Parfit's Metaphysics and What Matters in Survival

Eric T. Olson

University of Sheffield

Abstract

Derek Parfit takes the central principle of his discussions of personal identity to be "reductionism": that our existence and persistence are not basic facts, but consist in something else. A number of striking claims, including the famous unimportance of identity, are supposed to follow from it. But they don't follow. The main principle in Parfit's arguments is something far more contentious that is never mentioned: a capacious ontology of material things. But the capacious ontology makes trouble for Parfit: it weakens his claim about the unimportance of identity and undermines his arguments for the stronger claim he intended.

Keywords: Identity, Ontology, Derek Parfit, Personal identity, Reductionism.

1. Reductionism

Derek Parfit's discussions of personal identity have been enormously influential. His most celebrated claim is that our numerical identity over time has no importance in itself: a future person's being me rather than someone else is no reason for me to have any special concern for him.

He takes the central premise of these discussions to be a principle he calls *reductionism*. Nearly all his other claims about personal identity are supposed to follow from it (together with a few ancillary premises too obvious to mention). Reductionism says that our existence and persistence through time are not basic facts, but always consist in something else:

> Our existence consists in the existence of a body, and the occurrence of various interrelated mental processes and events. Our identity over time consists in physical and/or psychological continuity (1999: 218; see also 1984: 210f., 1995: 16, 19).

There are two claims here, one about our existence at a given time and one about our continued existence from one time to another.

The first is that a person's existence consists in the existence of a body or organism and mental events or activities of the right sort (presumably, though this is not actually stated, activities that that organism is engaged in). Parfit

doesn't say what it is for one thing to consist in another, but he clearly means at least that these things are metaphysically sufficient for there to be a person. More strongly, the facts about the organism in some sense *ground* the person's existence: there is a person *because* there is an organism engaged in mental activities.

Though this is put in terms of mental events or activities, it would be better to speak more broadly of mental *properties*. To engage in any mental activity is to have a mental property, and I don't think Parfit meant to deny that a person can exist without engaging in mental activity—in deep sleep, say. So the claim is that a person exists because there is an organism having certain mental properties.[1]

The second claim is that a person persists just if she bears some sort of physical or psychological continuity to herself as she is at other times. And here too the facts about continuity ground the facts about persistence. So:

Reductionism

A person exists at any one time because there is an organism having certain mental properties at that time.

A person persists through time because of physical or psychological continuity.

Or rather a *human* person: reductionism could be true of *us* even if the gods were exempt from it. Parfit infers a number of fascinating claims from reductionism: not only the unimportance of identity, but also that questions about our persistence are "empty", that their answers are uninformative, and that we can arbitrarily choose those answers without fear of getting it wrong.

Ever since I first read Parfit, I have wondered how these claims are supposed to follow from such a modest premise. I know now why I was puzzled: they don't follow. The fascinating claims require a completely different and far more controversial principle. Although he never mentions it, it is presupposed in nearly everything he says about personal identity. Parfit's central metaphysical principle is not reductionism, but a powerful and capacious ontology of material things. Reductionism itself plays at best a minor role in his arguments. But this capacious ontology is trouble for Parfit: it makes his famous claim about the unimportance of identity less striking than it appears, and deprives him of any argument for the stronger claim he intended.

2. Emptiness

I will confine my discussion to reductionism about identity over time: that our persistence consists in physical or psychological continuity. One of the most interesting claims Parfit infers from it is that questions about persistence are empty or verbal or nonsubstantive. Call it the *emptiness claim*. It is interesting both in its own right and because of its role in his arguments for the unimportance of identity.

---

[1] Parfit later denies that a human organism could be a person (2012), and this denial is implicit in his earlier work. It implies that no human organism could actually have the mental properties characteristic of personhood. (This may be why he spoke instead of mental processes.) So the claim must be that a person exists either because an organism has certain *non*mental properties, or because an organism stands in some relation other than having to certain mental properties. I will ignore this messy complication.

Suppose I have an adventure that results in someone who is continuous with me in certain ways. Even if we know all this, we may be uncertain whether this being is me or someone newly created, and thus whether I still exist. The emptiness claim is that this is not something further about me that we don't know. If we know what the adventure consists of and in what ways the resulting person is continuous with me, we know all the facts. The further question of whether he *is* me is empty:

> Our question is not about different possibilities. There is only a single possibility, or course of events. Our question is merely about different possible descriptions of this course of events (Parfit 1995: 23; see also 1999: 218).

In asking whether I continue to exist,

> there is only one fact or outcome that we are considering. Different answers to our question are merely different descriptions of this fact or outcome. This is why without answering this empty question, we can know everything that there is to know (Parfit 1984: 214; see also 242, 258, 260, 277; 2008: 184, 202f.).

There are underlying facts about physical and psychological continuity, but no additional fact about whether I survive. We can describe these underlying facts by saying that I still exist, or by saying that I no longer exist and the resulting person is numerically different from me. Both descriptions can be true. Whether I still exist is, in the pejorative sense, a merely verbal question (Parfit 1995: 25).

Parfit seems to be saying that a question is empty just if different answers to it—apparently incompatible answers such as Yes and No—describe the same fact or state of affairs. So if the question of whether I survive is empty, my surviving and my not surviving are not different states of affairs, but the same state of affairs described differently. It is, he says, the state of affairs we could describe more perspicuously by stating the underlying facts about continuity. They are all the facts there are about what happens.

It follows that in order to know all the facts, we needn't *also* know whether I survive. That's not an additional fact to be known. If we know the relevant underlying facts, the answer to an empty question tells us nothing further.

Parfit compares asking whether I survive the adventure to asking whether nausea is a type of pain (Parfit 1995: 36). Imagine boarding a boat, knowing that you will be seasick but not suffer otherwise. To say that you will be seasick is to assert one state of affairs, while to say that you won't be would assert another, incompatible with the first. They are, as Parfit says, different possibilities. But to ask whether you will be in pain would not be a serious question. Either Yes or No would be an equally good answer. We could describe your queasiness either as pain or as something equally unpleasant but not pain. Your being seasick and in pain is not a different state of affairs from your being seasick but not in pain: it's just the same thing in other words. Different answers to the question—Yes and No—are compatible. And they're uninformative: if we know that you'll be seasick but not suffer otherwise, being told that you will be in pain, or that you won't be, says nothing further about your condition.

Here is a more familiar way of characterizing empty questions (Chalmers 2011). A question is empty when parties giving apparently incompatible answers to it are engaged in a merely verbal dispute. They're not disagreeing about the

facts of the case, but only in some way about how to describe them. They appear to disagree only because they understand some expression occurring in the question differently. They could resolve their dispute by distinguishing those different understandings and thus clarifying the question, or by restating it without using the disputed expression.

So if you say you're going to be in pain on the boat and I say you're not (and we both know the underlying facts), it's because you take the word 'pain' to mean something broader that includes seasickness and I take the word to mean something narrower that excludes it. We both agree that you will be in pain in this broader sense and not in the narrower sense. Once we see this, we'll stop disagreeing. And our apparently incompatible answers can be made consistent by restating them in this more perspicuous way.

It's appropriate to respond to an empty question not by giving a straight answer but by asking for clarification. If you ask whether you'll be in pain on the boat, the best response is something like, "It depends on what you mean by 'pain'. If you mean it in the broader sense, then yes; if you mean it in the narrower sense, then no".

Substantive questions are not like this. If you ask whether you're going to be seasick, it's not usually appropriate to respond by saying, "It depends on what you mean by 'seasick'", as there are not different ways of understanding that word yielding different answers. We cannot resolve disputes over substantive questions simply by restating them in a way that avoids equivocation. Nor can removing the equivocation make apparently incompatible answers to them consistent. Disputes over substantive questions are disagreements about the facts of the case and not about how to describe those facts.

Parfit notes that questions about persistence are not always empty, but only when the underlying facts about continuity are presupposed. Suppose I ask in an ordinary context whether Derek Parfit is still living today. (No one would ordinarily ask whether he still *exists*, but the point is the same.) It would be inappropriate to respond, "It depends on what you mean by 'living'", or by any other word. The question is substantive if any is. What makes it substantive is that I was unaware of Parfit's death in 2017 and the current state of his remains. I asked because I didn't know the underlying facts. It's only when the underlying facts are presupposed that questions are empty. It is, we might say, only *philosophical* questions about persistence that are empty, not ordinary ones. Likewise, whether you will be in pain on the boat is empty only for those who know that you will be nauseous but not suffer otherwise. If you're worried about being injured at sea, the question may be perfectly substantive.

3. Metaphysical Implications

The emptiness claim, then, is that all questions about our persistence are empty or verbal given the underlying facts about continuity. This claim is surprising. To take one of Parfit's examples, imagine that I use what the characters in *Star Trek* call the "transporter". It scans me and records the precise nature and arrangement of my atoms, thereby vaporizing me, then sends this information to a faraway place—Mars, perhaps—where it is used to arrange new atoms just as mine were when I was scanned. This, let us assume, produces someone physically and mentally just as I was then.

Parfit says that whether he is me is an empty question: we can describe the facts just as well by saying that he is me or by saying that he isn't, but rather someone else newly created. Suppose my mother says he's me and my father says he's not, though both know the underlying facts. Then their dispute is merely verbal: it arises only because they disagree about the meaning of a word. But what word? What sort of linguistic clarification would resolve the dispute and remove the apparent inconsistency of their answers? If a good response to the question is, "It depends on what you mean by '__'", what expression should fill the blank?

Parfit doesn't say. You might think it's the word 'is'. (Or the word 'survive': to ask whether I survive something is equivalent to asking whether anything existing afterwards is me.) Whether the man appearing on Mars is me depends on what you mean by 'is'. Once we distinguish all the things it could mean in this context, we will see that some apply and some don't. He "is" me in one sense of the word and not in some other sense. These claims are perfectly consistent. And once my mother and father understand this, they'll stop disagreeing. (Or at least they'll stop disagreeing about whether the man on Mars is me. They may now disagree about something else: about what the word 'is' ordinarily means in such contexts, say. But that would no longer be a verbal dispute.)

What, then, are the different things that the word 'is', flanked by singular terms, could mean? Well, it sometimes expresses numerical and sometimes qualitative identity: exact similarity (or nearly exact). And it could be that my mother means that the man appearing on Mars is exactly similar to me as I was when I was scanned and my father means that we're numerically distinct: we're two and not one. These statements could of course both be true. And if my mother and father realize that this is what they mean, they will stop disagreeing. Whether the man appearing on Mars is me can depend on whether 'is' means qualitative or numerical identity.

But this would not be an account of how questions about *persistence* are empty. For something to persist is for that very thing to continue existing—for a thing existing at later times to be numerically the same as a thing existing earlier.

For questions about persistence to be empty owing to an ambiguity in the word 'is', there must be more than one sort of numerical identity. It must be possible for the man on Mars to be numerically identical with me in one way but numerically distinct in another way. Perhaps he could be the same *person* as me but not the same *organism* or the same *mass of matter*, even though he and I *are* people, organisms, and masses of matter. The question of whether he and I are just plain one, without qualification, might be meaningless. The thought is that there is no such thing as absolute, unqualified numerical identity, but only different sorts of qualified or relative identity: same this or same that. Then disagreements about persistence could be due to disagreement about which qualified identity relation is expressed by the word 'is', making them merely verbal.

The claim that there is no such thing as absolute identity is known as the relative-identity thesis, and has occasionally been defended (Geach 1967). It has far-reaching logical and metaphysical implications (van Inwagen 1988). But Parfit never suggested that it might be his view.

Or maybe there *is* such a relation as absolute numerical identity, but it admits of indeterminacy. This thing and that one could be "sort of" one and "sort of" two. Whether the man appearing on Mars is me or someone else might de-

pend on what degree of numerical identity you mean. Suppose my mother takes the word 'is' in this context to express numerical identity to at least degree 0.4 and my father takes it to express numerical identity to at least degree 0.9. If they agree that we are identical to some degree between the two, they will stop disagreeing when they became aware of this linguistic difference. That the man on Mars is me to at least degree 0.4 would be perfectly consistent with his not being me to degree 0.9.

This would be indeterminacy in the identity relation itself: that the man on Mars is neither definitely me nor definitely not me would lie in the connection among that relation, me, and the man on Mars. It would be "metaphysical" or "*de re*" vagueness. But Parfit appears to deny that the identity relation admits of indeterminacy (Parfit 1971a: 685, 1971b: 11, 1995: 29). And it would fit badly with his claim that the only facts about persistence are those to do with continuity. I should have to be identical with the man on Mars to a certain degree—0.5, say—and that would appear to be something beyond the facts about continuity. It would tell us something further. Even if the question, 'Is the man appearing on Mars me?' were empty, the question, 'To what degree is he me?' could be perfectly substantive.

Another thought is there are not really any people at all. Talk of people is just a loose way of speaking about mental processes and events.[2] Whether it is correct to say that the man appearing on Mars is me would be a matter of how *these* mental processes and events (the ones we call "mine", though of course they would not really be anyone's) relate to certain mental processes and events on Mars. And there are different relations among these events that might make such statements correct: relations of similarity and relations of causal dependence, for instance. Then the question, 'Is the man on Mars me?' could be empty insofar as disagreements about the answer are due entirely to disagreement about which such relations these are (Olson 1997: 149-52).

But this would make Parfit's claim that our persistence has no practical importance rather an anticlimax. That we do not exist would be the headline news; the unimportance of identity would be a subsidiary story on the inside pages. In any event, Parfit explicitly rejects this view (Parfit 1995: 17f.).

Suppose, then, that there really is someone who steps into the transporter and someone who steps out of it on Mars, that the word 'is' in the question 'Is he me?' expresses absolute numerical identity, and that this relation does not admit of indeterminacy. In that case it's hard to see how the appropriate response to the question could be, "It depends on what you mean by 'is'". How could clarifying the meaning of that word resolve the apparent disagreement between those who think the answer is Yes and those who think it's No, or make those answers consistent? If the question is in some way about words, it does not appear to be about the word 'is'.

The only other relevant words are the singular terms 'me' and 'the man appearing on Mars'. It could be that the answer to the question depends on which beings these terms are taken to denote. Perhaps two beings step into the transporter. One of them ends up on Mars; the other, who in some way coincides with him, is destroyed in the process. If the word 'me' refers to the first, then the

---

[2] I have often heard this suggestion in conversations about Parfit.

answer to the question is Yes; if it refers to the second, the answer is No. And that one of these beings survives and the other does not are perfectly consistent. Once those who disagree about the question see this, they may cease to disagree. Their disagreement may be due entirely to disagreement about which being the word 'me' refers to.[3] Questions about persistence could be empty not because of indeterminacy of identity, but because of indeterminacy of *reference*.

The suggestion, then, is that questions about whether a past or future being is me are empty or verbal because there is more than one being that the word 'me' could be taken to denote in the context—more than one "admissible referent" of the term, we might say—and only one of them exists at the relevant past or future time. Each of these beings would have to be intelligent and capable of consciousness, as any admissible referent the word 'me' must have these properties (at those times when it's a person, anyway).

This applies not just to science fiction, but to real cases as well. Because the embryological facts leave it uncertain whether each of us was once a zygote, there must be a being who is now conscious and intelligent and writing these words who was once a zygote, and another such being who was not, but appeared at a later stage of development. What makes the question empty is that our personal pronouns and proper names can be taken to denote either being.

So for any past time such that we can truly say that I began to exist then, there must be an admissible referent of the word 'I'—an intelligent being—who began to exist at that time. Otherwise the question of whether I began then would not be empty, and Parfit says that all questions about our persistence are empty. The same goes for the end of us: for every time in the course of our physical and mental decay when those informed of the underlying facts can truly say that we come to an end, there must be an intelligent being, or at least a being who was once intelligent, who ends at that moment. I can see no other way of accounting for the things Parfit says about empty questions.

4. The Capacious Ontology

For this to be the case, there have to be a lot more intelligent beings than we thought. For any period that someone knowing the underlying facts could take to be my lifetime, there must be an admissible referent of my name who exists only during that period. So there must be one who began at the moment I was conceived, one who began when consciousness first appeared, and one who began at birth. And as none of these events is instantaneous, there must be, for every reasonable candidate for the moment of my conception, a referent who began then. The same will go for the end of me.

And there will be empty questions about dogs, buildings, cities, and other concrete objects as well as well as people. It appears to follow from this proposal that for any period that someone informed of the underlying facts could take to

---

[3] Or the term 'the man appearing on Mars'. One being who appears on Mars was previously on Earth; another was never on Earth, but was created by the machine; and either could be taken to be the referent of the term 'the man appearing on Mars'. Or the term 'person', as its extension comprises just the beings denoted by the personal pronouns in their canonical uses. Or any other term that denotes a person. I will set these complications aside.

be the lifetime of anything whatever without being straightforwardly and factually mistaken, there must be a candidate for being that thing that exists only then.

We might call this view the *capacious ontology* of material things. The most familiar version of it is the ontology of temporal parts, also known as "four-dimensionalism" (Quine 1960: 170f.; Heller 1990: ch. 1; van Inwagen 1990b; Sider 2001). It says that all persisting thing are composed of arbitrary temporal parts. For every period of time when a thing exists, whether short or long, continuous or discrete, there is a temporal part of it that exists only then. A temporal part of something is an object that is just like that object but with a shorter temporal extent. So there is such a thing as my first half: a being physically and mentally just like I am during the first half of my life, which eats and sleeps and thinks about philosophy, but exists at no other time. And there is a temporal part of me that exists from midnight last night till midnight tonight. Given that time is continuous, there are an uncountable infinity of beings now sitting here and writing these words.

The ontology of temporal parts is almost invariably combined with *unrestricted composition*, the principle that any non-overlapping things, whatever their nature or arrangement, make up or compose a larger thing.[4] Together these principles imply that there is a material thing occupying every matter-filled region of spacetime. As Quine once put it, a physical object "comprises simply the content, however heterogeneous, of some portion of space-time, however disconnected and gerrymandered" (Quine 1960: 171). The temporal-parts ontology implies that for every even-numbered day of the calendar, there is a different temporal part of me that exists only during that day. And the principle of unrestricted composition implies that there is a being composed of these day-long parts: a being physically and mentally just like me except that it exists only on even days. There is, likewise, a thing made up of my temporal parts located on even days and yours located on odd ones, which discontinuously changes location each midnight and has a radically discontinuous mental life. And there are far more things—real, concrete objects like you and me—whose spatiotemporal boundaries are too complex to describe.

Whether Parfit accepted the temporal-parts ontology is unclear.[5] But he needs *some* sort of capacious ontology. In fact his claims about reductionism and emptiness require something stronger still. For our persistence to consist in some sort of continuity, it is not enough for this continuity to hold, as matter of contingent fact, throughout our careers—for any being existing in the past or future in fact to be me just if it is then continuous with me in that way. The claim ascribes to us the modal property of *necessarily* persisting just if such continuity

---

[4] Lewis 1986: 212f.; Sider 2001: 120-39; van Inwagen 1990a: 74-79. To say that the *x*s compose *y* is to say that each of the *x*s is a part of *y*, no two of the *x*s share a part, and every part of *y* shares a part with one or more of the *x*s.

[5] He never rejects it, even when discussing Lewis, who endorses it explicitly (Parfit 1976, Lewis 1976). At one point he says it may be compatible with "our ordinary scheme" of three-dimensional entities (Parfit 1999: 255) and suggests that the difference between affirming and denying it is merely verbal. His statement that a temporal part of a person "is just a person during a period" (Parfit 1976: 102, n. 3) supports this interpretation. I think he is mistaken about this (see Sider 2001: ch. 3), but that's a large topic that I can't take up here.

holds. And it's not only questions about persistence in real cases that Parfit sees as empty, but also questions about persistence in counterfactual cases. For good or ill, the *Star Trek* transporter will never exist. But we can ask whether we *could* travel with it if it did exist: whether the person materializing at the destination would be the one vaporized at the starting point. Presumably Parfit considers the question empty, given the underlying facts about continuity. It's not just that if the procedure existed, asking whether someone had survived it *would* be empty. The question we can actually ask—what would happen to the person—is empty already.

This requires more than the existence of a candidate for being me occupying each spacetime region that someone informed of the underlying facts could take to be the one I occupy without being factually mistaken. That's not enough to make it true to say that I could travel by transporter and also true to say that I couldn't. It appears to require further that for every set of modal properties that someone could take me to have without being factually mistaken, there is a candidate for being me—an admissible referent of the word 'I' in my mouth—who actually has those properties. There is one who could survive teleportation and another who couldn't. If no candidate for being me could survive teleportation, it would be straightforwardly and factually false to say that I could and the question would not be empty. The emptiness claim requires a "modally augmented" capacious ontology.

Parfit could avoid this further commitment by adopting a counterpart-theoretic account of modal predication (see Lewis 1971; Olson 2007: 108-114). But he doesn't do so when he has the chance: he says that if a statue and a lump coincide throughout their existence, they must differ in their persistence conditions—modal properties to do with what it's possible for them to survive (Parfit 2008: 194f.).

Unless I have badly misunderstood Parfit's reductionism, neither the emptiness claim nor the capacious ontology (modally augmented or not) follows from it. My persistence could consist in some sort of physical or psychological continuity even if there were only one conscious being writing these words. In that case there would be nothing "empty" about questions to do with my persistence through time, such as whether, given the facts about continuity, I was ever a zygote. I should either have been a zygote or not (or it might be indeterminate whether the identity relation holds between me and a certain zygote). If there were no candidates for being me with a different history, then just one answer to the question—Yes or No (or Indeterminate)—would be factually true, however difficult it may be to know which it is.

Likewise (counterpart theory aside), I should either have or lack the property of possibly travelling by transporter. If there were no candidates for being me with different modal properties, there would be nothing empty about this question either. Just one answer would be true, and any other would be factually false. We could not respond, "It depends on what you mean by 'I'", as no admissible interpretation of that pronoun would yield a different answer.

5. How Parfit Presupposes the Capacious Ontology

It's not just the emptiness claim that presupposes the capacious ontology. Several other views of Parfit's are unsurprising given the capacious ontology but baffling otherwise.

One is his claim that we can arbitrarily choose how to answer a question about persistence without risk of getting it wrong. ("When an empty question has no answer, we can decide to *give* it an answer" [Parfit 1984: 214; see also 240, 260].) We can say that the person stepping out of the transporter is me and I've survived, or we can say that he isn't me and I've perished. It's up to us.[6] But how can whether someone has survived or perished be a linguistic decision? The capacious ontology implies that there are many different admissible referents of the term 'me' or 'I' in my mouth, some who survive the adventure and some who don't. Which one the terms denote is a fact about our language, and this really is, to some extent anyway, up to us as a linguistic community to decide. We can give the question of whether I survive an answer in the sense of fixing that reference to a being (or beings) who survives, making the answer Yes, or by fixing it to a being who perishes, making the answer No.

Or consider what Parfit says about "day-persons" and "series-persons". Day-persons are just like you and me except that they can exist only during a single uninterrupted period of consciousness (Parfit 1984: 292; see also 1971a: 686). If a day-person nods off, that's the end of her. Parfit says, without giving any argument, that there actually are day-persons. At any moment there are as many as there are conscious people.

He introduces the idea of a series-person by imagining a scenario where everyone over the age of 30 has an annual scanning-and-replication procedure (Parfit 1984: 290). This vaporizes the person, which Parfit takes (at least for the sake of argument) to destroy her permanently, and creates a replica physically and mentally just like her but without the effects of ageing, illness, and injury that have accrued during the preceding year. Series-persons would be just like people except that they could survive this procedure. So nothing could be both a person and a series-person. A series-person might coincide with a succession of people, one who lives for thirty years and several who live for just one year each. Parfit says, again without argument, that in this scenario there would be series-persons. In fact there actually are series-persons—as many as there are people.[7] Because the scanning-and-replication procedure does not exist, series-persons have the same careers that people have. But they *could* diverge.

The obvious question about these claims is why anyone should take the concepts *day-person* and *series-person* to apply to anything. I know that I exist, and that I am a person—a being with certain special mental properties that can survive temporary periods of unconsciousness but not being vaporized. (I may be wrong about that, but Parfit is not suggesting that I am.) But why suppose that there is *another* being now thinking my thoughts who exists for only a day and cannot survive unconsciousness? Or a third who has the same career as me but can survive being vaporized? I cannot see what answer Parfit could give other than to appeal to the capacious ontology. He must say that there are day-persons and series-persons because every matter-filled region of spacetime contains a material thing—and not just one, but many, with different modal properties.

---

[6] Ayer (1963: 127), Nozick (1981: 69), Johnston (1989), Unger (1990: 239), and others make similar claims. For further references and critical discussion see Olson 1997, Merricks 2001.

[7] "The concept of a series-person applies just as often as the concept of a person" (1984: 291). Shoemaker tells much the same story and makes a similar claim (Shoemaker 1984: 109f.; see also Olson 2001).

The capacious ontology plays a central role in Parfit's reasoning. And it is no principle of logic, but a substantive and highly contentious metaphysical claim.[8] So it may seem surprising that he never mentions it. At one point he comes close. He says, "If the concept of an X is coherent, and the conditions for its application are fulfilled, we cannot deny that there are Xs" (Parfit 2008: 197). As an example he defines a "squigillion" as "the combination of any bullfrog and the planet Jupiter". He infers that if bullfrogs and Jupiter exist, squigillions must exist too. (His claim that there are day-persons and series-persons appears to be based on similar thinking.) This looks like an assertion of some sort of capacious ontology.

In fact Parfit seems to have taken this to be an *argument* for the existence of squigillions. But there is no real argument here. His definition of 'squigillion' appears to mean 'something *composed* of a bullfrog and Jupiter'. If so, the conditions for applying the concept *squigillion* are not fulfilled by the mere existence of a bullfrog and the planet Jupiter. The existence of something composed of two things x and y, and thus numerically distinct from either, does not follow by any principle of logic from premises about the existence, nature, and arrangement of x and y. The inference requires a further principle about the conditions under which things compose something. And such principles are precisely what metaphysical debates over the capacious ontology are about.

In any event, these remarks of Parfit's are brief and isolated. He clearly did not think the capacious ontology required any defense. I can only surmise that it never occurred to him that anyone might question it. Somehow he did not see the debates over the metaphysics of material things that fill the library shelves as calling the capacious ontology into question. Probably he never consciously formulated it in his mind. That made it invisible and prevented him from seeing his reliance on it. My aim, however, is not to criticize Parfit for this, but to show how the capacious ontology figures in his thinking.

6. The Argument from Below

I turn now to Parfit's claim that numerical identity over time has no rational or moral importance. We usually suppose that if something is going to happen to me, the fact that it will happen to *me* rather than someone else gives me a reason to care about it—a selfish or prudential reason. Even someone who would not lift a finger to save his own mother from unspeakable agony has a reason to save *himself* from it. I have a "special anticipatory or prudential concern" for my future welfare and mine alone (Parfit 1995: 28). Not to care about it would be irrational. Likewise, I have a selfish reason to continue existing for as long as living is good.

Parfit rejects this traditional view. The fact that the person having a future experience is me, he says, is no reason for me to care about it, or at least no more reason than I have to care about what happens to anyone else. What gives me a special reason to care about someone's welfare is not that he *is* me, but rather that he is in some way psychologically continuous with me. Likewise, what I have reason to want in wanting to survive is not that I myself continue existing, but that someone exists in the future who is psychologically continuous

---

[8] See e.g. Lewis 1986: 202-204, van Inwagen 1990a, Sider 2001, Olson 2007: Chs. 3 and 5.

with me, whether or not he is me. What matters practically in questions of persistence—*what matters in survival* for short—is not identity itself, but continuity.

Parfit says this follows from reductionism, perhaps together with some uncontentious ancillary premises (Parfit 1984: 217). It doesn't. He gives two arguments for the unimportance of identity. The central premise in the first—the "argument from below"—is not reductionism, but the emptiness claim. The second, based on fission cases, does not require either reductionism or emptiness. I discuss it in the next section.

The argument from below says that if our persistence consists in a sort of continuity, as reductionism says, then it must be this continuity that matters and not identity (Parfit 1995: 29). This is not because whenever one fact consists in another, it can only ever be the second fact that matters practically and not the first (Parfit 1995: 32f.). Our being conscious might consist in our atoms' being arranged in a certain way. It doesn't follow that we have no reason to care about our being conscious, but only about the way our atoms are arranged. We may still have a reason to care about consciousness itself.

What's special about persistence, Parfit says, is not just that it consists in something else, but that this fact about what it consists in is "merely conceptual" (Parfit 1995: 33). That persistence consists in some sort of continuity is not a fact about persistence itself, but about our words or concepts. When there is continuity of the right sort, we describe it by saying that someone has survived. But we could equally have described it by saying the opposite. So the connection between continuity and persistence is not direct, Parfit says, but goes by way of our linguistic decisions. This is different from the connection between atomic configurations and consciousness. That we are conscious just when our atoms are arranged in the right way has nothing to do with linguistic decisions, but is a necessary truth or a law of nature.

Parfit says that if psychological continuity were to matter practically by making it the case that we survive, it would be important only "because it allows a certain word to be applied". That would be "treating language as more important than reality" (Parfit 1995: 32). But "what matters is reality, not how it is described" (Parfit 2008: 203; see also 1984: 265, 271). So if identity were practically important, and it consisted in some sort of continuity, it would be important because of its relation to this continuity. But this relation is nothing more than a matter of how we choose to describe that continuity—by saying that someone has survived or by saying the opposite. And identity could not get its significance from a linguistic decision. Yet it could not get its significance from anything else. So it must be continuity that matters and not identity.

That's the argument from below. Whatever merits it may have, it relies on the emptiness claim: that statements about someone's persistence tell us nothing beyond the underlying facts about continuity. That I survive a certain adventure and that I don't, the argument says, are merely different descriptions of the same facts. That's what makes the connection between persistence and continuity "merely conceptual", and thus why survival cannot have practical significance. But for this to be true, whether I survive must depend on what we mean by 'I'. There must be an admissible referent of that word who survives the adventure and another who doesn't survive; the linguistic decision is about which one we're referring to. Suitably generalized, that gives us the capacious ontology.

Without the capacious ontology, the argument from below is hopeless. Suppose I exist tomorrow. And suppose, as reductionism has it, that this is not a

brute fact, but consists in there being someone continuous in a certain way to-morrow with me as I am now. Suppose further that he has what matters to me in survival: it would be irrational for me not to care, now, about his welfare then. But now suppose, contrary to the capacious ontology, that there is only one admissible referent of the word 'I' in my mouth. Then the fact that he exists tomorrow has nothing to do with linguistic decisions. And in that case his being continuous with me in the ways that he is might matter to me precisely by mak-ing it the case that he *is* me. This would not require any linguistic decision to be practically significant. The connection between his being continuous with me and his being me would be a metaphysical fact about what it takes for me to persist. It would be like the connection between the arrangement of my atoms and my being conscious.

So the central premise in the argument from below is that the connection be-tween continuity and persistence is just a matter of how we choose to describe things. That doesn't follow from reductionism. It requires the capacious ontology.

7. The Fission Argument

Parfit's other argument for the unimportance of identity is also officially based on reductionism (Parfit 1971b; 1976: 91; 1984: 253-266, 278; 1993: 24-28; 1995: 41-43). Suppose half my brain is destroyed and the other half transplanted into someone whose original brain has been entirely destroyed. The resulting person, with half my brain and a new body, will be psychologically continuous with me to a high degree. (In real life, losing a cerebral hemisphere causes drastic cogni-tive disabilities, so the story requires a certain amount of idealization. But let's not be difficult about this.) Parfit says this person will have what matters to me in survival: I shall have the same reason to be concerned about his welfare after the operation that I have to be concerned about my own welfare.

Now suppose that my brain is divided and each half is transplanted into a different empty head, resulting in two people, Lefty and Righty, each psycholog-ically continuous with me in the same way that the resulting person in the "sin-gle" transplant would be. In this case, Parfit argues, both will have what mat-ters. It could hardly be that just one of them has it, as they relate to me in the same way. If either has it, both do. But the person resulting from the single transplant would have what matters. And the mere fact that a second person is continuous with me in the same way could hardly mean that neither of them has it. So both must.

Yet they cannot both be me. Lefty is distinct from Righty, and one thing cannot be identical with things that are distinct from each other. If Lefty and I are one and Righty and I are one, Lefty and Righty cannot be two: that's an el-ementary fact about the numerical concepts *one* and *two*. Yet Lefty and Righty clearly *are* two. So I can't be both. Parfit says I am neither, as nothing could make me one and not the other.

So both Lefty and Righty have what matters to me in survival, yet neither is me. It follows that someone else can have what matters to me. The relations of rational and moral concern that in real life we bear only to ourselves can hold between one person and another. What matters in this case, Parfit says, is "psy-chological connectedness and/or continuity" (Parfit 1984: 215). And he infers from this that even in cases where there *is* identity, only psychological continui-ty matters.

What role does reductionism play in this argument? Parfit says it requires our persistence to consist in a sort of continuity, as reductionism says. Specifically, it consists in a sort of that I bear to the person resulting from the single transplant, so that he is me and I survive, and to both Lefty and Righty in the double transplant. That prevents me from being Lefty and not Righty or vice versa, as both are continuous with me in the same way. And it means that what prevents me from being Lefty is the mere existence of Righty. More generally, anyone who is appropriately continuous with me will be me, unless this continuity takes a branching form where two people bear it to me. Yet the existence of a second person continuous with me could not prevent the first from having what matters: how could my reason for caring about someone's welfare be entirely cancelled out by the mere existence of someone else (Parfit 1976: 101; 1984: 263; 1995: 43)?

But the argument does not actually require reductionism. There are two main premises. First, the resulting people in the double transplant are not both me. This (given that there are two of them) is a principle of logic and nothing to do with reductionism. Reductionism may rule out my being Lefty rather than Righty or vice versa, but the argument does not require that claim. It would make no difference if I *were* Lefty and not Righty in the double transplant: it would still result in a person numerically distinct from me who has what matters to me in survival.

The second premise is that each resulting person has what matters to me. Why suppose that? Well, because they're strongly psychologically continuous with me. Each has apparent memories, beliefs, plans, preferences, and personality more or less identical to mine, and will be convinced, at least until the situation is explained to him, that he is me. If that were not a reason to suppose that they have what matters to me, the argument would have no force. But this doesn't require reductionism either. Adding the reductionist premise that Lefty *would* be me were it not for the existence of Righty and vice versa[9] provides no further reason to suppose that Lefty has what matters. And given that both do have it, it follows (given that they're not both me) that someone other than me can have it.

Consider a non-reductionist: someone who denies that our persistence consists in any sort of continuity. She might think it consists in the survival of an immaterial substance, or that it's a brute fact not consisting in anything else (Merricks 1998). She will have the same reason that anyone else has to suppose that Lefty and Righty would have what matters to me, and thus that you can have what matters without identity. Again, the reductionist premise that what prevents someone from surviving fission is merely that it produces two people rather than one makes the argument more striking, and has no doubt boosted its popularity. But it contributes nothing to the argument's rational force.

At one point Parfit says the fission argument was not meant to assume that psychological continuity is what matters (Parfit 1993: 24). One version of it, anyway, says that identity cannot be what matters because it is a one-one relation and what matters is not. Two people can have what matters to me in survival, but I cannot be numerically identical with two people. But his reason for saying

---

[9] Of course, Lefty and I cannot be contingently identical or distinct: no distinct things could have been one, and no identical things could have been two. My statements can be rewritten to accord with this, but I wanted to smooth the reader's path.

that what matters can be many-one is that both Lefty and Righty would have what matters. And why suppose that, unless psychological continuity is at least sufficient for mattering?

8. Why Parfit Should Give Up the Capacious Ontology

I have argued that the unimportance of identity does not follow from Parfit's reductionism. The argument from below requires the emptiness claim and the capacious ontology as well. And reductionism makes no substantial contribution to the fission argument.

But the trouble goes deeper. The capacious ontology actually undermines the unimportance of identity, by implying that in every case where we have what matters in survival, there *is* identity. It may be that a future person can have what matters to me practically without being me. But the capacious ontology implies that there will always be an intelligent being coinciding both with me now and with that future person then. *Someone* always survives, even if it isn't me.

Suppose Parfit is right in saying that although I cannot survive a double transplant, both resulting people, Lefty and Righty, will have what matters to me. The capacious ontology says that every matter-filled spacetime region is occupied by at least one material thing. So there is someone located where I am before the transplant, thinking my thoughts then, and located where Lefty is afterwards and thinking *his* thoughts then. (Otherwise there would be no reason to believe in the existence of Parfit's series-persons.) And there is likewise someone coinciding with me beforehand and with Righty afterwards. Call them *Left-survivor* and *Right-survivor*. Although I myself don't survive the operation, they do.

Now recall Parfit's view that what matters in survival—the reason for special, prudential concern that we usually have only for ourselves—can come apart from identity. Call it the *independence claim*. We can see now that it comes in two versions.[10] The weaker version says that a future person other than me can have what matters to me in survival. I can get what I want in wanting to survive even if I don't survive. What matters to me in survival can come apart from my own survival:

Weak Independence
Someone can have a reason for special prudential concern, at an earlier time, for the existence and welfare at a later time of someone else.

The stronger version says that a future person can have what matters to me in survival even if she doesn't exist at all now. I could l get what I want in wanting to survive even if no one survives. What matters to me in survival can come apart from *anyone's* survival, not just my own:

Strong Independence
Someone can have a reason for special prudential concern, at an earlier time, for the existence and welfare of someone at a later time even if no one exists at both times.

The fission argument supports only the weak independence claim. In fact the capacious ontology provides a much simpler argument for that claim. Recall

---

[10] Parfit himself makes this distinction in his debate with Lewis (Parfit 1976: 94).

Parfit's "day-persons": beings just like us except that they can exist only during a continuous period of consciousness. The capacious ontology implies that they actually exist. The day-person coinciding with me tomorrow has what matters to me now in survival if Lefty and Righty do in the fission story (he's strongly psychologically continuous with me); yet he's not me. If what matters is some sort of psychological continuity, the weak independence claim is a trivial corollary of the capacious ontology. Cases where someone else has what matters to me in survival do not occur only in science fiction, but are utterly commonplace.

But the capacious ontology actually rules out the strong independence claim. It implies that whenever a future person has what matters to a person existing now, there is someone coinciding with both the present person and the future person (Left-survivor or Right-survivor in the double transplant). No matter what happens to *me*, as long as someone exists afterwards, someone survives.

It follows that for all Parfit has said, identity might play an essential role in prudential concern. It could be that whenever someone has a special prudential reason to care about someone else's future welfare, this is because someone has a special prudential reason to care about *her own* future welfare. Whenever someone has what matters in survival without herself surviving, it's because someone coinciding with her does survive. All prudential concern is based on numerical identity. Identity over time *is* practically important, even if not in the direct way we might have thought.

Suppose I have a prudential reason to care about Lefty's existence after the operation even though he's not me. According to the capacious ontology, Left-survivor coincides with me before the operation and with Lefty afterwards. Presumably *he* has a reason, before the operation, to care about *his* existence afterwards. For all Parfit says, this might be because it's he himself who will exist then. What matters in his case might be identity. And what gives me a reason to care about Lefty's existence after the operation might be that Left-survivor coincides with me now, sharing my thoughts, and coincides with Lefty afterwards. What matters to me in survival may be grounded in the literal survival of someone else.

Or think of the day-person now coinciding with me—call him D1—and the one coinciding with me tomorrow, D2. By Parfit's lights, D1 has a special prudential reason to care about D2, as D2 will be psychologically continuous with him. But that might give him a reason to care about D2 by making it the case that *I* survive the night and coincide with D2 tomorrow.

You might think this is ruled out by the argument from below. The argument says that whether I survive the night, given the underlying facts about continuity, is only a matter of how we choose to describe those facts. And the mere fact that we describe them by saying that I survive cannot be a prudential reason for anyone to care about the resulting person. The reason must lie in something independent of our linguistic decisions. Parfit says it can only be the facts about continuity.

But given the capacious ontology (which the argument from below requires), the reason for D1's concern could just as easily be the fact that I survive the might. And the truth or falsity of the capacious ontology is in no way due to linguistic decisions. (At any rate Parfit says nothing about how what material things there are could be up to us to decide.) The argument from below shows at best that what matters in survival is either some sort of continuity (rather than

numerical identity), or else the fact that someone or other survives (which *is* a matter of numerical identity). But it doesn't tell us which.

So it could be that although the facts about psychological continuity are practically important and give us a special prudential reason to care about certain future people, they are important by making it the case that someone survives. I haven't argued that this is actually the case. My point is only that Parfit has said nothing against it.

Parfit opposes the traditional view of the importance of identity over time:

Strong Importance of Identity

What matters in survival is always identity. Psychological continuity is practically important because it secures identity. Whenever someone has a special prudential reason to care about someone's future welfare, it's because she is that person.

He says instead that identity has no practical importance at all:

Strong Unimportance of Identity

What matters in survival is never identity, but only some sort of psychological continuity. Whenever someone has a special prudential reason to care about someone's future welfare, it's not because anyone survives, but only because that future person is then psychologically continuous with her.

But given the capacious ontology, Parfit's arguments do not support the strong unimportance claim. In fact it's hard to see how any argument could ever support it: the capacious ontology makes it impossible for there to be a future person having what matters to me in survival in a case where no one survives. A thought experiment separating the two factors that might ground what matters—someone's surviving, on the one hand, and mere psychological continuity on the other—is metaphysically impossible. Even the argument from below cannot support the strong unimportance claim: the capacious ontology, which the argument requires, implies that someone always survives, and this fact has nothing to do with our linguistic decisions. And for all the argument says, this might be what secures what matters practically.

So for all we know, identity may still bear on what matters in survival, though in a less direct way than the strong importance claim has it. Parfit's arguments are consistent with this:

Weak Unimportance of Identity

What matters in survival is always identity. Psychological continuity is practically important because it secures identity. Whenever someone has a special prudential reason to care about someone's future welfare, it's either because she is that person and thus survives, or because someone coinciding with her survives. But it's always because someone survives.

Whatever the significance of this claim, it's not what we were promised. It lacks the elegant simplicity of the strong unimportance claim. Strong unimportance is a radical challenge to our ordinary thinking about value. The weaker claim is much less interesting. The most likely reaction to it is simply bafflement. It's unlikely to change our thinking. Or if it does, that's because it presupposes the capacious ontology. The capacious ontology really is a challenge, and its consequences for practical ethics are poorly understood (Olson 2010). Most people, on the first encounter, find it outrageous.

Parfit never considered the weak unimportance claim. He did not distinguish it from strong unimportance or see that his arguments, by presupposing the capacious ontology, support only the weaker claim. I think the reason is that he never consciously formulated the capacious ontology. That made it hard for him to see its consequences.

But we *could* argue for the strong unimportance claim if we gave up the capacious ontology. Without it, there is no reason to suppose that anyone survives in the fission case[11]—that there are any such beings as Left-survivor or Right-survivor. In that case, Lefty's and Righty's having what matters to me in survival could not be due to anyone's surviving. It could only be due to facts about continuity. In this case at least, identity over time would have no practical importance at all, direct or indirect, just as Parfit claims.

The capacious ontology leaves us with no reason to accept Parfit's celebrated claim about the unimportance of identity. The only way to argue for that claim is to reject the capacious ontology. Despite the central role it plays in his thinking, he'd be better off without it.[12]

References

Ayer, A. J. 1963, *The Concept of a Person and Other Essays*, London: Macmillan.

Chalmers, D.J. 2011, "Verbal Disputes", *Philosophical Review*, 120, 515-66.

Geach, P.T. 1967, "Identity", *Review of Metaphysics*, 21, 3-12.

Heller, M. 1990, *The Ontology of Physical Objects*, Cambridge: Cambridge University Press.

Johnston, M. 1989, "Human Beings", *Journal of Philosophy*, 84, 59-83.

Lewis, D. 1971, "Counterparts of Persons and their Bodies", *Journal of Philosophy*, 68, 203-11.

Lewis, D. 1976, "Survival and Identity", in Rorty, A. (ed.), *The Identities of Persons*, Berkeley: University of California Press.

Lewis, D. 1986, *On the Plurality of Worlds*, Oxford: Blackwell.

Merricks, T. 1998, "There are No Criteria of Identity Over Time", *Noûs*, 32, 106-24.

Merricks, T. 2001, "Realism About Personal Identity Over Time", *Philosophical Perspectives*, 15, 173-87.

Nozick, R. 1981, *Philosophical Explanations*, Cambridge, MA: Harvard University Press.

Olson, E.T. 1997, "Relativism and Persistence", *Philosophical Studies*, 88, 141-62.

Olson, E.T. 2001, "Personal Identity and the Radiation Argument", *Analysis*, 61, 38-43.

[11] Or rather, that anyone survives and remains psychologically continuous with herself as she was previously. If people are animals, three people might survive the operation: the donor, who continues existing with an empty head, and the two who receive the transplants.

[12] This paper is a distant descendant of a talk given at the American Philosophical Association, Central Division in 2014 to mark the 30th anniversary of Parfit's *Reasons and Persons*. I thank those involved in the session. I am especially grateful to Jens Johansson, Todd Karhu, Eli Pitcovski, and Karsten Witt for comments on more recent versions.

Olson, E.T. 2007, *What Are We?*, New York: Oxford University Press.

Olson, E.T. 2010, "Ethics and the Generous Ontology", *Theoretical Medicine and Bioethics*, 31, 259-70.

Parfit, D. 1971a, "On 'The Importance of Self-Identity'", *Journal of Philosophy*, 68, 683-90.

Parfit, D. 1971b, "Personal Identity", *Philosophical Review*, 80, 3-27.

Parfit, D. 1976, "Lewis, Perry, and What Matters", in Rorty, A. (ed.), *The Identities of Persons*, Berkeley: University of California Press.

Parfit, D. 1984, *Reasons and Persons*, Oxford: Oxford University Press.

Parfit, D. 1993, "The Indeterminacy of Identity: A Reply to Brueckner", *Philosophical Studies*, 70, 23-33.

Parfit, D. 1995, "The Unimportance of Identity", in Harris, J. (ed.), *Identity*, Oxford: Oxford University Press.

Parfit, D. 1999, "Experiences, Subjects, and Conceptual Schemes", *Philosophical Topics*, 26, 217-70.

Parfit, D. 2008, "Persons, Bodies, and Human Beings", in Sider, T., Hawthorne, J. and Zimmerman, D. (eds.), *Contemporary Debates in Metaphysics*, Oxford: Blackwell.

Parfit, D. 2012, "We Are Not Human Beings", *Philosophy*, 87, 5-28.

Quine, W.V.O. 1960, *Word and Object*, Cambridge, MA: MIT Press.

Sider, T. 2001, *Four-Dimensionalism*, Oxford: Oxford University Press.

Unger, P. 1990, *Identity, Consciousness, and Value*, New York: Oxford University Press.

van Inwagen, P. 1988, "And Yet There are not Three Gods but One God", in Morris, T. (ed.), *Philosophy and the Christian Faith*, University of Notre Dame Press. (Reprinted in his *God, Knowledge, and Mystery*, Ithaca: Cornell University Press 1995.)

van Inwagen, P. 1990a, *Material Beings*, Ithaca: Cornell University Press.

van Inwagen, P. 1990b, "Four Dimensional Objects", *Noûs*, 24, 245-55.