even if it is conceivable).

Newton (or a hypothetical defender of Newton) might naturally enough reply that this dismissal of the argument is too quick. After all, he or she might say:

Because space without matter is conceivable, space without matter appears to be genuinely possible. To be sure, apparent genuine possibility does not imply genuine possibility (there is no contradiction in the idea of something merely seeming to be genuinely possible). Even so, the fact that space without matter is conceivable is a *pro tanto* or defeasible reason to think that space without matter is genuinely possible (just as the *inconceivability* of matter without space is a *pro tanto* or defeasible reason to think that matter without space is genuinely *impossible*). In the absence of special reasons to deny or doubt that what is conceivable and looks genuinely possible is in fact genuinely possible, it is (not just *pro tanto* but all-things-considered or on-balance) reasonable to believe that what is conceivable and looks genuinely possible is in fact genuinely possible. Moreover, we have no special reasons to deny or doubt that space without matter is genuinely possible. So it is on balance reasonable for us to accept that space without matter is genuinely possible, and thus to accept that space and matter are distinct. At least, it is on balance reasonable for us to do that, if (like Newton) we find space without matter conceivable.

This reply to our hypothetical objection seems fair enough. One cannot dismiss the modal argument for the distinctness of space from matter simply on the grounds that conceivability does not guarantee genuine possibility.

In which case, it would seem, one cannot dismiss the modal argument for human beings having souls simply on the grounds that conceivability does not guarantee genuine possibility. If we can in fact conceive of a human being's surviving instantaneous dematerialization, then, it is at least arguable, we have at least a *pro tanto* or defeasible reason to suppose that a human being really could survive

instantaneous dematerialization. In which case, if we have no special reasons to deny or doubt that a human being could survive instantaneous dematerialization, then we have an all-things-considered or on balance reason to think a human being really could survive instantaneous dematerialization. In which case, we have an all-things-considered or on balance reason to accept (5), despite the fact that conceivability is no guarantee of genuine possibility.

Granted, if we cannot conceive of a human being surviving the instantaneous loss of all its (proper) material parts, then it seems we lack a good reason to accept (5), and consequently lack a good reason to accept the cogency of the argument. In fact though, I think relatively few people (be they philosophers or not, be they materialists or not) are happy to say that they cannot so much as conceive of human beings surviving instantaneous dematerialization.

Think of it this way: suppose someone told me that today, at noon exactly, all the matter in the world would instantaneously go out of existence. Nevertheless, she said, a certain mountain – say, Pizzo d’Evigno – would go on existing (as a “dematerialized” being). My natural response would be along the lines of:

What you’re describing seems to me inconceivable and impossible.

What could make it true that something immaterial existing this afternoon was the very same thing as Pizzo d’Evigno (the mountain existing this morning)?

Suppose on the other hand that someone told me that, at noon exactly, God would instantaneously annihilate all the matter in the universe. Nevertheless, she said, I, along with all other human beings, would go on existing (in “dematerialized” or disembodied form). I would not be inclined to respond that what the person described was inconceivable and impossible. For the idea that I or another human being could go on in dematerialized form seems intelligible and conceivable in a way that the idea a mountain could go on in dematerialized form does not.

Why is this? I take it that it is because I think of myself as a “locus of experience”, and I seem to be able to conceive of myself going on existing, as the very same locus of experiences (as the “old” locus of “new experiences”), even after the annihilation of all matter. By contrast, I don’t think of mountains as loci of experience. So I am unable to conceive of Pizzo d’Evigno going on existing, as the very same locus of experience, after the annihilation of all matter. And I don’t have any other (non-experience-involving) way of making sense of the idea of Pizzo d’Evigno existing, after the annihilation of all matter. So I can’t make sense of that idea. Given the way I conceive of Pizzo d’Evigno, I cannot conceive of its surviving instantaneous dematerialization; given the way of conceive of myself or my wife, I can conceive of my or of her surviving instantaneous dematerialization.

Now assuming that (i) we are able to conceive of human beings’ surviving instantaneous dematerialization, and (ii) if we are able to conceive of that, we have a *pro tanto* or defeasible reason to think that that is possible, it follows that (iii) we have a *pro tanto* or defeasible reason to think that human beings could survive instantaneous dematerialization. But it evidently does not follow that (iv) we have good enough reason to accept (5), since it is compatible with (iii) that we also have

good reasons to deny or doubt that human beings could survive instantaneous dematerialization. (Indeed, it is compatible with (III) that we have better reasons to deny or doubt that human beings could survive instantaneous dematerialization, than we do to think that they could survive it.) To put this more concisely, if (I) and (II) are true, then we have a defeasible reason to think human beings could survive instantaneous dematerialization, but it is perfectly compatible with (I) and (II) that that reason be not only defeasible but defeated.

For example, it might be argued that we have good reason to accept all of the following claims: (a) if human beings could survive instantaneous dematerialization, they are not material beings, (b) if human beings are not material beings, then human minds are immaterial beings that causally interact with material beings, (c) if human minds are immaterial beings that causally interact with material beings, momentum is not conserved, and (d) momentum is in fact conserved. Or it might be argued that we have good reason to accept all of the following claims: (a) if human beings could survive dematerialization, then human minds are immaterial beings, (b') the minds of (all) non-human animals are material beings, and (c') the minds of at least some non-human animals (e.g. our close primate relatives) are similar enough to human minds that it cannot plausibly be maintained that the former are material, and the latter immaterial.

For reasons of space, I shall not explore either of these lines of argument. I shall instead explore a different strategy for showing that, despite the apparent conceivability of human beings' surviving instantaneous dematerialization, we do not have good enough reason to accept that that is genuinely possible, because we have good reasons to doubt (as opposed to deny) that it is genuinely possible.

John Locke believed that human minds were immaterial. But, he says, we don't know that all minds are immaterial:

We have the ideas of matter and thinking, but possibly shall never be able to know, whether any mere material being thinks, or no; it being impossible for us, by the contemplation of our own ideas, without revelation, to discover, whether omnipotency has not given to some systems of matter fitly disposed, a power to perceive and think, or else joined and fixed to matter so disposed a thinking immaterial substance: It being, in respect of our notions, not much more remote from our comprehension to conceive, that God can, if he pleases, superadd to matter a faculty of thinking, than that he should superadd to it another substance with a faculty of thinking; since we know not wherein thinking consists, nor to what sort of substances the Almighty has been pleased to give that power, which cannot be in any created being, but merely by the good pleasure and bounty of the Creator. For I see no contradiction in it, that the first eternal thinking Being or omnipotent Spirit should, if he pleased, give to certain systems of created senseless matter, put together as he thinks fit, some degrees of sense, perception, and thought.... What certainty of knowledge can any one have that some perceptions, such as, v.g.. pleasure and pain, should not be in some bodies themselves, after a certain manner modified and moved, as well as that they should be in an immaterial substance, upon the motion of the parts of body? ... I say not this, that I would any way lessen the belief of the soul's immateriality: I am not here speaking of probability, but knowledge; and I think not only, that it becomes the modesty of philosophy not to pronounce magisterially, where we want that evidence that can produce knowledge; but also, that it is of use to us

to discern how far our knowledge does reach; for the state we are at present in, not being that of vision, we must, in many things, content ourselves with faith and probability; and in the present question, about the immateriality of the soul, if our faculties cannot arrive at demonstrative certainty, we need not think it strange. (John Locke, *An Essay Concerning Human Understanding*, IV.iii.6).

As I understand Locke, he is saying here that there is no manifest incompatibility between the property of being a material being, and the property of having a mind (of having sense, perception, thought). Initially, at least, this seems right. The property of thinking does not manifestly imply the property of being a material thing (otherwise the idea of God or an angel would be manifestly incoherent), but neither, it seems, is the property of thinking manifestly incompatible with the property of being a material thing: the property of thinking does not seem in any obvious way to exclude the property of being composed of atoms, in the way that, for example, the property of being a number, or the property of being a hole, manifestly excludes the property of being composed of atoms.

If this is so, then it looks as though we can conceive of material beings with minds. In which case, it seems, we can also conceive that we (human beings) are material beings with minds. At any rate, as best I can see, I am just as able to conceive that I am a being both human and material, as I am to conceive of my surviving instantaneous dematerialization.

So it looks as though an opponent of the modal argument for human souls can challenge premiss (5) along the following lines:

Granted, we can conceive of a human being's surviving instantaneous dematerialization. Granted, this gives us a defeasible reason to suppose that (5) is true. But we can also conceive that human beings are material beings. And this gives us a defeasible reason to suppose that it is genuinely possible that human beings are material beings. Moreover, as per (4), a material being could not survive instantaneous dematerialization. Whence we may conclude that either it is conceivable but not genuinely possible that human beings survive instantaneous dematerialization, or it is conceivable but not genuinely possible that human beings are material beings. Moreover, we have no more reason to suppose that *human beings are material beings* is conceivable but not genuinely possible than we do to suppose that *human beings survive instantaneous dematerialization* is conceivable but not genuinely possible. In these circumstances, although we have a *pro tanto* or defeasible reason to accept (5), we do not have an all-things-considered or on-balance reason to accept it; the reasonable thing for us to do is to withhold assent from (5), even if (a) it is conceivable that human beings survive instantaneous dematerialization, and (b) we have no (all-things-considered, on-balance) reason to suppose that it is *not* genuinely possible for human beings to survive instantaneous dematerialization. Since we should withhold assent from (5), we should not find the argument cogent.

How might a defender of the modal argument for human souls respond to this challenge to (5)?

The challenge only goes through if it is in fact impossible for a material being to survive instantaneous dematerialization. Some philosophers (e.g. Sydney Shoemaker) would deny this assumption. But it would obviously be self-defeating

for the defender of the modal argument for human souls to take issue with that assumption, since it's a premiss of that argument (premiss (4)).

As far as I can tell, this leaves the defender with the following options:

(O₁) She could challenge the assumption that whenever we can conceive that *p*, we have a *pro tanto* or defeasible reason to think that *p* is genuinely possible. (Absent that assumption, we no longer have a route from the claim that we can conceive that human beings are material beings to the claim that we have a defeasible reason to think that *human beings are material beings* is genuinely possible, and thus (given (4)) a reason to doubt that *human beings survive instantaneous dematerialization* is genuinely possible.)

(O₂) She could challenge the assumption that we can conceive that human beings are material beings. (Absent that assumption, we again no longer have a route to the claim that *human beings are material beings* is genuinely possible; so we no longer have a reason to think that *human beings survive instantaneous dematerialization* might not be genuinely possible, even if it is conceivable).

(O₃) She could challenge the assumption that we have no more reason to think it is conceivable-but-not genuinely-possible for a human being to be a material being than we do to suppose that it is conceivable-but-not genuinely possible for a human being to survive instantaneous dematerialization. (If we have more reason to regard *human beings are material beings* as conceivable-but-not-genuinely-possible than we do to regard *human beings survive instantaneous dematerialization* as conceivable-but-not-genuinely possible, then our "total evidence" favors (5), and it's not evident that we should withhold our assent from it.)

If the defender of the modal argument for human souls plumps for (O₁), she can certainly challenge the assumption that we have (even a defeasible) reason to think that *human beings are material beings* is genuinely possible. But then she has to face the question of why we have (even a defeasible) reason to think that *human beings survive instantaneous dematerialization* is genuinely possible. It was suggested earlier that we have a (defeasible) reason to think that *human beings survive instantaneous dematerialization* is genuinely possible, because we can conceive of human beings surviving instantaneous dematerialization. But if the defender of the modal argument for human souls plumps for (O₁), she cannot give this explanation, and it's unclear to me what other one she might give.

I have already indicated some of the reasons why I regard (O₂) as a problematic option: like Locke, I am inclined to think we can conceive that human beings are material beings. But suppose that we can motivate the idea that we cannot so much as conceive that human beings are material beings. In that case, it is unclear why we need the complications of the modal argument for human souls. We could more straightforwardly argue that since it is inconceivable and impossible that that human beings are material beings, and it is incontestable that human bodies are material beings, and there is nothing more to a human being than a human body and a human mind, human minds are not material beings. So we can get to premiss (6) without appeal to the (alleged) possibility of human beings' surviving instantaneous dematerialization.

Analogously, I don't know of any reason to think we have more reason to regard *human beings are material beings* as conceivable-but-not-genuinely-possible, than we do to regard *human beings survive instantaneous dematerialization* as conceivable-but-not-genuinely possible. In any case, suppose the defender of the modal argument for human souls says that we can reasonably accept (5), (only) because we have reasons to think that of the pair {*human beings survive instantaneous dematerialization, human beings are material beings*} it is the second, rather than the first that is conceivable-but-impossible. Then the defender of the argument is in effect conceding that the modal argument for human souls is dramatically enthymematic, since it says nothing about why *human beings are material beings*, rather than *human beings can survive instantaneous dematerialization*, is conceivable-but-impossible. If we supplemented the modal argument for human souls with an explicit argument in favor of the claim that of the pair {*human beings survive instantaneous dematerialization, human beings are material beings*} it is the second, rather than the first that is conceivable-but-impossible, it seems that the supplement (rather than (1) – (7)) would be the core of the argument for human souls.

Summing up: given the conceivability (or at very least, the apparent conceivability) of a human being's surviving instantaneous dematerialization, I don't think we can dismiss the modal argument for human souls on the grounds that its crucial premiss (5) can be seen to be unmotivated, once we take account of the distinction between conceivability and genuine possibility. But I do think we can challenge that argument on the grounds that (5) is insufficiently motivated, once we take into account the distinction between conceivability and possibility, and take account of the conceivability (or at least the apparent conceivability) of things that are possible only if (5) is false. So, although we appear to have some reason to think (5) is true, it is by no means clear that we have good enough reasons to think it is true to make it reasonable for us to accept it. I don't know what such reasons might be, but perhaps they could be provided. However that may be, they are not provided in the argument under consideration, so that argument is at very least (seriously) deficient because (seriously) under-explicit. If such reasons were provided, they might vindicate the argument, but they might equally supplant it, in the sense of rendering it otiose.

Setting these considerations to one side, there is another worry about our argument. The argument depends on the idea that, since human beings are capable of surviving an instantaneous passage from materiality to non-materiality, but material beings are not, human beings are not material beings.

Now we might understand materiality in either of two ways. We might say that something is material if and only if all its (proper) parts are material. Or we might say that something is material if at least some of its (proper) parts are material. If we construe materiality in the first (and stronger) way, it seems plausible enough that a material being could not survive the instantaneous passage from materiality to non-materiality. For, on the first understanding of materiality, a material being could survive the instantaneous passage from materiality to non-materiality only if it could survive the instantaneous loss of all its (current) (proper) parts, and it

is hard to see how a being can lose all its (current) (proper) parts *at once*, without thereby going out of existence (even if a being can gradually lose all its (current) (proper) parts, without thereby going out of existence (as long as the parts are replaced in a way that preserves the (current) “structure” or “(Aristotelian) form” of that being). On the other hand, if we understand materiality in the second (and weaker) way, there is nothing obvious about the idea that a material being could not survive the instantaneous passage from materiality to non-materiality. Why couldn't a being that currently has some material parts and some immaterial parts instantaneously become a being with no material parts, as long as it held onto all its immaterial parts?

So the modal argument for human souls needs the stronger construal of materiality, according to which being material = being wholly material = having only material parts. Now on that construal, there will be a difference between non-materiality and immateriality. A being will be non-material if and only if it is not material, i.e. if and only if it is not wholly material, if and only if not all of its parts are material. A being will be immaterial if and only if none of its parts are material. Thus immateriality will be a stronger property than non-materiality (since *being non-material* will be equivalent to *being immaterial* or *being partly but not wholly material*).

If, however, immateriality is stronger than non-materiality, then (7) is open to question. For (7) moves from *human minds are not material* (which follows from (3) and (6)) to *human minds are immaterial* (which does not).

Here is another way to put essentially the same point. (On the contemporary understanding of the term) a human soul is an immaterial human mind. A mind that was partly material and partly immaterial would be at best a “quasi-soul” (something that resembles a soul more than a material mind does). And it seems that the most that the modal argument could show is that human beings either have souls or quasi-souls. In order for the argument to arrive at the conclusion that human beings have souls, we would need some extra premisses. If, for example, we supposed that no beings are partly material and partly immaterial (or, more weakly, that no human beings are partly material and partly immaterial), we could move from *human beings have souls or quasi-souls* to *human beings have souls*. Or, if we supposed that any being that has material parts has at least one essential material part, we could get from *human beings can survive instantaneous dematerialization* (and the other premisses of the modal argument) to *human beings have souls*. But neither these assumptions, nor any other assumptions I can think of that would allow us to eliminate the possibility that human minds are quasi-souls, seem obvious. So again, the modal argument for human souls is seriously enthymematic.

Suppose we call the scholastic argument for human souls sketched earlier *the epistemic argument for human souls* (since it depends on the idea that human beings have epistemic powers they could not have, unless their minds were souls). To my mind, both the epistemic and the modal argument for human souls are unsuccessful, in a strong sense. It is not just that those arguments don't demonstrate or establish that human beings have souls. It's that it is hard to see why anyone who had his or her wits about her, and was exposed to the epistemic or the modal argument,

would be more inclined to believe in human souls, than he or she was before exposure to it. (For a contrast case, consider Kripke's arguments for the falsity of descriptivist accounts of the meaning of proper names. Quite a few people with their wits about them who had no initial inclination to regard those descriptivist accounts as false, acquired a strong inclination to regard them as false, as a result of exposure to Kripke's anti-descriptivist arguments).

Why are two of the most celebrated arguments for human souls so unsuccessful? One possible explanation would be that the idea that humans have immaterial (or at least non-material) minds has zero intuitive appeal, and is accordingly very hard to argue for. But I don't think that this is how things are.

I'm inclined to think that the view that the epistemic and modal arguments are unsuccessful because, even though the idea that human minds are not material has intuitive appeal, those arguments don't directly or successfully engage the intuitions that favor the idea that human minds are not material.

Here are some things that we never say, and it would be distinctly odd to say:

His mind is half a degree hotter than it was three hours ago.

Her mind is shrinking at the rate of 1/1000-th a centimeter per second.

The bullet is lodged smack in the middle of his mind.

If human minds are material beings, then (in the right circumstances), any of these statements could be (strictly and literally) true. Yet we don't make such statements. (If, in the course of a conversation, someone said to me, "stop: my mind is overheating!", I would never consider the possibility that my interlocutor was using the term "overheating" literally.) Why don't we ever say things are sometimes (strictly and literally true), if human minds are material beings?

Perhaps it's because – whatever our "official" views on the mind-body problem – at a deep level, we believe, or at least are inclined to believe, that (unlike a cerebral cortex) a human mind is not the sort of thing that can have properties such as growing hotter, or shrinking, or having a bullet lodged in it, because a human mind is not a material being at all. In support of this idea, it might be said that if someone says to us, "my glass is full of H₂O", it sounds a bit odd (in the sense of "unusual") – arch, or pedantic, perhaps – but it doesn't sound *wrong* to us; by contrast, if someone says to us "there's a fragment of metal lodged in my mind" it does sound wrong to us (at least, it sounds wrong to me). Nobody has any inclination to regard "my glass is full of H₂O" as what in the 1950s philosophers called "a category mistake"; but we do have some inclination to regard "there's a fragment of metal lodged in my mind" as a category mistake (at least, I do).

(Compare: we never say things like: "a human body is playing chess", "a human body has fallen in love". Perhaps that's because – whatever our "official" views on the identity or otherwise of human beings with their bodies – at a deep level, we believe, or at least are inclined to believe, that a human body is not the sort of thing that can play chess or fall in love.)

If this is right, then the view that human beings have souls – or at least, quasi-souls – has considerable intuitive appeal. But neither the epistemic argument nor the modal argument as it were harness that intuitive appeal.

If this is how things are, why do some of the most celebrated arguments for human souls fail to directly engage our intuitions in favor of the existence of souls (or quasi-souls)? Why don't believers in souls such as Descartes or Aquinas leave epistemic and modal considerations out of it, and appeal directly to our intuitions about what sort of properties minds do and do not have?

Perhaps it is because arguments of that sort would be dialectically unsatisfactory. An analogy may help bring out what I mean.

Willard van Orman Quine believed that numbers were beings that lacked mass and spatial location, and were accordingly non-material (and immaterial) beings. But, so far as I know, nowhere does he provide the following argument for the non-materiality of numbers:

Numbers do not have mass.

Numbers do not have spatial location.

If numbers were material beings, they would have mass and spatial location.

Numbers are not material beings.

It is not surprising that Quine does not provide the argument just set out. For it's not clear what the dialectical point of the argument would be. After all, the argument wouldn't induce anyone to acquire the belief that numbers are not material beings, because only someone who already accepted the non-materiality of numbers would accept either the first or the premiss.

Analogously, someone could argue for the non-materiality of human minds as follows:

Human minds don't (ever) grow hotter, or shrink, or get bullets lodged in them.

If human minds were material beings, they would (sometimes) grow hotter, or shrink, or get bullets lodged in them.

Human minds are not material beings.

But again, it's not clear what the dialectical point of this argument would be. It wouldn't induce anyone to acquire the belief that human minds are non-material, because only someone who already accepted the non-materiality of human minds would accept the argument's first premiss.

Philosophers arguing for the non-materiality of human minds (including the proponents of the epistemic and modal arguments) have typically aimed to put together premisses, each of which would and should secure the assent even of those

who do not already believe that human minds were non-material, in such a way that the premisses jointly implied the non-materiality of the human mind. So they have constructed arguments that depend on premisses such as *it is at least possible for human beings to survive instantaneous dematerialization*, and *no material being could survive instantaneous dematerialization* (each of which could, and indeed has, secured the assent of some who believe that human minds are material), rather than arguments that depend on premisses such as *a human mind is the not the sort of thing that can grow hotter, or shrink, or have a bullet lodged in it* (which would only secure the assent of those who already believe that human minds are non-material). Arguments depending on premisses like *a human mind is not the sort of thing that can grow hotter, or shrink, or have a bullet lodged in it* could and would be dismissed by those who do not accept the non-materiality of human minds on the grounds that they “beg the question”. (“Question-begging-ness” is simply the complement of dialectical usefulness.)

I have already said that I don’t think the epistemic argument for human souls is a non-question-begging and compelling argument for the non-materiality (much less the immateriality) of human minds. And I have already said why I don’t think that the modal argument for human souls is a non-question-begging and compelling argument for the non-materiality (much less the immateriality) of human minds. In fact, I don’t know of any non-question-begging and compelling (or even fairly compelling) arguments for the non-materiality (much less the immateriality) of human minds – despite the (to my mind, often overlooked) fact that, on the face of it, the non-materiality of human minds is a thesis not without intuitive appeal. (In some of the contemporary literature in the philosophy of mind, one comes across something like the following view:

the idea that mental properties are not physical properties is problematic, because it is not clear that we can reconcile it with a defensible account of the causal role of mental properties. Be that as it may, that idea has a great deal of intuitive appeal (for reasons brought to light by Kripke). On the other hand, the idea that human minds are not material beings is just a mistake – a mistake that (given the decline of religion, and progress in the sciences, and in particular neuroscience) people are less and less inclined to make.

For reasons I have tried to bring out, I think this view underestimates the intuitive pull of the idea that human beings have souls or at least something like souls (quasi-souls). As Sydney Shoemaker remarks, even [scientifically literate and non-religious] contemporary materialist philosophers often harbor (recognized or unrecognized) inclinations to embrace mind-body dualism. And non-philosophers often endorse the view that human minds are material (because they see it as part of “the scientific world view”), whilst continuing at some (possibly un-articulated) level to think – or at least to “kind of think” – that human minds lack properties that (they would agree) human minds would have, if they were material.)

Christopher Hughes
King’s College London
christopher.hughes@kcl.ac.uk

Christopher Hughes ha svolto i suoi studi presso la Wesleyan University e la University of Pittsburgh, e insegna presso la King's College London e il London Programme della Notre Dame University. Ha interessi in metafisica, logica filosofica, filosofia medioevale, e filosofia della religione. I suoi libri includono *Kripke: Names, Necessity, and Identity* (Oxford, 2004); *Filosofia della Religione: La Prospettiva Analitica* (Laterza, 2005), e il prossimo anno verrà pubblicato il suo libro sulla metafisica di Tommaso d'Aquino.