



Johnston versus Johnston

Kacper Kowalczyk¹

Received: 25 April 2021 / Accepted: 25 January 2022

© The Author(s) 2022

Abstract

Personites are like continuant people but shorter-lived. Johnston argues that personites do not exist since otherwise personites would have the same moral status as persons, which is untenable. I argue that Johnston's arguments fail. To do that I propose an alternative way to understand intrinsicness and I clarify the meaning of reductionism about persons. I also argue that a plausible ethical theory is possible even if personites have the same moral status as persons. My arguments draw on Johnston's earlier debate with Parfit about personal identity and the place of ordinary concerns in a naturalistic world. I also describe an important but metaphysics-free problem that arises from Johnston's discussion.

Keywords Personal identity · Four-dimensionalism · Intrinsic properties · Value theory

1 Introduction

How do we persist through time? According to *worm theory*, we persist through time because we have different parts at different times, much like we extend through space because we have different parts at different places. So when you look at me now, what you see is only a part of me, my present *person-stage*. But I have other person-stages at other times. I am a *spacetime worm*.

A typical worm theorist also says that, in addition to worms like you and me, there are other more gerrymandered worms. One is the worm that consists of my person-stages up to now and your person-stages from now on. Another is the worm that consists of my person-stages up to yesterday, but which has no further person-stages.¹

I differ from the first worm in that my person-stages are *R*-related, where *relation R* is psychological continuity and/or connectedness, *psychological connectedness* is the

¹ See Lewis (1976) and Sider (2001a). Atypical worm theories are possible but I will not discuss them. See Magidor (2016).

✉ Kacper Kowalczyk
kacper.kowalczyk@ucl.ac.uk

¹ University College London, London, UK

holding of particular direct psychological connections such as memory and intention, and *psychological continuity* is the holding of overlapping chains of strong connectedness.² And I differ from the second worm in that I am *maximal*: I am not a proper part of another worm that is also *R*-interrelated.

I am what Lewis calls a ‘person’: a maximal *R*-interrelated aggregate of person-stages.³ The first worm is what Olson calls a ‘crossperson’: an aggregate of person-stages that is not *R*-interrelated.⁴ The second worm is what Johnston calls a ‘personite’: an aggregate of person-stages that is *R*-interrelated but not maximal.⁵

Johnston argues that personites do not exist since otherwise personites would have the same moral status as persons, which is untenable.⁶ If both have moral status, then we have too many things with moral status. For example, when you promise to meet me on the other side of the lake, are you not also making unkeepable promises to the masses of personites that happen to coincide with me at that time but cease to exist before I cross the lake? So is it ever alright to make promises?⁷ And if persons lack moral status, then we have too few things with moral status. For example, when you promise to meet me on the other side of the lake, why should you keep your promise given that I lack moral status?

This is Johnston’s

Personite Problem. There are persons and personites. If persons have moral status, then so do personites. But every plausible ethical theory implies that persons have moral status and personites do not.⁸

Johnston’s conclusion is that there is something wrong with worm theory and, by extension, with a naturalistic picture of the world that it exemplifies.⁹

Johnston has two main arguments for the claim that persons have moral status only if personites do. The first I call

The Intrinsicness Argument. Every personite is a duplicate of some possible person. Moral status is intrinsic. An intrinsic property is shared between possible duplicates. So if persons have moral status, then so do personites.

The second I call

² These are Parfit’s (1987: p. 206) definitions. I follow Johnston (2016, 2017) in that I focus on psychological rather than physical relations and on continuity rather than connectedness. But no generality is lost either way.

³ See Lewis (1976).

⁴ See Olson (2010).

⁵ See Johnston (2016, 2017). Olson (2010) would call it a ‘subperson’.

⁶ What do I mean by ‘moral status’? For simplicity’s sake, I will follow Johnston (2017: p. 621) in thinking of moral status as conferring moral weight on one’s interest and perhaps generating side-constraints on other people’s conduct.

⁷ Based on Johnston’s (2016: p. 208; 2017: p. 630) example.

⁸ Olson (2010) discusses what is essentially the same problem, and his arguments are similar to Johnston’s. But I will focus on Johnston’s more recent and more extensive presentation.

⁹ Johnston does not spell out how the extension is supposed to work, so I will mostly ignore it in this paper.

The Triviality Argument. Every personite differs from some possible person only in a way that does not matter. Whether something has moral status cannot depend on a difference that does not matter. So if persons have moral status, then so do personites.

Johnston has another similar argument that involves a different person-like object in a worm theorist's metaphysics. I will get to it in Sect. 4.

I argue that none of these arguments succeed. I cover the Intrinsicity Argument in Sect. 2, the Triviality Argument in Sect. 3, and a variant of the Triviality Argument in Sect. 4. I often draw on Johnston's own earlier work on personal identity.¹⁰ Hence the paper's title. This is not to say that the early Johnston is always right or that the later Johnston is always wrong. It is merely to say that the early Johnston noted some important distinctions and objections that bear on the later Johnston's arguments. In Sect. 5, I also argue that much of commonsense ethics could be salvaged even if Johnston is right that persons have moral status only if personites do. In Sect. 6, I conclude with some general lessons about metaphysical arguments for ethical conclusions and I identify an important but metaphysics-free problem similar to Johnston's.

2 The Intrinsicness Argument

Take Tweedledee, an ordinary person. According to worm theory, Tweedledee is a spacetime worm. But now consider Tweedle, a worm that is exactly like Tweedledee, except shorter by a few years. Tweedle is also, uncontroversially, a person. Both are shown in Fig. 1.

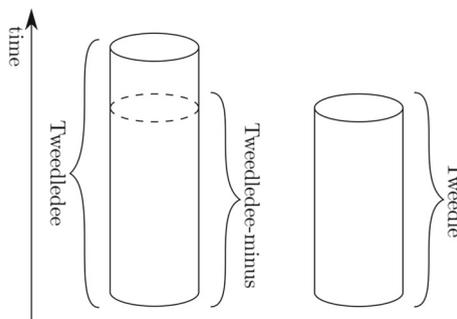


Fig. 1 Tweedledee etc.

But note that Tweedledee himself includes a shorter spacetime worm that looks very much like Tweedle. That's Tweedledee-minus, one of Tweedledee's many personites. Tweedle and Tweedledee-minus are indistinguishable in microphysical, chemical, and biological terms.

Johnston would say that they are intrinsically indistinguishable. It would be more neutral to say that they are *duplicates* in Lewis's sense, that is, that there is a one—one

¹⁰ See Johnston (1989a, 1989b, 1992, 1997).

correspondence between their parts (temporal and spatial) which preserves perfectly natural relations and properties, where perfectly natural properties are those that make for objective similarity and figure in our best science.¹¹ It's just that the demise of Tweedledee-minus is followed by some further person-stages which carry on with his psychology and his body.

But moral status is plausibly intrinsic. As Johnston puts it, it is “quite different from such things as being famous, being vindicated or being the best Ping-Pong player ever” (2017: p. 622). So it seems that if Tweedle has moral status, then so does Tweedledee-minus.

We can formulate Johnston's argument as follows:

- (1) Persons have moral status.
- (2) Moral status is intrinsic.
- (3) If a property is intrinsic, then every possible duplicate of something that has it also has it.¹²
- (4) Every personite has a possible duplicate that is a person.¹³

Therefore:

- (5) Personites have moral status.¹⁴

Given that I both have moral status and host personites, this would mean that there are as many things with moral status that share my spatiotemporal career as there are possible continuant persons that differ from me by a few years, days, and so on, at either end. Surely that's too many things with moral status.

Johnston concludes that worm theory is to blame and, by extension, naturalism. But, contrary to Johnston, the problem has nothing to do with ethics or the metaphysics of persistence. To see this consider a similar argument about cats.

Take Tibbles the cat and Tib, a cat that is exactly like Tibbles except shorter by a tail. Tib is also, uncontroversially, a cat. They are both shown in Fig. 2.¹⁵

¹¹ See Lewis (1983) and Lewis (1986: pp. 50–69).

¹² This is weaker than Lewis's (1983: pp. 355–358, 1986: pp. 61–63) *analysis* of intrinsicness in terms of duplication (and of duplication in terms of perfect naturalness and parthood).

¹³ Johnston (2016: p. 204) notes that this can be weakened to say that every personite has a duplicate that has moral status.

¹⁴ This is close to the formulation in Johnston (2016: pp. 203–205), except that here and below I write as if Lewis's modal realism is true, so that there is no need to repeatedly relativize duplication and property instantiation to possible worlds. I hope that no confusion will result.

¹⁵ Tibbles's story is told by Wiggins (1968) who credits it to Geach. A slightly different version appears in Geach (1980: pp. 215–218). Similar examples go back to the Stoic Chrysippus, on which see Burke (1994). It is worth noting, however, that Wiggins's, Geach's and Burke's arguments that involve Tibbles are different from Johnston's argument which is actually more similar to the argument of Olson (1995).

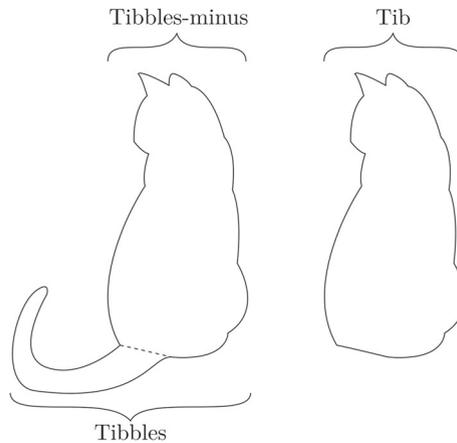


Fig. 2 Tibbles etc.

But note that Tibbles himself includes a smaller cat-like thing that looks very much like Tib. That's Tibbles-minus, one of Tibbles's many cat-like proper parts. Tib and Tibbles-minus are indistinguishable in microphysical and chemical terms. They are arguably duplicates in Lewis's sense. It's just that Tibbles-minus happens to be attached to a tail.

We can now formulate an argument similar to Johnston's own:

- (1') Cats have cathood.
- (2') Cathood is intrinsic.
- (3') If a property is intrinsic, then every possible duplicate of something that has it also has it.
- (4') Every cat-like proper part of a cat has a possible duplicate that is a cat.

Therefore:

- (5') Cat-like proper parts of cats have cathood.

Given the existence of both cats and cat-like proper parts of cats, this would mean that there are as many cats within my cat as there are possible cats that differ from it by a whisker, a paw, and so on. Surely that's too many cats.

I claim that if Johnston's argument works, then so does this argument. Similar arguments would then work in the case of other ordinary properties such as rock-hood, house-hood, applehood, personhood, and so on.¹⁶ But this argument has nothing to do with ethics or the metaphysics of persistence.¹⁷ And most of us think it is unsound

¹⁶ These arguments might have similar ethically revisionary implications as Johnston's argument. See Hudson (2001: pp. 34–44) and Jones (2010: pp. 55–62). Indeed, given that moral status at a time is intrinsic, an argument similar to Johnston's would show that, at a given time, a person might include proper spatial parts that have moral status.

¹⁷ This, by the way, shows that Kaiserman's (2019) stage-theoretic response to Johnston's argument is not general enough. On stage theory, see, for example, Sider (1996, 2001a). Another problem is that stage theory shares worm theory's metaphysics but merely maps our person talk differently onto things in that metaphysics. Since "in certain circumstances, such as when we take the timeless perspective, reference is to worms rather than stages" (Sider 1996: p. 448), personites can still cause trouble even for a stage theorist.

since most of us believe in cats and cat-like proper parts of cats yet do not believe that there are many cats within my cat.¹⁸

What an argument like this shows, I take it, is that either most ordinary properties such as cathood and, indeed, moral status, are not intrinsic,¹⁹ or that an intrinsic property need not be shared between duplicates in Lewis's sense.²⁰ I think it would be far-fetched to conclude that Tib and Tibbles-minus are not duplicates in Lewis's sense.²¹ Either way, Johnston's argument would also fail.

I think it would be unhelpful to conclude that most ordinary properties are not intrinsic, since there is still an intuitive difference between cathood and obviously extrinsic properties such as being the best Ping-Pong player ever. It would be better to find another way to understand intrinsicness. A promising approach is to restrict duplication as in, for example:

(*) If a property is intrinsic, then every possible duplicate of something that has it also has it, provided it is not a proper part of something that also has that property. Put differently: a possible duplicate might lack an intrinsic property of the original but only if it is a proper part of something that still has that property. So there is a sense in which duplication preserves intrinsic properties.²²

Johnston once gave a similar response to a similar problem for the principle that personal identity through time is intrinsic. He noted that the principle might lead to too many continuant persons. His response was that it should be understood to mean that if a process secures the survival of a person, then so does every possible intrinsically indistinguishable process that "is not a part of some more inclusive process which secures the survival of a person" (1989a: p. 396).²³

A restricted duplication principle like this has a number of advantages. First, it gives a sense in which cathood and moral status might be intrinsic but being the best Ping-Pong player is not, without at the same time threatening a multiplication of cats or things with moral status.

Second, it does not imply that a cat cannot be a proper part of a cat or that something with moral status cannot be a proper part of something with moral status. It does not preclude the possibility of Johnston's Twenty-Oners: imaginary onion-like creatures whose twenty-one layers all have separate conscious lives.²⁴ What follows instead is that we cannot appeal to the intrinsicness of moral status to argue that a Twenty-Oner's layer has moral status on the grounds that it has a peeled duplicate that clearly has moral

¹⁸ But see Van Inwagen (1981) and Olson (1995) who argue that cats lack large proper parts, and Lewis (1993) who argues that there are many almost-identical cats within my cat.

¹⁹ See Hawley (1998), Sider (2001b, 2003), Weatherson (2001), Wasserman (2005), Walsh (2011), and Francescotti (2014).

²⁰ See Williams (2013). Ralf Bader develops a similar idea in his unpublished work. Sosa (1990) and Hawley (2005) seem sympathetic to this response but ultimately unconvinced. See also Jones (2010: pp. 36–39).

²¹ But see Merricks (1998, 2003) and Burke (2003) about similar examples that involve consciousness rather than cathood.

²² Compare Williams (2013) who argues that 'intrinsic' should often be understood to mean 'part-intrinsic', where a property is *part-intrinsic* if, and only if, every possible duplicate of something that has it also has it or is a proper part of something that has that property.

²³ Italics removed. Johnston credits Mark Hinchliff with bringing the problem to his attention.

²⁴ See Johnston (2017: p. 625). See also similar examples in Sutton (2014).

status. We must argue that a Twenty-Oner's layer has moral status on other grounds. So there is no need to disagree with Johnston's claim that "we would recognize the moral status of a person embedded within another larger person" (2016: p. 205).

Third, it can underwrite the role that Johnston wants intrinsicness to play in the story of humanity's moral progress in overcoming racism, sexism, and speciesism. As he puts it, "[h]ow (...) may we rationally reconstruct the deeply admirable, even if shamefully belated, expansion of the protected circle, if not by way of the historically crucial appeal to intrinsic similarities?" (2016: p. 211). Expanding the protected circle by appeal to intrinsic similarities is plausible relative to a range of ways to understand what an intrinsic similarity is. At any rate, I doubt that the expansion of the protected circle has to be underwritten by the principle of intrinsic moral status rather than, say, the principle of equal consideration of interests, according to which "the interests of every being affected by an action are to be taken into account and given the same weight as the like interests of any other being" (Singer, 2011: p. 20).

But while I think that it is a good idea to restrict duplication, I am not sure whether I identified the correct restriction. Consider, for example, the property of being a misshapen cat and the property being a lightweight cat. These seem intrinsic if cathood is, at least partly because shape and size properties seem intrinsic and intrinsic properties are supposed to be closed under conjunction. Now note that a misshapen cat can have a duplicate that is a proper part of a cat that is not misshapen. But such a duplicate is not a cat, let alone a misshapen cat, contrary to the restricted principle I described. Likewise a lightweight cat can have a duplicate that is a proper part of a cat that is not lightweight. But such a duplicate is not a cat, let alone a lightweight cat.²⁵ One response is to deny that these properties are intrinsic. Another response is to find a different restriction, for example, a restriction to duplicates that are not proper parts of anything which either does or *could* have the property in question. Then since every cat could be misshapen, a duplicate of a misshapen cat that is a proper part of a cat does not satisfy the new restriction. Likewise for a duplicate of a lightweight cat that is a proper part of a cat. Perhaps other restrictions are better still.²⁶

3 The Triviality Argument

I conclude that Johnston's Intrinsicness Argument does not succeed. It is an instance of a problem in the theory of intrinsicness unrelated to ethics or the metaphysics of persistence.

²⁵ There are also diachronic and ethical examples such as the property of being a short-lived person and the property of having a life that gets better over time. A similar problem arises if cats can cease to be cats due to extra parts, for example, if a tail-less cat with a dog's tail is neither a cat nor a dog but a cat-dog chimera. I learned a similar example from Ralf Bader. Sider's example (2001b) of a golden sphere embedded in a golden cube is also similar.

²⁶ Another problem is that a duplicate of a cat might be created by a freakish lightning hitting a swap rather than by the usual methods of cat propagation. Many people intuit that such a duplicate cannot be a cat. I do not share the intuition and, besides, the problem is orthogonal to Johnston's argument and irrelevant to arguably nonbiological properties such as personhood and moral status, which is why I do not discuss it in the main text.

What about Johnston's other argument, the Triviality Argument? We can formulate it as follows:

- (1) Persons have moral status.
- (2) If one thing has moral status, and another differs from it only in ways that do not matter, then the other also has moral status.
- (3) It does not matter that something is maximal rather than not.
- (4) For every personite there is some possible person that differs from it only in being maximal rather than not.

Therefore:

- (5) Personites have moral status.

The problem with Johnston's argument is that it could just as well be run in reverse to conclude that maximality does matter after all. I take it that Johnston's reason to believe that maximality does not matter is that, when considered independently of its purported relationship to personhood, it does not seem to matter whether an R -interrelated aggregate is maximal. But that is not enough for a good argument since someone might agree with Johnston's reason, but also believe that maximality matters in view of its purported relationship to personhood. To make his argument irreversible, Johnston owes us a better reason to believe that maximality does not matter.

Johnston once gave a similar response to Parfit's similar argument about what matters in a case like

Teletransportation. Derek's body is scanned and destroyed in a booth on Earth. In another booth on Mars Derek's Replica is soon created using Martian matter and the information received from Earth. Replica seems to have all of Derek's memories, intentions, and so on, and a body that is exactly like Derek's body on Earth just before it was destroyed.²⁷

Parfit claimed that in this case it does not matter that relation R has an abnormal cause (the scanner) rather than its normal cause (the brain). His reason was that the cause of R does not seem to matter when considered independently of its purported relationship to personal identity through time. He concluded that, insofar as personal identity through time can depend on whether R has a normal cause, it cannot be what matters in survival.²⁸

Johnston once argued that personal identity through time does depend on whether R has a normal cause.²⁹ But he did not agree with Parfit that personal identity through time is not what matters in survival. Against Parfit's "argument from below", Johnston suggested, "we should 'argue from above'. Brain-based realization of R , although not significant in itself, is of great derivative significance. For it is a necessary condition of our continued existence" (1992: p. 605).³⁰ To make his argument irreversible, Parfit owes us a better reason to believe that the cause of R does not matter.

²⁷ Based on Parfit's (1987: pp. 199–200) example.

²⁸ See Parfit (1987: pp. 285–6).

²⁹ Johnston (1989b) also argued that it is, in some sense, up to us whether we survive teletransportation. Johnston (2010: pp. 241–304) agrees and elaborates.

³⁰ See also Johnston (1997: pp. 162–169).

There are two other observations about Parfit's argument that also bear on Johnston's latest argument. First, we seem ready to argue from above elsewhere.³¹ Suppose, for example, that pain turns out to be C-fibre firing. It is in fact true that there exist A-Delta fibres that are like C-fibres but slightly thicker because myelinated. When considered independently of its purported relationship to pain, a fibre's myelination does not seem to matter. Still, I think that after such a discovery, we would conclude that the difference matters after all.³² Second, an argument like Parfit's cannot always be successful unless nihilism is true. This is because most if not all that matters are macroscopic phenomena with microscopic bases that do not seem to matter when considered independently of their purported relationship to macroscopic phenomena.³³

I do not claim that the early Johnston is correct about Teletransportation. Nor do I claim that arguments from above always win against arguments from below. But I do claim that the later Johnston's argument fails, much like Parfit's earlier argument, because, given his reason to believe its key premise, that argument could just as well be run in reverse.

4 The Continuity-Variant Argument

I take it that the later Johnston does, in effect, try to provide such a reason in the context of a similar argument that involves not personites but a different kind of person-like object in a worm theorist's metaphysics.³⁴

To see how it goes, consider

Tweedledee's Adventure. Tweedledee undergoes psychological change at an accelerating rate as he grows older. One month he goes to university, the next month he takes psychedelic drugs, reads Nietzsche, falls in love, the month after he takes on a new job, starts a family, gets drafted into an army, goes into a warzone, and converts to a new religion. And so on.³⁵

The story is illustrated in Fig. 3 where shading represents the strength of psychological connectedness.

³¹ See Shoemaker (1985), Sosa (1990), and Garrett (1992).

³² The (fictional) example of pain = C-fibre firing has been popularized by Kripke (1972). See Puccetti (1977) on why the identity is scientifically doubtful and on the difference between C- and A-Delta fibres.

³³ See Johnston (1992: p. 592) and Johnston (1997: pp. 167–8).

³⁴ The argument can be found in Johnston (2016: pp. 218–222).

³⁵ Based on Johnston's (2016: p. 222) example.

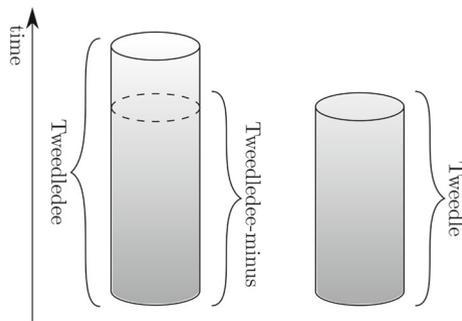


Fig. 3 Tweedledee's Adventure

Given that R (psychological continuity and/or connectedness) is the constitutive unity relation of persons, chains of strong psychological connectedness will continue as long as Tweedledee does. But chains of slightly stronger psychological connectedness will break shortly before that. Let Tweedle be a worm that is like Tweedledee except that it ends just when these chains of slightly stronger connectedness are about to break.

But note that Tweedledee himself includes a shorter spacetime worm that is a duplicate of Tweedle. That's Tweedledee-minus, an example of what Johnston calls a 'continuity-variant', a person-like worm whose constitutive unity relation is not relation R but relation R' , defined like R except in terms of a slightly stronger degree of connectedness. Continuity-variants are like persons but less able to withstand psychological change. Both are maximal with respect to their respective constitutive unity relations.

When considered independently of its purported relationship to personhood, it might seem not to matter whether something's constitutive unity relation is R rather than R' . So it is tempting to argue as follows:

- (1) Persons have moral status.
- (2) If one thing has moral status, and another differs from it only in ways that do not matter, then the other also has moral status.
- (3) It does not matter that something's constitutive unity relation is R rather than R' .
- (4) For every continuity-variant there is a possible person that differs only in that its constitutive unity relation is R rather than R' .

Therefore:

- (5) Continuity-variants have moral status.

This is what I call 'the Continuity-Variant Argument'. It is another argument from below, so we could again try to argue from above: it matters whether something's constitutive unity relation is R rather than R' because R and not R' is the constitutive unity relation of persons.

But this time Johnston suggests a reason not to reverse the argument, at least if we accept a naturalistic view such as Parfit's reductionism, according to which "a person's

existence just consists in the existence of a body, and the occurrence of a series of thoughts, experiences, and other mental and physical events” (Parfit, 1995: p. 16).³⁶

Johnston claims that, on such a view, the fact that a person’s constitutive unity relation is R rather than R' “cannot arise from some fact about the essential conditions of survival of some enduring soul pellet, Cartesian ego or separately existing mental entity ‘distinct from our brains and bodies’. It can only be a fact that arises from how we use terms like ‘person’ and ‘numerically the same person’” (2016: p. 225).

If Johnston is right, then his latest argument does not seem reversible like the earlier one. To reverse it would be to argue that it matters whether something’s constitutive unity relation is R rather than R' because the constitutive unity relation of persons is R rather than R' . But the constitutive unity relation of persons is R rather than R' because of how we use some words. And it plausibly cannot be the case, outside of literature or law, that some fact matters ultimately because of how we use some words.³⁷

But I think that Johnston is wrong about what follows from a naturalistic view such as Parfit’s reductionism. What follows is that facts about persons reduce to facts about mental and bodily continuity (as opposed to facts about immaterial souls). But reductionism does not constrain the character of the bridge laws that effect the reduction. It could be that these bridge laws are not linguistic in character. Indeed, it is hard to see how they could be linguistic since it is compatible with facts about mental and bodily continuity that there are no persons at all (at any rate, no persons that persist through time). And it plausibly cannot be the case, outside of literature or law, that words can create new things.³⁸

Johnston once gave a similar response to Parfit’s similar discussion of reductionism. He maintained that in addition to *eliminativism*, according to which persons are at best linguistic fictions, and *non-reductionism*, according to which persons are Cartesian egos, souls, or the like, there is also room for *reductionism with ordinary further facts*, according to which persons are neither linguistic fictions nor superlative entities. But then the bridge laws that connect facts about persons to facts about bodily and mental continuity could also be neither linguistic nor superlative. Johnston now seems to assume otherwise.³⁹ But that is a false dilemma.⁴⁰

³⁶ See also Parfit (1987: p. 211).

³⁷ In response to the early Johnston’s suggestion that we should argue from above rather from below in the case of personal identity through time, Parfit (1995) claimed that to do so would be to treat language as more important than reality. See also Parfit (2007) and Johnston (2010: pp. 305–377). The later Johnston seems to agree with Parfit’s claim, at least in the case of continuity-variants.

³⁸ This is one of Schaffer’s arguments for his claim that “in all concrete transitions from more to less fundamental, the dependence functions involved provide substantive information” (2017: 10). Perhaps a use/mention fallacy might make it seem plausible to think otherwise (this is not to say that Johnston is guilty of such a fallacy). It is obvious that whether I am in the extension of the word ‘person’ depends on how we use it. (Imagine we used ‘person’ to refer to cabbages.) But it is unobvious that whether I am a person depends on how we use the word ‘person’.

³⁹ Johnston is in good company, however. See Chalmers (1996: pp. 71–89), Sidelle (2007), and Heller (2008).

⁴⁰ This also shows, I think, how to respond to Johnston’s appeal to possible communities with different opinions about how much change a person can withstand: “Would they have missed some distinguished set of facts in the reductionist base? No; they would simply be extracting one kind of information about identity over time from the reductionist base” (2016: 226). They would not have missed anything in the

5 Personite ethics

So far I argued that Johnston fails to show that if persons have moral status, then so do personites. But suppose that I am incorrect. I will now argue that the personite problem can still be avoided since a plausible ethical theory is compatible with the claim that persons have moral status only if personites do.

As an example of ethical problems that personites supposedly cause, consider

The Budapest Case. Mark is invited to go to Budapest for three months next summer. Since he enjoys talking to the locals in their native language when abroad, he decides to take a month-long intensive course in Hungarian. He finds learning the language very difficult. When he arrives in Budapest, Mark finds it was worth the effort after all.⁴¹

Note that many of Mark's personites go through the trouble of learning Hungarian but cease to exist before reaching Budapest. Intuitively, it would be wrong for Mark to coerce many people to work for him for a month, just so that he can have a pleasant stay in Hungary. But that is relevantly like what Mark is doing given that persons have moral status only if personites do.⁴²

Johnston concludes from cases like this that it is immoral to be prudent, that is, to make present sacrifices for the sake of larger future gains. Personites likewise appear to threaten the commonsense ethics of compensation, promising, punishment, and so on.

Johnston admits that some ethical theories are not threatened by similar problems. He says that the Personite Problem "might be read as an argument for [the] most primitive form of hedonistic account of the good-making features" which treats both persons and personites as mere receptacles of pleasures and pains (2017: p. 642).

But that is an overstatement. The Personite Problem shows at best that it cannot matter how person-stages are packaged into persons. But that is compatible with the importance of non-hedonic features of person-stages (for example, knowledge), the existence of deontic prohibitions that protect person-stages (for example, the prohibition on violent interference with a person-stage), and with the importance of other relations between person-stages (for example, relation *R* or welfare equality). Personite ethics need not be as crude as Johnston makes it out to be.

Still, we might wonder how much of the commonsense ethics of compensation, promising, punishment, and so on, can also be salvaged. A promising approach is to reformulate that part of commonsense ethics in terms of what persons and personites share: *R*-interrelatedness.⁴³

For example, commonsense ethics has it that

- (*) *x* at some time is bound by a promise made by *y* at another time if, and only if, *x* is the same person as *y*.

Footnote 40 continued

reductionist base, but they might have missed a bridge law that connects the reductionist base with facts about persons.

⁴¹ Based on Johnston's (2016: p. 212, 2017: p. 623) example.

⁴² Johnston (2017: p. 632) argues that personites cannot give informed consent to what's happening.

⁴³ See Johnston (2016: pp. 211–215), inspired by Parfit (1987: pp. 307–347).

After reformulation we instead have the principle that

- (**) x at some time is bound by a promise made by y at another time if, and only if, x 's person-stage at the first time is R -related to y 's person-stage at the second time.

Likewise instead of the principle that

- (†) a benefit to x at one time fully compensates a burden to y at another time if, and only if, x is the same person as y ,

we have the principle that

- (††) a benefit to x at one time fully compensates a burden to y at another time if, and only if, x 's person-stage at the first time is R -related to y 's person-stage at the second time.

And so on. Additional adjustments will have to be made if, in light of the existence of continuity-variants, we conclude that persons and continuity-variants have the same moral status. For example, instead of the last principle we will have the principle that

- (†††) a benefit to x at one time compensates a burden to y at another time to the extent that x 's person-stage at the first time is R -related to y 's person-stage at the second time.

So why not think that a plausible (indeed, commonsense) ethical theory is possible even if persons and personites have the same moral status? Johnston's main complaint is that commonsense ethical principles, when reformulated in terms of relation R , are unfairly biased against personites and in favour of persons, so they do not really respect their equal moral status:

“a person can only be compensated by experiencing or otherwise receiving the benefit in question; whereas a personite can be compensated by another being experiencing or otherwise receiving the benefit in question” (2016: p. 204).⁴⁴

But that is incorrect. To see why, we can consider cases of division and amnesia. As an example of division, consider

⁴⁴ Johnston has two other complaints. One is that these principles would imply that, in the Budapest Case, Mark is obliged to learn Hungarian even if the net benefit to Mark is zero because his personites which come into existence after arriving in Budapest receive “a special benefit in happening to know Hungarian ‘innately’” (2016: 215), and, so, tip the scale in favour of learning Hungarian. I doubt that follows. First, if Mark's post-arrival person-stages have the benefit of knowing Hungarian without having to learn it, that would seem to be offset by the harm to pre-arrival person-stages of studying Hungarian without ever getting to master it. Second, to say that I know Hungarian innately is to say that I know Hungarian and was never someone who did not know Hungarian and, so, it is to say something about personal identity through time. Since we now have to reformulate our ethical principles in terms of relation R , we must replace the principle that there is a special benefit in knowing Hungarian without being R -related to a person-stage that did not know Hungarian. Mark's post-arrival person-stages do not get *that* benefit. Johnston's other complaint is that an aggregative principle of beneficence reformulated in terms of relation R will lead to implausible double-counting since personites overlap (2017: pp. 635–641). But that is not an implication of all such principles. See Holtug (2010: pp. 290–340) and Thomas (2016: pp. 161–254).

Tweedle's Division. Tweedle's body is fatally injured. Tweedle's brain is divided in half, and each half is successfully transplanted into two healthy but brainless bodies of his twin brothers. Each of the resulting people, believes that he is Tweedle, seems to remember living Tweedle's life, and so on. And each has a body that is very much like Tweedle's.⁴⁵

Let's use, as is customary, 'Lefty' to refer to the recipient of Tweedle's left hemisphere and 'Righty' to refer to the recipient of Tweedle's right hemisphere. Figure 4 provides a worm-theoretic depiction of the story.

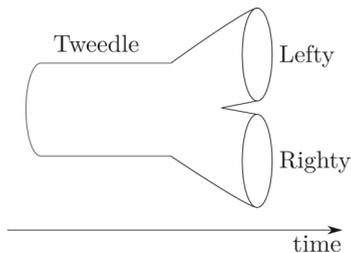


Fig. 4 Tweedle's Division

What happens in Tweedle's Division? One view is that there are two persons, Lefty and Righty, sharing Tweedle's body all along. 'Tweedle' is really an ambiguous name.⁴⁶ But even though Lefty is not Righty, Lefty's pre-division person-stage is strongly *R*-related to Righty's post-division person-stage, and, so, according to our new principle of compensation, a burden to Lefty before the division can be fully compensated by a benefit to Righty after the division.⁴⁷

Another view is that Tweedle ceases to exist and Lefty and Righty are two new persons.⁴⁸ But even though Tweedle is not Righty, Tweedle's pre-division person-stage is strongly *R*-related to Righty's post-division person-stage, and, so, according to our new principle of compensation, a burden to Tweedle before the division can be fully compensated by a benefit to Righty after the division.

Either way, we have a case where a burden to a person can be fully compensated by a benefit to some other being. Contrary to Johnston, this might happen to persons as well as to personites if either of these views of Tweedle's Division is correct.

But perhaps Tweedle is both Lefty and Righty in Tweedle's Division, so that after the division he has two bodies and two minds.⁴⁹ If we agree, we should consider another example,

⁴⁵ Based on Parfit's (1987: p. 254) example.

⁴⁶ See Lewis (1976).

⁴⁷ Compare Parfit's (1976) response to Lewis.

⁴⁸ See Nozick (1981: pp. 29–70). Parfit (1987: p. 260) thought that this would be the best description even if not the truth. Note that if this is our view, we can no longer claim that a person is a maximal *R*-interrelated aggregate of person-stages but must replace *R* with some other relation that does not allow for branching.

⁴⁹ The early Johnston (1989a) argued that all views about what happens in division cases are on a par and, so, it is indeterminate what happens. The later Johnston (2010: p. 308) thinks, similarly to Dainton (1992),

Tweedledee's Amnesia. Tweedledee lives a normal life, except for frequent episodes of amnesia, mood swings, and character shifts around his 40th birthday. The magnitude of these psychological changes is significant but not enough to make it the case that Tweedledee before his 40th birthday is a different person from Tweedledee afterwards.

This case involves at least two personites: Tweedle, who expires at Tweedledee's 40th birthday, and Dee, who begins right after. This is shown in Fig. 5 where shading again represents the strength of psychological connectedness.

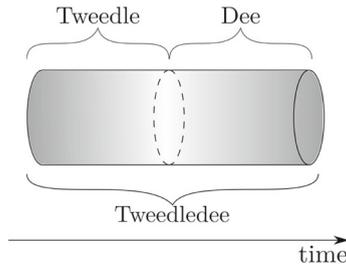


Fig. 5 Tweedledee's Amnesia

Suppose that Tweedledee works hard in his twenties to save for a future vacation. He can have a month-long vacation if he withdraws his savings at age 35 but, due to the magic of compound interest, a two-month-long vacation if he withdraws at age 45.

Since his youthful person-stages are R-related to his 35-year-old person-stage more strongly than to his 45-year-old person-stage, it could be that, according to our new principle of compensation, a month-long vacation at 35 better compensates his hard work than a two-month-long vacation at 45. But if he goes on vacation at age 35, that will mean that the career of Tweedledee the person will contain one month of vacation when it would otherwise have been two, while the career of Tweedle the personite will contain one month of vacation when it would otherwise have been zero. This is another case where personites seem to be favoured over persons.

6 Minimalism

I conclude that Johnston fails to show that persons have moral status only if personites do. Johnston's Intrinsicness Argument is an instance of a problem in the theory of intrinsicness unrelated to ethics or the metaphysics of persistence. Johnston's Triviality Argument is unconvincing since Johnston gives us no reason not to reverse it. Johnston's Continuity-Variant Argument is arguably irreversible but presupposes a mistaken view about what follows from a naturalistic view such as Parfit's reductionism. And even if Johnston is right that persons have moral status only if personites do,

Footnote 49 continued

that the best view is that Tweedle survives as both Lefty and Righty, much like the tiger survives both in Bengal and Sumatra.

a plausible ethical theory is still possible. On all these issues I tend to side with the early Johnston against the later Johnston. What went wrong?

The early Johnston urged minimalism, according to which “we can do better in holding out against various sorts of scepticism and unwarranted revision when we correctly represent ordinary practice as having given no crucial hostages to metaphysical fortune” (1992: p. 590). One of the early Johnston’s examples was that commonsense ethics need not presuppose anything as exotic as libertarian free will or Cartesian egos.⁵⁰ The later Johnston thinks that minimalism is powerless in the case of personites where “the supposed ontology, when taken together with our established and *deeply admirable* ethical commitment to expand the protected circle to beings significantly like us, is at odds with something central to our ethical outlook” (2017: p. 628).

I think that the later Johnston underestimates minimalism. While the idea of intrinsic moral status is established and admirable, it is also easily misinterpreted. We can and should interpret it in a way that does not threaten a multiplication of things with moral status. And once we clearly distinguish the Intrinsicness Argument from the Triviality Argument and the Continuity-Variant Argument, we see that the latter two do not draw on established and admirable ethical commitments. Lastly, commonsense ethics can also be reinterpreted in terms of relation *R*, making it more robust to the threat of personites and continuity-variants.⁵¹

So I am not too worried about the Personite Problem. But I am more worried about what I call ‘the Person Problem’, which is that we seem to lack a good reason to think that the aggregates that are persons are morally relevant in the first place (whether or not some non-maximal person-like aggregates are also morally relevant). Why does it matter, for example, whether a burden and its compensation fall within the life of a single person rather than within some gerrymandered aggregate of person-stages?

A solution to the Personite Problem is likely to be a corollary of the solution to the Person Problem. For example, in a different context Johnston himself asks “What, then, is so good or important or valuable about persons?” and answers that it is “because persons are thinking, reflective beings who thus experience the demand to *live* their lives, to give their lives shape according to their idea of the good, that they deserve a kind of respect that no mere thing, however appealing, does” (2010: p. 269). I am not sure whether that’s correct. But since we seem to experience the demand to live our entire lives rather than their non-maximal or gerrymandered proper parts, what Johnston says suggests a solution to both the Person Problem and the Personite Problem. The Person Problem also does not depend on the special metaphysics of worm theory, unlike the Personite Problem. For example, even if the only things around were immaterial souls (no organisms, no psychological continuers, and so on), we might still intelligibly wonder why it matters whether a later benefit comes to the same soul that suffered an earlier burden.

⁵⁰ See also Johnston (1987, 1989b, 1997).

⁵¹ In my other work, I argue, in a way that would be acceptable to a minimalist, that cases of fission, fusion and superlongevity still have ethically revisionary implications.

7 Conclusion

I conclude that Johnston's latest arguments should lead us to get clearer on the theory of intrinsicness, the meaning of reductionism, and the place of relation R in commonsense ethics. And they should lead us to ask why the aggregates that are persons are morally relevant in the first place. But they should not lead us to revise our ethics or the metaphysics of persistence.

Acknowledgements I would like to thank Ralf Bader, Teru Thomas, Tomi Francis, Elliott Thornley, Farbod Akhlaghi, Michal Masny, Andreas Mogensen, Alex Voorhoeve, and three anonymous reviewers for this journal.

Funding Not applicable.

Declarations

Conflicts of interest Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Burke, M. B. (1994). Dion and Theon: An essentialist solution to an ancient puzzle. *Journal of Philosophy*, 91(3), 129–139. <https://doi.org/10.2307/2940990>
- Burke, M. B. (2003). Is my head a person? In K. Petrus (Ed.), *On human persons* (pp. 107–125). Ontos Verlag.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.
- Dainton, B. (1992). Time and division. *Ratio*, 5(2), 102–128. <https://doi.org/10.1111/j.1467-9329.1992.tb00041.x>
- Francescotti, R. M. (2014). Editor's introduction. In R. M. Francescotti (Ed.), *Companion to intrinsic properties* (pp. 1–16). De Gruyter.
- Garrett, B. (1992). Persons and values. *Philosophical Quarterly*, 42(168), 337–344. <https://doi.org/10.2307/2219684>
- Geach, P. (1980). *Reference and generality: An examination of some medieval and modern theories* (3rd edition, 1st edition 1962). Cornell University Press.
- Hawley, K. (1998). Merricks on whether being conscious is intrinsic. *Mind*, 107(428), 841–843. <https://doi.org/10.1093/mind/107.428.841>
- Hawley, K. (2005). Fission, fusion and intrinsic facts. *Philosophy and Phenomenological Research*, 71(3), 602–621. <https://doi.org/10.1111/j.1933-1592.2005.tb00473.x>
- Heller, M. (2008). The donkey problem. *Philosophical Studies*, 140(1), 83–101. <https://doi.org/10.1007/s11098-008-9227-z>
- Holtug, N. (2010). *Persons, interests, and justice*. Oxford University Press.
- Hudson, H. (2001). *A materialist metaphysics of the human person*. Cornell University Press.

- Van Inwagen, P. (1981). The doctrine of arbitrary undetached parts. *Pacific Philosophical Quarterly*, 62(2), 123–137.
- Johnston, M. (1987). Is there a problem about persistence? *Aristotelian Society Supplementary Volume*, 61(1), 107–135. <https://doi.org/10.1093/aristoteliansupp/61.1.107>
- Johnston, M. (1989a). Fission and the facts. *Philosophical Perspectives*, 3, 369–397. <https://doi.org/10.2307/2214274>
- Johnston, M. (1989b). Relativism and the self. In M. Krausz (Ed.), *Relativism: Interpretation and confrontation* (pp. 441–472). Notre Dame University Press.
- Johnston, M. (1992). Reasons and reductionism. *The Philosophical Review*, 101(3), 589–618. <https://doi.org/10.2307/2186058>
- Johnston, M. (1997). Human concerns without superlative selves. In J. Dancy (Ed.), *Reading Parfit* (pp. 149–179). Blackwell.
- Johnston, M. (2010). *Surviving death*. Princeton University Press.
- Johnston, M. (2016). Personites, maximality and ontological trash. *Philosophical Perspectives*, 30(1), 198–228. <https://doi.org/10.1111/phpe.12085>
- Johnston, M. (2017). The personite problem: Should practical reason be tabled? *Noûs*, 51(3), 617–644. <https://doi.org/10.1111/nous.12159>
- Jones, N. K. (2010). *Too Many Cats: the Problem of the Many and the Metaphysics of Vagueness* (PhD Thesis). Birkbeck, University of London.
- Kaiserman, A. (2019). Stage theory and the personite problem. *Analysis*, 79(2), 215–222. <https://doi.org/10.1093/analys/any074>
- Kripke, S. (1972). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantics of natural language* (pp. 253–335, 763–769). D. Reidel Publishing.
- Lewis, D. (1976). Survival and identity. In A. Oksenberg Rorty (Ed.), *The identities of persons* (pp. 17–40). University of California Press.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61, 343–377. <https://doi.org/10.1080/00048408312341131>
- Lewis, D. (1986). *On the plurality of worlds*. Wiley-Blackwell.
- Lewis, D. (1993). Many, but almost one. In K. Cambell, J. Bacon, & L. Reinhardt (Eds.), *Ontology, causality and mind: Essays on the philosophy of D. M. Armstrong* (pp. 23–38). Cambridge University Press.
- Magidor, O. (2016). Endurantism vs. perdurantism?: A debate reconsidered. *Noûs*, 50(3), 509–532. <https://doi.org/10.1111/nous.12100>
- Merricks, T. (1998). Against the doctrine of microphysical supervenience. *Mind*, 107(425), 59–71. <https://doi.org/10.1093/mind/107.425.59>
- Merricks, T. (2003). Maximality and consciousness. *Philosophy and Phenomenological Research*, 66(1), 150–158. <https://doi.org/10.1111/j.1933-1592.2003.tb00248.x>
- Nozick, R. (1981). *Philosophical explanations*. Harvard University Press.
- Olson, E. T. (1995). Why I have no hands. *Theoria*, 61(2), 182–197. <https://doi.org/10.1111/j.1755-2567.1995.tb00496.x>
- Olson, E. T. (2010). Ethics and the generous ontology. *Theoretical Medicine and Bioethics*, 31(4), 259–270. <https://doi.org/10.1007/s11017-010-9148-7>
- Parfit, D. (1976). Lewis, Perry, and what matters. In A. Oksenberg Rorty (Ed.), *The identities of persons* (pp. 91–107). University of California Press.
- Parfit, D. (1987). *Reasons and persons* (2nd repr. with corrections of 1984 ed.). Clarendon Press.
- Parfit, D. (1995). The unimportance of identity. In H. Harris (Ed.), *Identity* (pp. 13–45). Oxford University Press.
- Parfit, D. (2007). Is personal identity what matters. *The Ammonius Foundation*. http://www.stafforini.com/docs/parfit_-_is_personal_identity_what_matters.pdf
- Puccetti, R. (1977). The great C-fiber myth: A critical note. *Philosophy of Science*, 44, 303–305. <https://doi.org/10.1086/288746>
- Schaffer, J. (2017). The ground between the gaps. *Philosophers' Imprint*, 17. <http://hdl.handle.net/2027/spo.3521354.0017.011>
- Shoemaker, S. (1985). Critical notice of *Reasons and persons* by Derek Parfit. *Mind*, 94(375), 443–453.
- Sidelle, A. (2007). The method of verbal dispute. *Philosophical Topics*, 35(1), 83–113. <https://doi.org/10.5840/philtopics2007351/25>
- Sider, T. (1996). All the world's a stage. *Australasian Journal of Philosophy*, 74(3), 433–453. <https://doi.org/10.1080/00048409612347421>

- Sider, T. (2001a). *Four-dimensionalism: An ontology of persistence and time*. Oxford University Press.
- Sider, T. (2001b). Maximality and intrinsic properties. *Philosophy and Phenomenological Research*, 63(2), 357–364. <https://doi.org/10.1111/j.1933-1592.2001.tb00109.x>
- Sider, T. (2003). Maximality and microphysical supervenience. *Philosophy and Phenomenological Research*, 66(1), 139–149. <https://doi.org/10.1111/j.1933-1592.2003.tb00247.x>
- Singer, P. (2011). *Practical ethics* (3rd ed., 1st edition 1979). Cambridge University Press.
- Sosa, E. (1990). Surviving matters. *Noûs*, 24(2), 297–322. <https://doi.org/10.2307/2215530>
- Sutton, C. S. (2014). Against the maximality principle. *Metaphysica*, 15(2), 381–389. <https://doi.org/10.1515/mp-2014-0023>
- Thomas, T. (2016). *Topics in Population Ethics* (PhD Thesis). University of Oxford. Retrieved from <https://ora.ox.ac.uk/objects/uuid:fa2a09aa-e784-4126-bd4a-0487d3653add>
- Walsh, S. D. (2011). Maximality, duplication, and intrinsic value. *Ratio*, 24(3), 311–325. <https://doi.org/10.1111/j.1467-9329.2011.00502.x>
- Wasserman, R. (2005). Humean supervenience and personal identity. *Philosophical Quarterly*, 55(221), 582–593. <https://doi.org/10.1111/j.0031-8094.2005.00417.x>
- Weatherson, B. (2001). Intrinsic properties and combinatorial principles. *Philosophy and Phenomenological Research*, 63(2), 365–380. <https://doi.org/10.1111/j.1933-1592.2001.tb00110.x>
- Wiggins, D. (1968). On being in the same place at the same time. *Philosophical Review*, 77(1), 90–95. <https://doi.org/10.2307/2183184>
- Williams, J. R. (2013). Part-intrinsicity. *Noûs*, 47(3), 431–452. <https://doi.org/10.1111/j.1468-0068.2011.00837.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.