

A dilemma for the soul theory of personal identity

Jacob Berger¹

Received: 22 August 2016 / Accepted: 14 November 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The problem of diachronic personal identity is this: what explains why a person P_1 at time T_1 is numerically identical with a person P_2 at a later time T_2 , even if they are not at those times qualitatively identical? One traditional explanation is the soul theory, according to which persons persist in virtue of their nonphysical souls. I argue here that this view faces a new and arguably insuperable dilemma: either (a) souls, like physical bodies, change over time, in which case the soul theory faces an analogue problem of diachronic soul identity, or (b) souls, unlike physical bodies, do not change over time, in which case the soul theory cannot explain why souls relate to particular bodies over time and so at best only partially explains personal identity. I conclude that the soul theory fares no better than physicalist-friendly accounts of personal identity such as bodily- or psychological-continuity-based views.

Keywords Personal identity · Souls · Persistence over time · Bodily and psychological continuity

‘Never tell a child’, said George Macdonald, ‘you have a soul. Teach him, you are a soul; you have a body’ (Alexander 1892, p. 157).

Introduction

I raise here a new difficulty for what I will call the ‘soul theory’ of personal identity. Though there are many questions regarding personal identity (for a review, see, e.g., Kind 2015; Olson 2015), my focus is on the *problem of diachronic personal*

✉ Jacob Berger
jfberger@gmail.com

¹ Department of English and Philosophy, Idaho State University, 921 South 8th Avenue, Stop 8056, Pocatello, ID 83209-8056, USA

identity: what explains why a person P_1 at time T_1 is numerically identical with a person P_2 at a later time T_2 , even if P_1 at T_1 and P_2 at T_2 are not qualitatively identical? What makes an adult in her thirties the same person as herself in her twenties, despite being physically and psychologically different?¹ I assume that people do exist, persist over time, and that there could be an answer to this question (for dissent, see, e.g., Hume 1739/2000; Unger 1979). Because *souls* play a prominent role in many religious traditions from Christianity to Hinduism, one of the classical responses to this question (often put forth by religiously minded philosophers and theologians) is that the persistence of one's nonphysical soul explains personal identity. As Richard Swinburne puts it, in addition to the physical matter of our bodies:

We may say that there is stuff of another kind, immaterial stuff, and that persons are made of both normal bodily matter and of this immaterial stuff but that it is the continuity of the latter which provides that continuity of stuff which is necessary for the identity of the person over time (1984, p. 27).

In other words:

Soul Theory: Person P_1 at time T_1 is numerically identical with a person P_2 at a later time T_2 iff there is a chain of overlapping soul-continuity linking P_1 and P_2 —that is, P_1 and P_2 have the same soul.²

According to the soul theory, the reason why a thirty-year-old adult is the same person as a particular twenty-year-old adult is that both have the same nonphysical soul.

Why care about the soul theory? Haven't most nonreligious metaphysicians abandoned it in favor of other, typically physicalist-friendly, accounts of personal identity such as *bodily-* or *psychological-continuity-based* theories? But while the soul theory may be currently out of fashion, it boasts prominent historical and contemporary adherents (see, e.g., Butler 1736/2005; Madell 1981), perhaps most notably Thomas Reid (1785/1969) and Swinburne (1984, 1997, 2013, 2014). Moreover, as Eric Olson (2007, chapter 7) observes in his recent treatment of soul-based accounts of *synchronic* identity,³ metaphysical theories that advert to souls

¹ Going forward, uses of 'same', 'identical', and related expressions should be understood in the *numerical* sense and references to the problem of personal identity refer to the problem of *diachronic* identity, unless otherwise specified.

² When I say there is a 'chain of overlapping X-continuity linking P_1 and P_2 ', I mean that there are various person stages between P_1 and P_2 and at each junction between stages the adjacent stages can be said to have the same relevant feature X (in this case, the soul) (see, e.g., Parfit 1984, p. 206).

³ Introductory texts perhaps more frequently mention souls as a solution to the problem of synchronic identity—the question of what makes a person *that person* at a particular time (see, e.g., Rachels and Rachels 2012, pp. 53–55; Olson 2015, Sect. 1). And as Olson (2007, p. 150) observes, such views come in two basic forms: views on which we are simply souls (which Olson (2007, p. 151) calls 'immaterialism') and views on which we are complexes in some way of souls and material bodies (the view defended by Swinburne (1984) and suggested by his quote above). But whatever account we give of synchronic identity, the soul theory as I construe it claims that the soul or soul component of persons explains diachronic identity.

are too often casually dismissed by contemporary philosophers,⁴ though they deserve to be taken seriously. Most physicalist accounts of personal identity face well-known objections (for review, see, e.g., Kind 2015; Olson 2015)—and so one might wonder whether the soul theory might once more seem attractive. It is in that spirit (pun intended) that I aim to grapple seriously with the soul theory.

Although there are many familiar objections to nonphysical souls/minds in general (for review, see, e.g., Mandik 2013) and the soul theory of personal identity in particular (see, e.g., Locke 1690/1975; Quinton 1962; Parfit 1984; Shoemaker 2009), I will not dispute that souls exist. Rather, I argue that, even if we have souls, the soul theory faces an arguably insuperable dilemma. The dilemma is this: either (a) souls, like physical bodies, change over time, in which case the soul theory faces an analogue problem of diachronic *soul* identity, or (b) souls, unlike physical bodies, do not change over time, in which case the soul theory faces a related problem insofar as it cannot explain why souls inhere in particular bodies⁵—and so the soul theory at best only partially explains personal identity. The soul theory is plainly not *reductive* insofar as it does not seek to explain personal identity in purely physical terms, but one might think that this is a small price to pay for an account of personal identity. But I argue that the soul theory faces serious difficulties and is thus no improvement over reductive views.

Before unpacking the dilemma, I should point out that I assume no particular conception of a soul, except that it is a nonphysical entity related to a physical body (in life) that purports to explain why a person persists over time. I assume no specific religious conception of a soul—for example, whether or not it is given to us by God—nor any thicker metaphysical conception of the soul—for example, whether the soul supervenes with metaphysical necessity on the physical or can exist without it. I intend my arguments to be maximally general, applying to any account of the soul. Where the nature of the soul is relevant to the dilemma, I flag it. With that said, let us turn to the first horn.

The first horn: a changing soul

The soul theory must maintain either that souls change over time or that they do not. What if it holds the former? I argue that it faces a problem of soul identity, analogous to the problem of personal identity: what explains why a changing soul remains the same soul over time? As Trenton Merricks has observed, “None of the prominent dualists (for instance, Swinburne or [Roderick] Chisholm) offers anything like a criterion of identity over time for... souls” (1998, p. 121, fn. 1).

⁴ For example, consider this remark: “Souls might seem to provide quick answers to many philosophical perplexities about identity over time, but there is no good reason to believe that they exist” (Conee and Sider 2005, pp. 10–11).

⁵ I often write as though a soul *inheres in* or *occupies* a body, or that a body *possesses* a soul, but these expressions are of course a kind of loose talk. Since souls are nonphysical, they do not have spatiotemporal locations, and so cannot be located in a body or anywhere else. What I mean, strictly speaking, is that a soul is paired with or related to that body.

The problem for the changing-soul theory is that no criterion of identity can be offered, or so I will argue.

I should first be clear about the way(s) in which souls might change over time. Since souls are often identified with the *dualist* conception of minds of the sort defended by Descartes (1641/1984), a natural suggestion would be to think of souls as nonphysical particulars (substances), which are the bearers of mental properties that change over time. This seems to be the view defended by the English bishop and philosopher Joseph Butler (1736/2005); it is more recently defended by Swinburne, who writes:

Souls are immaterial subjects of mental properties. They have sensations and thoughts, desires and beliefs, and perform intentional actions. Souls are the essential parts of human beings, and humans have sensations etc. and perform intentional actions in virtue of their souls doing so (1997, p. 333).⁶

According to this view, if I have the thought that Philadelphia is in Pennsylvania, it is because my soul takes on the mental property of having the thought that Philadelphia is in Pennsylvania. Moreover, on this view, a soul S_1 at time T_1 need not have the same mental properties at a later time T_2 . If souls explain why human beings have thoughts, then the soul of a person who thinks through an inference will have a series of numerically distinct thoughts.

But if souls can have distinct sets of mental properties over time, then the soul theory is arguably vulnerable to objections analogous to various standard problems for physicalist accounts of personal identity. For example, a classic objection to a simple bodily-continuity view—according to which a person P_1 at time T_1 is numerically identical with a person P_2 at time T_2 just in case there is a chain of overlapping bodily continuity linking P_1 and P_2 —depends on Locke's (1690/1975, II.xxvii) famous thought experiment involving the *prince and cobbler*. Locke imagines that the minds of a prince and cobbler might radically change (perhaps as the result of magic, drugs, or neurological intervention), with the result that the prince's body behaves like the cobbler and vice versa. A common intuition is that the prince now inhabits the cobbler's body. And Locke concludes that such scenarios illustrate that the body is not the proper criterion for personal identity, lest we regard the cobbler as simply having become insane.

A soul theory on which souls have changing psychological features faces a similar problem. Could a soul S_1 bearing a set of mental states M_1 and a soul S_2 bearing a set of mental states M_2 at time T_1 radically change mental properties? To the extent that we can conceive of things quickly and profoundly altering the mental states of a person, it seems reasonable to assume that something (e.g., God or magic) could immediately and unexpectedly alter the mental states of souls in this way. In that case, would S_1 continue to exist, though at T_2 it has the psychological profile of S_2 at T_1 , or would S_1 *become* S_2 ?

⁶ On Swinburne's subtle view, souls are what he calls 'pure mental substances' that possess only 'pure mental properties' (see, e.g., 2014, p. 149; cf. 2013, p. 173)—that is, the existence of these substances and properties do not entail the existence of any physical substances or properties. Mental substances are necessary for, and the subject of, thoughts and experiences.

In order to answer this question, soul theorists must decide whether or not souls are individuated by their mental properties. Swinburne is clear that they are not: he maintains that, though souls can take on or lose mental properties, the underlying soul substances are themselves *indivisible* and *unchanging*; each soul is individuated simply by its own *haecceity* or *thisness*.⁷ This so-called *simple view* of the soul has a long history (see Zimmerman 1991); as Swinburne (2013, p. 173) notes, his account echoes Plato's (e.g., 1997, pp. 77b–80c) view of the soul in the *Phaedo*, on which souls are immutable and metaphysically simple insofar as they have no (nonphysical) parts. And it would seem that this view would hold that S_1 remains S_1 , despite taking on psychological features of S_2 , because its unchanging substance persists. Just as the same lump of clay might be the substance that persists throughout the process of being fashioned into different sculptures, a soul substance might persist despite at times taking on or losing different mental characteristics. But endorsing this kind of account amounts to embracing the second horn of the dilemma, to which I return in the “[The second horn: an unchanging soul](#)” section—arguing that it is problematic for other reasons.

Although this conception of souls is often assumed, it is not clear there are good reasons for it. After all, lumps of clay do not persist through all changes: exposing clay to a powerful acid may separate it into its composite substances or destroy it altogether. And some theists have urged that nonphysical entities can change. Peter Forrest (1998), for example, proposes that the Trinity arose as a result of the dividing of the primordial unitary God. Thus it would seem open that, while soul substances may persist throughout the process of taking on or losing mental properties, those underlying substances can themselves change too.

But if souls can change, and if they are not individuated by their psychological properties, then it is unclear how to answer the question of whether or not souls persist in the kinds of cases described above. It might seem that soul theorists can simply maintain that, since souls are not individuated psychologically, S_1 does persist. Such a view, of course, delivers what many regard as the wrong result in Locke's case.⁸ But even if soul theorists were willing to accept this consequence, the view faces another explanatory challenge: *why* does the underlying soul substance persist? It would seem equally open to the soul theorists to claim that the substance of S_1 persists, changes into the substance of S_2 , or even ceases to exist—and without an additional criterion of soul identity, it is unclear how to settle the issue.

A soul theorist might maintain instead that souls *are* individuated by their mental properties—that is, at T_2 S_1 becomes S_2 . Just as a sculpture may come into or go out

⁷ Swinburne writes that “my soul has its own thisness, independently of any thisness possessed by any brain to which it is connected. For clearly my soul could have had a different mental life from the one it had.... Hence the mental substance is not the substance that is in virtue merely of the properties which it has” (2014, p. 151).

⁸ I acknowledge that some have argued that thought experiments such as Locke's are question begging (e.g., Williams 1970). After all, why not think that the cobbler has become insane? My point, however, is not that these scenarios establish a particular criterion as being relevant to personal identity (e.g., psychological-, rather than bodily-, continuity). My point is that Locke's thought experiment reveals that there is a question about the relevant criterion—and so we need a theory to answer whether or not the cobbler and prince swap bodies. But the soul theory cannot explain what happens in such scenarios.

of existence depending on how its clay is arranged, so too might souls cease to exist or change into other souls depending on which psychological properties they gain or lose. But this position faces two related difficulties. First, the view seemingly amounts to a simple version of a psychological-continuity-based view, of the sort defended by Locke himself. On this kind of view, a person P_1 at time T_1 is identical with a person P_2 at time T_2 just in case there is a chain of overlapping psychological continuity linking P_1 and P_2 .⁹ A soul theory that individuates souls psychologically makes the same predictions as a psychological-continuity-based view and so encounters the same difficulties.

This kind of soul theory faces, for example, a variant of the oft-discussed *fission* problem (see, e.g., Parfit 1971; Lewis 1976). This problem imagines that P_1 at T_1 might at T_2 split (magically, as a result of surgery, etc.) into two psychologically equivalent individuals P_3 and P_4 . The question is: with whom is P_1 identical? P_3 , P_4 , neither, or both? An answer of ‘both’ seems ruled out, as one cannot be numerically identical with two distinct individuals. And there does not seem to be any reason to privilege P_3 or P_4 . So it might seem that P_1 is identical with neither, but this is problematic for the psychological-continuity view, as P_1 is psychologically continuous with both.

An analogue problem undermines this version of the changing-soul theory. Could a soul S_1 at T_1 “fissure” into two psychologically equivalent souls S_3 and S_4 —and if so, which resultant soul is identical with S_1 ? It would seem that a soul theory cannot offer a principled reason to answer one way or the other—and it cannot, as I argue in the next section, comfortably avail itself of the kinds of replies that non-soul theorists have offered.

The central problem raised by this horn of the dilemma is this: *changing-soul theorists cannot offer an account of diachronic soul identity and thus souls cannot explain diachronic personal identity*. Since the soul theory cannot explain whether and how a soul persists in cases of soul fission or the radical altering of its properties, the soul theory cannot explain whether and how a person persists in such cases. Does the prince come to inhabit the cobbler’s body, or does the cobbler become insane (or die)? It would seem that the changing-soul theorist cannot say.

Some potential replies

Merricks (1998) does not regard the fact that a theory of personal identity cannot offer an informative criterion of identity as a problem because he denies the assumption that the persistence of persons requires such criteria (see also, e.g., Reid 1785/1969, essay 3, chapter 4; Madell 2014). Similarly, on the basis of fission-type cases and related considerations, theorists such as Parfit (1984) have argued that strict identity is not what matters relative to the question of whether or not we *survive* in the manner that we care about over time. On this kind of view, while P_1 in a fission scenario is not *identical* with either P_3 or P_4 , P_1 stands in an appropriate

⁹ Locke himself casts the relevant continuity in terms of *memory*, but it can be understood in broader psychological terms (cf. Parfit 1984, p. 207).

relation (Parfit claims a relation of sufficient psychological continuity) to P_3 and P_4 such that P_1 survives in *both*. Alternatively, some physicalists urge that we endorse a *multiple-occupant view*, on which we identify persons with *space–time worms* composed of numerically distinct temporal parts (e.g., Sider 2001; Olson 2007). On this view, there may be multiple numerically distinct individuals occupying the same body; fission cases simply reveal where these space–time worms diverge. There are two people in P_1 's body at T_1 , one of which is identical with P_3 and the other with P_4 . There are many related kinds of solutions (for review, see Kind 2015), but I cannot explore them here.

But even at this schematic level, it is clear why, even if such positions were attractive to physicalists, related solutions are unappealing for soul theorists. First, souls do not seem like the kinds of things that merely survive over time albeit in multiple forms. Rather, souls were posited to give *strict criteria* for identity over time. Though one could grant that souls merely survive over time, one might equally well adopt the metaphysically simpler physicalist view that material persons merely survive over time. And while soul theorists could follow Merricks in holding that there need not be criteria of soul identity, one might as well assume a physicalist view on which persons are wholly physical and persist, but not because of any particular criteria. Likewise, while one could grant that the same material body might harbor numerous or even innumerable souls that would inhabit distinct bodies were that body to fissure, if one is willing to adopt a multiple-occupant view, then one should adopt a physicalist multiple-occupant view. That is to say, while the problem raised above need not decisively undermine the soul theory, adopting these alternative strategies renders it quite unattractive and undermines its major motivation (in this context).¹⁰

A soul theorist might object, however, that outlandish scenarios such as cases of soul fission may be conceivable but are not genuinely possible.¹¹ Soul theorists might protest that souls are simply not the kinds of things that can fissure or that can exchange psychological properties. Perhaps, soul theorists might urge, such cases would involve unacceptable violations of *Natural Law*.¹² Thus if P_1 and P_2 were to acquire each others' mental properties, they would either have gone insane or died—and the resultant bodies would either be soulless (akin to philosophical zombies; see, e.g., Chalmers 1996) or would acquire new souls. That is, there is no puzzle for the soul theory because such scenarios do not involve continuity of soul or personhood in new bodies.

¹⁰ I note that some might nonetheless prefer a soul theory—such as a soul-based multiple-occupant view—for other reasons. For example, they might prefer such a view because only a soul theory can secure life after bodily death. And while that may be a reasonable consideration, notice that such an argument supports a soul theory not on the grounds that it explains ordinary (living) personal identity, but rather that it explains how persons might exist separately from their bodies. But I do not deny that there may be additional reasons to posit nonphysical souls—my point is that souls are not needed to explain *personal identity*.

¹¹ There is of course debate regarding whether or not conceivability is a good guide to possibility (see, e.g., the essays in Gendler and Hawthorne 2002). For reasons to doubt that prince-cobbler-type scenarios are genuinely possible, see, e.g., Brison (1996).

¹² I thank Kristopher Phillips for suggesting to me this possible reply.

It is unclear, however, why such cases are impossible. To presume that such scenarios, for example, violate Natural Law would seem to assume that these cases involve (unacceptable) disruptions to souls' identities. And perhaps soul theorists might maintain that souls cannot exchange mental characteristics or fissure because such changes would disturb a soul to such a degree that it would cease to exist. But such replies depend on *assuming* a theory of soul identity; they do not provide one.

One might object instead that the problems here arise only because I assume a questionable Cartesian conception of the soul, a conception on which souls are *particulars* that have properties that might change over time. After all, perhaps only particulars can fissure. And that conception may be problematic for the reasons outlined here.

So what if the soul is not a substance or object at all, but rather itself a property? Since this proposal echoes Aristotle's (1968) *hylomorphic* view in *De Anima*—according to which the soul is the *form* of the body's *matter*; not a distinct kind of substance related to the body—Olson (2007, Sect. 7.7), calls it 'hylomorphism'. And Olson distinguishes two varieties of hylomorphism: a weaker version, on which soul properties require bodies, and a more radical *Thomistic* version, on which soul properties can persist without bodies (for a defense of the latter view, see, e.g., Stump 1995).¹³ Either way, the hylomorphic account of diachronic personal identity holds that persons persist over time not because they are related to distinct particulars, souls, but because they are particulars with persisting soul properties.

The problem with this suggestion is that it is hard to make sense of the idea of a property that changes over time. Properties are typically understood either to be *universals*—unchanging and eternal property-types—or *tropes*—particulars, property-instantiations, that constitute ordinary objects (for review, see, e.g., Conee and Sider 2005, chapter 8). Either way, properties are the things in virtue of which objects can be said to change over time; objects change over time by gaining or losing properties. Properties do not themselves change.

Perhaps soul properties change insofar as a person gains or loses *related* soul properties over time. Tropes, for example, are thought to be more or less qualitatively similar to one another. A red trope is more similar to an orange trope than to a green trope. Likewise, one might think that a soul trope A_1 at T_1 might be more or less similar to a distinct soul trope B_1 at T_2 —and in that way the soul tropes are qualitatively distinct but numerically the same. The problem with this suggestion, however, is that it simply moves the difficulties I have raised to the level of soul properties. Suppose a body bearing A_1 fissures into two distinct bodies, each seemingly bearing qualitatively identical soul tropes A_2 and A_3 . In which resultant person does the former persist? It would seem that no answer is (or could be) available. In any case, a hylomorphic view would seem to fit better with an account of the soul as unchanging—an account to which I return shortly.

¹³ One might think that this hylomorphic view is inconsistent with the central assumption that I make that souls are nonphysical, but it is not. As so-called *property dualists* have urged, it is possible that physical objects have nonphysical properties or perhaps that objects are actually metaphysically neutral substrates that have both physical and nonphysical properties (see, e.g., Chalmers 1996).

Distinguishing the objection

At this stage, it is worth distinguishing the forgoing argument from some related objections to the soul theory; doing so will clarify the nature of the objection presented here.

First, my objections are not the oft-observed *epistemic* point that we cannot *know* which body a particular soul inhabits. As James and Stuart Rachels put it in their introductory textbook:

If the Soul Theory were true... it would be hard to know whether the same kernel [of soul] were present at different times; the soul can't be seen or touched or detected in any normal way. However, we would at least know what personal identity over time consists in (2012, p. 57; cf. Shoemaker 1977).

The problem I raise, by contrast, is a *metaphysical* issue. It is not that we cannot know which body a soul inhabits, though there is some fact of the matter about which body it does inhabit. The problem is that the soul theory cannot offer a principled reason to think that a soul persists through time and so cannot explain why a particular soul might relate to one body or another.

Second, consider Locke's prince/cobbler thought experiment, which is often read as the basis for an objection to the *necessity* of souls for personal identity. Locke of course raises this possibility as a problem not only for bodily-continuity theories but also for the soul theory. If, as one might assume, the prince's soul remains in the prince's body, then the soul theory would seem to deliver the wrong result. It would seem that both physical *and* spiritual persistence is irrelevant to personal persistence. What is relevant is the continuity of their psychological states. Thus souls would seem not to be necessary for personal identity.

Similarly, consider Leibniz's (1686/1962, *Discourse on Metaphysics* xxxiv) *King of China* thought experiment. Leibniz asks: if you had the opportunity to transfer your soul into another individual's body (perhaps the body of the King of China) but with the understanding that the new body retained all of its physical and psychological properties, would you do so? The intuitive answer is 'no'. This is because, as Shoemaker (2009, p. 33) recently puts it, our practical concerns typically track our psychological or bodily features, not whatever soul features we may have. That is, souls would seem not to be *sufficient* for personal identity. Shoemaker concludes that such considerations decisively undermine the soul theory.

These points boil down to the claim that souls are irrelevant to personal identity. The problem raised here, by contrast, is that there are no criteria for *soul* identity and that consequently souls cannot explain personal identity, even if souls were relevant to it. Even if our practical concerns did not track our psychological or bodily features—and so we should not privilege those as criteria for personal identity over souls—the soul theory would face the present objections.

In this regard, my point is also distinct from Anthony Quinton's (1962) worries about the identification and individuation of souls. Quinton argues that we have no grip on the nature of a soul except by way of its psychological features.

Consequently, we have no way to identify a soul without reference to the person bearing the soul: the soul theory is at bottom circular. Quinon's argument echoes Hume's (1777/2007) discussion in *On the Immortality of the Soul*, wherein Hume maintains that he has no idea what a soul might be other than a bundle of perceptions, which are plainly changeable and tied to a mortal body.

But even if there were some way to construe souls independently of their psychological features, the soul theory would be problematic. I have been arguing thus far as though souls could vary only in terms of their mental properties (as, for example, Swinburne seems to hold), but the problem generalizes. Whatever kinds of nonphysical properties souls may have, if they change over time, then there will be fission-type worries. If a soul S_1 with nonphysical property K at T_1 splits at T_2 into two K -equivalent souls S_3 and S_4 , which resultant soul is identical with S_1 ? Whatever K might be, the soul theory seemingly cannot answer this question.

The second horn: an unchanging soul

As previously noted, perhaps the most familiar way to understand the soul is the Platonic view that the soul is an unchanging (and possibly eternal) entity that persists through time. The Hindu conception of the *Ātman* might serve as a religious example: an everlasting spiritual core of a person, independent from the body and particularized interests and projects (see, e.g., Johnson 2010, p. 37).

There are several ways to unpack this proposal. First, souls might be *bare particulars*—particulars stripped of any properties other than what minimally individuates them. Parfit (1984, p. 228) calls this the 'featureless-Cartesian view'. Second, they might be as Butler and Swinburne imagine: unchanging substances that bear (mental) properties. This is the simple view of souls (see also, e.g., Reid 1785/1969; Madell 2014). While such views do permit souls to change in certain ways, these theories count as unchanging-soul views insofar as such changes cannot alter souls' identities over time. Third, souls might involve sets of *unchanging* features: perhaps souls have fixed psychological properties, such as a static version of one's best self.¹⁴ And fourth, souls might be, as hylomorphism holds, unchanging properties of bodies. All versions are problematic.

The difficulty for such views is not that they cannot explain personal identity, for in a way they do so. Proponents of unchanging-soul views can maintain that what makes a person persist over time is the presence of her unchanging soul (itself perhaps individuated by its unique thisness).

But answering the problem of personal identity in this manner comes at a high cost, insofar as it raises a related problem—namely, *that soul theorists have no grounds for maintaining that a soul is related to a particular physical body over time*. Kim (2001) has called this the 'pairing problem' (for discussion, see Olson 2007, p. 165; cf. Shoemaker 1977). Kim raises this issue as an objection to the soul

¹⁴ This is arguably the *Mormon* conception of the soul that persists after bodily death: "[After bodily death the] soul shall be restored to the body, and the body to the soul; yea, and every limb and joint shall be restored to its body; yea, even a hair of the head shall not be lost; but all things shall be restored to their proper and perfect frame" (Church of Jesus Christ of Latter-day Saints 2013, Alma 40: 23).

theory of *synchronic* personal identity—at a particular time why is *my* soul paired with *my* body and not someone else’s? But the problem is arguably deeper for the soul theory of diachronic identity. Even if there were a principled explanation of why a soul occupies a particular body at a time, if material bodies change but souls do not, then it is unclear what could explain why that unchanging soul remains related to that changing body over time.

Views on which souls change at least have a hope of answering the pairing problem. If souls have (or are) changing psychological properties that drive bodily actions, for example, then it is plausible that each soul would occupy the body on which those psychological properties supervene. But on unchanging-soul views, it would seem that nothing could pair a particular soul with a particular body. The problem is especially clear in the case of unchanging-soul views such as the bare-particular formulation, on which souls have no changing features. We might imagine that souls are, on this view, akin to fundamental physical particles that inhere in bodies. But what then could explain why my fundamental particle does not leave my body and occupy another’s, if nothing about my body or psychology characterizes or corresponds to features of the particle? Put another way, since unchanging souls are unchanging in respect of their *intrinsic* properties, only a change in their *relational* properties could answer the pairing problem. But it is unclear what relational properties could do the trick.

Even Swinburne-type views on which unchanging souls have changing psychological features cannot address the issue. To illustrate the problem, consider what such soul theorists might say in the prince/cobbler case. Swinburne’s view might seem to offer a straightforward account: since the soul substance’s psychological properties change, the prince and cobbler’s souls swap bodies. But things are not so straightforward. It would seem equally open that their souls would remain in their respective bodies or would cease to be tethered to either body. Of course, the latter two possibilities result in the arguably unattractive consequences that the individuals either become insane or die. The point, however, is that it is unclear what resources unchanging-soul views could have for explaining *why* any of these options is correct. More importantly, to assert that one’s soul would remain in the same body or change bodies would seem to amount to hewing to either bodily-continuity or to psychological-continuity accounts, thereby rendering the soul explanation of personal identity otiose. Indeed, in the absence of an explanation of what fact(s) associate souls with bodies over time, the soul theorists’ assertion that a soul remains in a particular body (but not because the body or psychology is the criterion of personal identity) would seem to fare no better than the circular stipulation that a person persists in a particular body over time because she persists over time. Talk of souls seems merely to rename the mystery, not explain it.

The objection here echoes Locke and Leibniz’s worries, insofar as it depends on the possibility that a soul might jump from body to body—and Swinburne (e.g., 2013, p. 173) does acknowledge that his view entails this consequence. Shoemaker (2009, p. 33) argues that this possibility is absurd because it entails that we might regularly regard people as persisting when they do not. But, again, the objection here is neither the problem that souls do not seem relevant to personal identity nor is it the epistemic objection that we cannot know which body a soul inhabits. Rather,

the concern here is metaphysical: soul theorists provide no informative explanation of why a soul would inhabit a particular body over time—and hence offer an incomplete account of personal identity.

One might reply that, since unchanging-soul theorists do in a way offer an answer to the problem of personal identity, the pairing problem as I've construed it is an additional issue, one that soul theorists need not address (in this context). But this pairing problem is not an extraneous concern. It arises only as a result of this variety of soul theory's particular attempt to explain personal identity. No analogous problem arises for standard bodily- or psychological-continuity-based theories. Even if what explains why P_1 at T_1 and P_2 at T_2 are the same person is that they have the same soul, we would like to know *why* P_1 's and P_2 's respective bodies possess the same soul, given that their bodies are (potentially quite) distinct. Without an explanation of that fact, the soul theory's account of personal identity is at best partial and arguably stipulative.

As noted previously, however, some metaphysicians seem willing to claim that the fact that a person persists over time is brute and unanalyzable; thus some soul theorists might claim that the fact that a soul is tethered to a particular body is similarly inexplicable. But those theorists who deny strict criteria for personal identity, such as Parfit or Merricks, typically grant that psychological or bodily factors are at least relevant to explaining a person's persistence over time (see, e.g., Merricks 1998, p. 118ff). Since unchanging-soul theorists eschew any such considerations, it is unclear what the soul adds to the explanation of personal identity—it would seem that one might as well instead adopt a simpler physicalist account on which identity is unanalyzable.

God's sovereignty and divine knowledge

One might, however, seek answer to this problem by appeal to *God* (or the Gods): God is what explains the persistence of souls in particular bodies. After all, proponents of souls often assume that souls are creations of God and that God is sovereign.

This approach is questionable because we can construct something analogous to a *Euthyphro* dilemma for it: does a particular body keep the same soul over time—and so that body remains the same person over time—because God commands that state of affairs to be the case, or does God make a particular body keep the same soul over time because that body remains the same person? If the former, then God is simply stipulating which body remains which person, without (good) reason. If the latter, however, then there is something that makes a particular body remain the same person over time, independent of its soul. That is, there is a solution to the problem of personal identity that does not depend on souls.

The only option for the proponent of the soul theory is the first horn of this dilemma. Endorsing the second horn plainly renders the soul theory unnecessary. But it might seem that endorsing the first horn does not render the soul theory stipulative. Couldn't God have *His own* reasons for keeping a soul within a particular body over time?

The worry here, though, is that even God could not have reason for linking a soul to a particular body. Again, what happens in a scenario such as the case involving the prince and cobbler? Does God move the soul to the other body—or not? If souls consistently change or stay tied to their respective bodies, then it would seem that God is applying either a psychological- or bodily- criterion, not a soul criterion. But perhaps God need not hew to either of these broad approaches to identity. Perhaps in some situations, the soul switches bodies and in others it does not. In that case, however, what reason could *God have* for linking souls to their bodies?

One might think that God decrees that each new body has a soul that accounts for its identity because, for example, it would otherwise be hard or impossible for people to be moral agents with consistent (moral) characters. Alternatively, perhaps God keeps a soul tethered to a body because to do otherwise would be to violate Natural Law—and so God cannot or will not move souls from one body to another. But like the attempts to deny the possibility of soul fission, these sorts of replies either assume versions of bodily- or psychological-continuity views or simply beg the question at hand. Why would moving a soul from one body to another break the Natural Law if the body were not the genuine criterion for personal identity?

This problem is distinct from the issue that there might be limitations on God's knowledge. Some have argued, for example, that there are kinds of self-knowledge that only a particular individual can have and that such knowledge is incompatible with God's omniscience (see, e.g., Wierenga 2003, p. 50). Thus one might worry that, even if an individual's identity were determined by her soul, only an individual can know her own soul.¹⁵ But even if we grant that God can know about individuals' souls, it remains unclear that God could have grounds for assigning a soul to a particular body.

It is of course often claimed that God may have reasons for God's actions that mere mortals simply cannot understand (e.g., American Bible Society 1999, *Ecclesiastes* 8:16–17). Thus one might suspect that God also has a kind of special knowledge pertaining to how souls would attach to bodies under various conditions.

But invoking such mysterious divine knowledge arguably commits the fallacy of invoking the *God of the Gaps*, the error of assuming that since we have no solution as of yet to a problem, God must exist and be the solution to that problem (see, e.g., Coulson 1955). Notice that (a) this kind of claim can be made in reply to *any* metaphysical puzzle and (b) it does not offer an *informative* account of personal identity. Suppose, like Peter van Inwagen (1990) and others, we sought an answer to the *special composition problem*—the problem of what explains why a particular plurality of things composes another thing. A deist could reply: pluralities compose an object because God composes them, but we simply cannot understand how or why God composes them. This reply does not in any illuminating way answer the special composition problem nor does it seem specific to that problem.

Indeed, the problem for the soul theory is deeper than that. If one were to willing to entertain the view that what tethers soul to a particular body over time is God's unknowable will, then perhaps one should instead endorse the more parsimonious view that what makes a person persist over time is God's will and abjure souls

¹⁵ I thank Gregory Spendlove for this interesting objection.

altogether. Again, the soul “explanation” would seem to be rendered unnecessary insofar as it adds nothing to the account of personal identity. Even if persons did not have souls, we could explain personal identity via God: God makes a particular soulless body the same person over time.

Conclusions

I have argued that the soul theory of personal identity faces a dilemma. Either the soul theory cannot offer an account of soul identity or it cannot explain why souls attach to bodies, thus offering at best an incomplete account of personal identity. Assuming that people persist over time, souls do not explain that persistence.

Acknowledgements I thank Ben Abelson, Ralph Baergen, Carl Levinson, Melissa Norton, Kristopher Phillips, Jim Skidmore, and Russell Wahl for their helpful commentaries on/discussions of this material. I especially thank Gregory Spendlove for his helpful suggestions about how to frame much of this material as well as encouragement for writing the piece.

References

- Alexander, W. H. F. (1892). Be not entangled again in a yoke of bondage. *The British Friend*, 1(1), 157–158.
- American Bible Society. (1999). *The Holy Bible, King James Version*. New York: American Bible Society.
- Aristotle. (1968). *De Anima: Books II, III (with passages from Book I)* (D. W. Hamlyn, Trans.). Oxford: Clarendon Press
- Brison, S. (1996). Outliving oneself: Trauma, memory, and personal identity. In D. Meyers (Ed.), *Feminists rethink the self* (pp. 12–39). Boulder: Westview Press.
- Butler, J. (1736/2005). Of personal identity. In J. Cummings (Ed.), *The analogy of religion*. New York: Cosimo.
- Chalmers, D. (1996). *The conscious mind. In search of a fundamental theory*. Oxford: Oxford University Press.
- Church of Jesus Christ of Latter-day Saints. (Ed.). (2013). *The Book of Mormon*. Salt Lake City: Church of Jesus Christ of Latter-day Saints. Web Edition: <https://www.lds.org/scriptures/bofm?lang=eng>. Accessed Feb 5, 2016.
- Conee, E., & Sider, T. (2005). *Riddles of existence: A guided tour of metaphysics*. Oxford: Clarendon Press.
- Coulson, C. A. (1955). *Science and Christian belief*. Oxford: Oxford University Press.
- Descartes, R. (1641/1984). *The philosophical writings of descartes* (J. Cottingham, R. Stoothoff & D. Murdoch, Trans.). Cambridge, UK: Cambridge University Press.
- Forrest, P. (1998). Divine fission: A new way of moderating social trinitarianism. *Religious Studies*, 34(3), 281–297.
- Gendler, T., & Hawthorne, J. (Eds.). (2002). *Conceivability and possibility*. Oxford: Oxford University Press.
- Hume, D. (1739/2000). In D. Norton & M. Norton (Eds.), *A treatise on human nature*. Oxford: Oxford University Press.
- Hume, D. (1777/2007). On the immortality of the soul. In S. Buckle (Ed.), *An enquiry concerning human understanding and other writings* (pp. 190–197). Cambridge, UK: Cambridge University Press.
- Johnson, W. J. (2010). *A dictionary of Hinduism*. Oxford: Oxford University Press.
- Kim, J. (2001). Lonely souls: Causality and substance dualism. In K. Corcoran (Ed.), *Soul, body, and survival: Essays on the metaphysics of human persons*. Ithaca, NY: Cornell University Press.
- Kind, A. (2015). *Persons and personal identity*. Cambridge, UK: Polity Press.
- Leibniz, G. (1686/1962). *Basic writings* (G. Montgomery, Trans.). La Salle, IL: Open Court Publishing.

- Lewis, D. (1976). Survival and identity. In A. Rorty (Ed.), *The identities of persons* (pp. 17–40). Berkeley: University of California Press.
- Locke, J. (1690/1975). In P. Nidditch (Ed.), *An essay concerning human understanding*. Oxford: Clarendon Press.
- Madell, G. (1981). *The identity of the self*. Edinburgh: Edinburgh University Press.
- Madell, G. (2014). *The essence of the self. In defense of the simple view of personal identity*. New York: Routledge.
- Mandik, P. (2013). *This is philosophy of mind: An introduction*. Malden, MA: Wiley-Blackwell.
- Merricks, T. (1998). There are no criteria of identity over time. *Noûs*, 32(1), 106–124.
- Olson, E. (2007). *What are we? A study in personal ontology*. Oxford: Oxford University Press.
- Olson, E. (2015). Personal identity. In E. Zalta (Ed.), *The stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/fall2015/entries/identity-personal/>. Accessed Nov 30, 2015.
- Parfit, D. (1971). Personal identity. *Philosophical Review*, 80, 3–27.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Plato. (1997). In J. Cooper (Ed.), *Plato: Complete works*. Indianapolis: Hackett.
- Quinton, A. (1962). The soul. *The Journal of Philosophy*, 59, 393–403.
- Rachels, J., & Rachels, S. (2012). *Problems from philosophy* (3rd ed.). New York: McGraw Hill.
- Reid, T. (1785/1969). In B. A. Brody (Ed.), *Essays on the intellectual powers of man*. Cambridge, MA: MIT Press.
- Shoemaker, S. (1977). Immortality and dualism. In S. C. Brown (Ed.), *Reason and religion*. Ithaca, NY: Cornell University Press.
- Shoemaker, D. (2009). *Personal identity and ethics: A brief introduction*. Peterborough: Broadview Press.
- Sider, T. (2001). *Four dimensionalism*. Oxford: Oxford University Press.
- Stump, E. (1995). Non-Cartesian substance dualism and materialism without reductionism. *Faith and Philosophy*, 12(4), 505–531.
- Swinburne, R. (1984). Personal identity: The dualist theory. In S. Shoemaker & R. Swinburne (Eds.), *Personal identity* (pp. 1–66). Oxford: Blackwell.
- Swinburne, R. (1997). *The evolution of the soul*. Oxford: Oxford University Press.
- Swinburne, R. (2013). *Mind, brain, and free will*. Oxford: Oxford University Press.
- Swinburne, R. (2014). What makes me me? A defense of substance dualism. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense*. New York: Routledge.
- Unger, P. (1979). I do not Exist. In Graham MacDonald (Ed.), *Perception and identity*. London: Macmillan.
- van Inwagen, P. (1990). *Material beings*. Ithaca, NY: Cornell University Press.
- Wierenga, E. (2003). *The nature of god: An inquiry into divine attributes*. Ithaca, NY: Cornell University Press.
- Williams, B. (1970). The self and the future. *Philosophical Review*, 79(2), 161–180.
- Zimmerman, D. W. (1991). Two Cartesian arguments for the simplicity of the soul. *American Philosophical Quarterly*, 28, 217–226.