Australasian Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rajp20

The Ethical Implications of Panpsychism

Joseph Gottlieb & Bob Fischer

To cite this article: Joseph Gottlieb & Bob Fischer (08 May 2024): The Ethical Implications of Panpsychism, Australasian Journal of Philosophy, DOI: 10.1080/00048402.2024.2350708

To link to this article: https://doi.org/10.1080/00048402.2024.2350708

AUSTRALASIAN JOURNAL OF PHILOSOPHY
https://doi.org/10.1080/00048402.2024.2350708

The Ethical Implications of Panpsychism

Joseph Gottlieb and Bob Fischer

Texas Tech University; Texas State University

ABSTRACT

The history of philosophy is a history of moral circle expansion. This history correlates with a history of expansionism about consciousness. Recently, expansionism about consciousness has exploded: to invertebrates, to plants, to logic gates, and to fundamental entities. The last of these expansions stems from a surge of interest in panpsychism. In an exploratory spirit, this paper considers some largely uncharted territory: the ethical implications of panpsychism. Our conclusion is that while panpsychism probably does significantly expand our moral circle, it's also probably short on anything of practical significance.

ARTICLE HISTORY Received 16 May 2022; Revised 2 January 2023

KEYWORDS consciousness; panpsychism; Russellian panpsychism; moral standing; welfare subjects; valence

1. Two Expansions

According to one influential philosophical narrative, the last three centuries have involved the steady expansion of the moral circle, where to be 'in' the circle is to be recognised as having moral standing.[1] We once overestimated the properties required for moral standing and underestimated the individuals who possess them; as a result, our moral circle was too small (Sebo, forthcoming). Now, we appreciate that it needs to be much larger, including not just all humans but many kinds of nonhumans.

This expansion correlates with another: namely, an expansion of attributions of phenomenal consciousness.[2] The correlation is not surprising, as consciousness matters morally.[3] There is an enormous difference between burning an old newspaper for fire and burning an old friend: however we account for the difference, consciousness is likely to be part of the story.[4] If, say, honey bees are conscious, then in one important respect, they are more like old friends than old newspapers—and our moral circle should adjust accordingly.

[1] In axiological terms, we could instead say: 'recognised as having intrinsic value'.
[2] Henceforth, save for when needed for clarity, the 'phenomenal' qualifier will be dropped.
[3] Granted, the correlation is imperfect. Some expansions of the moral circle have been based on accounts of moral status that are not rooted in consciousness. Biocentric views have pushed for moral expansion based on the property of being alive, not consciousness *per se*.
[4] On the other hand, there has also been some movement towards contractionism about consciousness. For example, on some higher-order theories, consciousness requires a kind of conceptual sophistication, exercised in an ability to token a conceptually thick higher-order thoughts or 'HOTs'. However, the extent to which HOT theory is contractionist is a matter of controversy; see, e.g., Rosenthal 2005, and Gennaro 2012. More extreme is illusionism, which says that nothing is conscious (e.g., see Kammerer 2019).

In recent years, this second expansion has become an explosion: to invertebrates, to plants, to logic gates, and beyond.[5] This paper is about the limit case (or a case close to the limits) of this second expansion: *panpsychism*, or the thesis that consciousness is fundamental and ubiquitous. All panpsychists are expansionist to some degree—they will all claim that consciousness extends beyond the usual suspects—but they disagree over *the proper extent* of expansionism. Some panpsychists insist that everything is conscious, from water molecules to mountain ranges (Roelofs 2019). We set such views aside. Instead, we are interested in forms of panpsychism whose distinctive expansionist elements concern fundamental entities—specifically, *Russellian panpsychism* (RP). And we will focus on 'smallist' (Coleman 2006) versions of RP, which takes particles to be fundamental (as opposed to the universe as a whole), taking electrons as our paradigmatic fundamental particle. While the thesis that there's consciousness at the fundamental level is true on most if not all versions of panpsychism, (smallist) RP is arguably the most popular form of panpsychism. On top of that, it's an increasingly popular view in the metaphysics of consciousness more generally. So, while our analysis of the ethical implications of panpsychism is not exhaustive—it leaves out, for example, the implications of cosmopsychism (Goff 2017)—it takes up a prominent view with features shared by many other views.[6]

This brings us to our main question, namely, what are RP's ethical implications? The early animal ethics literature focused squarely on (i) whether animals matter, (ii) how much they matter, and (iii) what we should do in light of the answers. At the same time, people considered (iv) whether the idea of having duties to animals was a bridge too far: was that reason enough to reject the idea of our having such responsibilities? Those four problems, it turned out, were more than enough at that stage of inquiry. Similarly, we're in the very early days of the literature on the ethics of RP.[7] So, we focus on whether electrons matter, how much they matter, and what we should do in light of the answers to these questions. Our tentative answers are that:

(1) If RP is true, then it's plausible that electrons have moral standing.
(2) If electrons have moral standing, then it is not easy to avoid the conclusion that individual electrons matter to a non-trivial degree.
(3) Despite (1) and (2), it isn't clear that panpsychism has any practical consequences.
(4) So, given (3), proponents of RP can reasonably accept (1) and (2).

To defend these claims, we need to first get a better grip on RP.

[5] On insects, see Barron and Klein 2016; on plants, see Maher 2017; on simple physical systems, see Tononi 2012 and Koch 2015; on AI, see Brunet and Halina 2020.

[6] Though see fn. 20 for more on cosmopsychism. Another place of interest that is set aside here is the space *between* electrons and organisms. If the panpsychist had combination principles that yielded conscious entities sized above electrons but below organisms (e.g., see Roelofs 2019), then that too may raise unique ethical questions. Many thanks to an anonymous referee for raising this possibility.

[7] Crummett (2022) argues that typical adult humans contain many 'morally relevant non-agential subjects' if just one of several various metaphysical theses are true, including panpsychism. Roelofs and Buchanan (2019: 3012–13) briefly discuss the moral implications of universalism, which says that (literally) everything is conscious. Tomasik (2014) discusses suffering in fundamental physics but approaches the issue quite differently than we do. Goff (2019, ch. 5) explores questions related to 'consciousness and the meaning of life' given his form of panpsychism, but they aren't all explicitly ethical (e.g., whether we have free will), and the ones that are (e.g., what grounds ethical objectivity) are different from those taken up here. For work on the ethical implications of panpsychism in environmental ethics, see, e.g., Mathews 2003 and Skrbina 2013.

2. Russellian Panpsychism

RP's popularity is partly due to its purported ability to thread the needle between physicalism and dualism. This is Chalmers's (2015) *Hegelian argument*: the synthesis of the causal argument for physicalism (thesis) and the conceivability argument for dualism (antithesis) and anti-reductionism, yields RP. The latter point turns on zombie-style cases, which are supposed to show that consciousness isn't *a priori* derivable and so can't be reduced to the narrowly physical (Chalmers 2015). How and whether RP can make good on the Hegelian synthesis doesn't matter here; what matters is that doing so is a motivation.

A version of panpsychism is 'Russellian' in so far as it assumes *structuralism* about physics, according to which physics tells us about the relational nature of matter, not about its intrinsic, categorical character. Physics tells us what mass and charge *do*, not what mass and charge *are*. For the Russellian, this intrinsic void is filled by *quiddities*, about which physics is (ostensibly) silent. Quiddities are just those fundamental categorical properties that occupy the relational roles specified by our physical theories. But at least some quiddities are identical to 'micro-phenomenal' properties. Micro-phenomenal properties play microphysical roles (for example, 'the mass role', 'the charge role'), but they're also the properties that *constitute* 'macro-level' consciousness—the kind humans possess. So, if the correct micro-level physics refers to electrons, then RP entails that electrons are subjects of conscious experience.

We should emphasise two points. First, the claim about electrons being subjects of experience is true by definition: all micro-physical entities have micro-phenomenal properties, whatever instantiates micro-phenomenal properties has micro-level experiences, and to be a thing that *has* or *undergoes* micro-experience is to be a subject of experience. Goff (2017: 171), for instance, is explicit about this, saying that the proponents of RP face a unique version of the so-called combination problem, since they think 'the basic building blocks of reality are *subjects*'.

Second, the 'micro' in 'micro-experience' is meant to pick out a kind of experience—that is, the kinds of experience had by fundamental physical entities. That's compatible with micro-experiences being quite unlike macro-experiences. However, it isn't compatible with the view that electrons aren't *really* conscious. So, while 'electrons are conscious' is shorthand for 'electrons are subjects of conscious micro-experience', it remains the case that 'micro-consciousness' is supposed to be a form of consciousness, full stop.

Admittedly, there are versions of RP that reject both the claim that electrons are conscious and that they're micro-conscious, like panprotopsychist variants of RP. Here, fundamental physical entities have protophenomenal properties instead of phenomenal properties, where merely instantiating protophenomenal properties isn't sufficient for consciousness. As Chalmers (2015: 259) puts it, 'protophenomenal properties are special properties that are not phenomenal (there is nothing it is like to have a single protophenomenal property) but that can collectively constitute phenomenal properties, perhaps when arranged in the right structure'. Whatever we think of panprotopsychism, it is not an example of an expansionist take on consciousness, which is what interests us here. So, we can set panprotopsychism aside.[8]

[8] Likewise, we set aside Coleman's (2015) *panqualityism*, which denies that the qualities at the fundamental level—such as the *sharpness* of a particular pain—are experienced.

What if we rejected the constitution relation and went with an emergentist variant of RP (see, for example, Mørch 2014)? This would have some interesting consequences, but not ones implying that electrons aren't conscious. On those views, macro-level subjects are 'fusions' of micro-level subjects—the latter of which cease to exist when they come to compose a macro-subject. Still, in the absence of fusion, micro-level subjects are subjects, not pseudo-subjects.

Going forward, then, we'll assume some version of RP that implies that electrons are conscious. What does this tell us? Perhaps surprisingly, not much—or not much on its own. To make any progress, we need answers to (at least) three other questions:

Standing Is consciousness (as such or of some kind) sufficient for moral standing?

Status If electrons have moral standing, then how much do they matter, morally speaking, relative to other individuals with moral standing?

Reasons Given particular answers to Standing and Status, would we have moral reason to act differently than we do now?

In other words, Standing asks whether electrons matter at all. Given that they matter, Status asks how much they matter. And for combinations of answers to Standing and Status, Reasons asks whether someone should act differently as a result. Let's take up these questions in turn.

3. Standing

We can't fully answer Standing in a paper that has other aims; it deserves more attention than we can give it here. We can make our job a bit easier, though, by helping ourselves to an assumption: namely, that all *welfare subjects*—all beings who can be made better and worse off—have moral standing. Then, we can ask whether, given RP, electrons meet a sufficient condition for welfare subjecthood.

Here's one: the *phenomenal theory* (PT) of welfare subjecthood. On PT, if *S* has the capacity for phenomenal consciousness, *S* is a welfare subject.[9] If RP is true, electrons have this capacity. So, if PT and RP are true, electrons are welfare subjects—which is an answer to Standing.

PT is *prima facie* plausible, but one might object as follows. Consider Grey. Unlike other conscious subjects, Grey is limited in the kinds of experiences he can have. Indeed, he can only have one kind: a visual experience of a grey, static, and uniform field. Perhaps Grey is not a welfare subject. If so, PT is false.

There are a few things to say here. One reason to think Grey isn't a welfare subject is if Grey can't accrue any welfare goods and bads. If we stipulate that an experience of a grey, uniform field is neither pleasant nor unpleasant *and* that this is the only kind of experience Grey can have, then it's unclear how Grey can be better or worse off.

Or so it might seem. Suppose that we adopt *deflationism* about welfare subjects, where our theory of welfare is prior to our theory of welfare subjects (van der Deijl 2021; Lin 2021).[10] And suppose we couple deflationism with an objective list theory of

[9] The term 'phenomenal theory' comes from Lee (ms), although he makes it a necessary and sufficient condition. We only need the latter. Other PT-sympathisers include Siewert (1998), Lin (2021), and van der Deijl (2021).

[10] Lee (ms) rejects deflationism. But he would say that Grey would be like his 'Zero': an entity that cannot accrue welfare goods and bads, but is a welfare subject all the same, because they have a welfare level zero, not no welfare level at all. We set aside the cogency of a welfare level zero.

welfare. Then, welfare subjects are entities that can have knowledge, friendship, pleasure, and whatever else is on the list of objective goods. Take knowledge. It's been argued, perhaps as early as Russell but also more recently (Duncan 2021), that some forms of knowledge are *constituted* by experience.[11] If that's right, then Grey is a welfare subject because he can accrue a welfare good. The same thing could be said of electrons: they accrue welfare in virtue of having knowledge-constituting conscious experiences. Electrons, then, would be like Chalmers's (2022) philosophical Vulcans: conscious entities with moral standing, but lacking in the capacity to have pleasure or pain.[12]

The general point here is we can't separate our assessment of PT from broader questions about the correct theory of welfare and, correspondingly, the correct theory of candidate welfare goods and bads. While there are certainly combinations of views that rule out Grey as a welfare subject, there are combinations that rule Grey in. Moreover, we readily grant that if electrons *are* just like Grey in having just this one kind of experience, their lives would be quite impoverished. Yet that's neither here nor there when asking about welfare subjecthood. That would speak, instead, to electron's *capacity* for welfare. As we'll see, that's relevant to Status and (to some extent) Reasons, not Standing.

With that said, we see no reason to think that electrons *are* like Grey. RP alone doesn't entail that electrons only have one kind of experience. Electrons' stock of experiences is fixed by the stock of experiential primitives, about which panpsychists disagree.

Perhaps the real worry, though, is that only *the right kind* of consciousness is sufficient for welfare subjecthood. Most obviously, someone might insist that all and only subjects *with valenced mental states* are welfare subjects, such as pleasures and pains—that is, the *Valenced State Theory* (VST).[13] Grey would thus not be a welfare subject because experiences of uniform grey fields aren't valenced.

This response leads us to two questions. The first is whether we ought to accept PT or VST (or neither). The second is whether, conditional on RP, we ought to think that electrons have valenced experiences. After all, that electrons are conscious doesn't entail that they instantiate every kind of conscious state. Blind humans are conscious but don't have visual experiences. If electrons have valenced experiences, then it doesn't matter whether PT or VST is correct: electrons are welfare subjects. So, let's put things this way: can electrons be in pain?

It's hard to know. Nagel (1974) famously emphasised our inability to grasp what it's like to be a bat. Presumably, the experiential gulf between bats and humans is smaller than the experiential gulf between electrons and humans. Bats perceive via echolocation, but they perceive. There's no clear sense in which electrons do.

Fortunately, RP's motivation can inform our assessment of electron consciousness. When we consider the explanatory burden that micro-subjects and micro-experiences are supposed to bear, it's plausible that electrons would have valenced experiences if RP were true—or, at least, it isn't clear that they'd lack them.

[11] On Duncan's (2021) view, experience is knowledge of things, a non-propositional form of knowledge.
[12] Chalmers stipulates that his Vulcans, while lacking in affect, have desires. While electrons-so-construed would have knowledge, they wouldn't necessarily have desires. For critical discussion of philosophical Vulcans, see Roelofs 2022.
[13] This view is sometimes called *narrow sentientism*. See Roelofs 2022.

To see why, consider several constraints on what micro or bottom-level consciousness could be like from Lewtas (2013).[14] One is that micro-level consciousness should be no more complex than necessary to explain macro-level consciousness. Another is that the explanation shouldn't have any gaps. Panpsychists want the macro-level facts (including facts about macro-level consciousness) to be in-principle derivable *a priori* from the micro-level facts (including facts about micro-level consciousness plus structure). This is part of the ostensible payoff of Chalmers's Hegelian synthesis argument.

Consider how these two constraints might work together. Jane—a macro-level entity—has the visual experience we associate with seeing a sunset. This macro-level experience has a distinct phenomenal character: what it's like for Jane to have this experience is that of being presented with a pinkish, yellowish, striped expanse. Now take a micro-level entity like an electron. By the first constraint, we should avoid attributing a phenomenally-type identical experience to the electron unless our explanatory demands require it. But as Lewtas (2013: 46) notes, given the second constraint, it isn't clear how we can move too far away from the experience we describe as *that of being presented with a pinkish, yellowish, striped expanse* if we want the *a priori* deduction to go through. There will probably be certain experiential facts that only obtain in entities with certain levels of complexity—perhaps entities with brains or their functional equivalents; so, the electron's experience may be pared down from Jane's. Still, the point is that panpsychism itself dictates that it can only be pared down so far (cf., Lewtas 2013: 48–49).

Suppose, then, that Jane falls down a flight of stairs and suffers a compound fracture. In normal, non-anesthetised humans, this would cause an immensely painful experience—an experience with a negative valence. What kind of micro-experience would correspond to this macro-experience? It's unclear. But we do know that *whatever* this micro-experience is like, it must be of the kind from which the existence of macro-experience is *a priori* derivable, perhaps in conjunction with some further structural-functional facts. We aren't sure how this would be possible without saying that micro-level experiences have forms of valence themselves.[15]

Of course, there is debate about whether valenced experiences are reducible. There are reductive theories of macro-level valence—for example, Heathwood (2007) reduces unpleasantness to *de re* desires and Klein (2015) reduces the painfulness of pain to higher-order imperatival states. But since RP is motivated by the purported *failure* of reductive accounts of experience, it's not obvious why proponents of RP would accept a reductive account of valenced experiences.

Relatedly, there's experiential primitives. One tendency in the philosophy of mind is to focus on visual experience (particularly colour experience) when thinking about what the most basic experience could look like. Against this, Shani (2022) argues that affective experiences are most basic. So, when we consider the experiential lives of the simplest entities—including electrons—we should expect them to have valenced experiences if they have experiences at all. If so, then per the VST, electrons are welfare subjects.

Let's take stock. We've found two routes from RP to the conclusion that electrons are welfare subjects: PT and VST. So, if all welfare subjects have moral standing, then RP implies that electrons have moral standing. Granted, there are ways for proponents of RP to resist this conclusion. Still, it seems credible given what motivates RP. This is a

[14] Roelofs (2019) explores similar questions.
[15] This is related to Roelofs's (2014, 2019) 'palette problem'.

striking result. For all that, however, we might conclude that electrons barely matter at all. We turn to that issue now.

4. Status

Standing concerns *whether* electrons matter morally. Status concerns *how much* they matter. So, if electrons have moral standing, then how much do they matter relative to other individuals with moral standing? To make any progress here, we need to map the conceptual terrain.

One answer to Status is: *exactly as much as any other entity with moral standing*. This answer amounts to the denial of differences in moral status, effectively collapsing the distinction between having moral standing and moral status. On this view, either you have moral status or you don't; there aren't higher and lower moral statuses. Kagan (2019) distinguishes between the *unitarian* view of moral status (just described) and *hierarchical* views, which posit differences in how much individuals matter. Peter Singer (1979) is probably the most famous proponent of a unitarian view, expressed neatly in the *principle of equal consideration of interest* (PECI). According to the PECI, if a human and a pig have pain experiences that create equal welfare impacts, the PECI implies that we ought to give equal weight to each individual's interest in pain relief. It doesn't matter that one is a human and the other is a pig: what matters is their welfare, not their species.

Hierarchical views, by contrast, could sort out this case in any number of ways. Strictly speaking, a hierarchical view is just any view that involves differences in moral status. So, one hierarchical view is the position that each unit of pig welfare matters exactly 4.783 times *more* than each unit of human welfare. As far as we know, no one endorses this view, but it's hierarchical all the same.

Most hierarchical views privilege humans. To illustrate, let's assume that pigs and humans are welfare subjects with moral standing. Suppose that if we don't act, 2500 pigs and 5 normal adult humans will all undergo an uninterrupted, 10-minute-long pain experience of comparable intensity. We can help the humans or the pigs but not both. Hierarchical views imply that we can't compare the welfare of the pigs and humans—and so work out whom to prioritise—until we know their relative moral statuses. If one welfare subject has a higher moral status than another, then a 1-unit drop in the former's welfare will make a larger difference to the moral value of an outcome than a 1-unit drop in the latter's welfare. To get the *status-adjusted value* of an individual's welfare, a given quantity of welfare must be discounted based on that individual's moral status (Kagan 2019: 108). Suppose that we can represent humans' moral status as 1 and pigs as 0.001. Suppose further that an intervention will avert the loss of 5 units of welfare per individual. Then, we get the status-adjusted values of our options as follows:

Humans: (5 units of welfare $\times$ 1) $\times$ 5 individuals $=$ 25;

Pigs: (5 units of welfare $\times$ 0.001) $\times$ 2500 individuals $=$ 12.5.

Given these assumptions, it is better to help the humans.

Despite appearances, though, the practical difference between unitarian and hierarchical views is smaller than it seems. This is because unitarians have developed their own way to justify privileging humans in such trade-off cases. Instead of appealing to differences in moral status, they appeal to differences in *capacity for welfare*.

An individual's capacity for welfare is how much welfare that individual can realize in total. An individual's capacity for welfare breaks down into two components: their welfare range, or how much welfare that individual can realize at a time, and their lifespan (Fischer forthcoming). Given the right combination of differences in welfare ranges and lifespans, it could work out that humans can realize far more positive and negative welfare than pigs can—our highs can be higher; our lows, lower; our lifespans, longer. As a result, the calculations above would go differently. There, we applied a discount rate to pig welfare. If we were unitarians who believed in differences in capacity for welfare, though, we wouldn't apply a discount rate, but we could think that there's less pig welfare to begin with. So, we might get:

Humans: (5 units of welfare) $\times$ 5 individuals $=$ 25;

Pigs: (0.005 units of welfare) $\times$ 2500 individuals $=$ 12.5.

Of course, we aren't defending these specific values. The point is just that both unitarians and hierarchicalists have ways of generating exactly the same practical conclusions via different methods. Hierachicalists simply discount some welfare based on whose welfare it is. Unitarians don't discount welfare, but they can attribute differences in capacity for welfare. Granted, these views may have very different merits; likewise, specific versions of these views may be more or less plausible. For present purposes, though, we just need these basic options on the table.

4.1 Electrons given Hierarchicalism

One problem for hierarchical views is to explain *why* beings have different moral statues. The problem is that many answers look like double counting. Kagan (2019: 279–284), for instance, suggests that an individual's moral status is based partly on that individual's capacity for welfare. That's a strange view: it says that because a pig can't realize as much welfare as a human, the welfare the pig *can* realize matters even less. Why would that be? Simply in virtue of the pig's species? It's hard to see how species membership could carry the explanatory burden. But if not in virtue of the pig's species, then what?

One possible answer—which we won't defend and will simply explore—is that individuals differ in the degree to which they're fungible, replaceable, or interchangeable. On this view, we might grant that while two individuals matter equally, there are, nevertheless, different deontic constraints on their treatment. So, for instance, the biocentric egalitarian—who thinks that all living things matter equally—might still say that if a rosemary bush dies, a new rosemary bush can replace it perfectly, whereas if a child dies, a new child doesn't replace the first. You can have another child, but that doesn't make up for the loss of the first. Given this, we could argue that being non-fungible makes some individuals matter differently than others, and in so far as there are different sets of deontic constraints that apply depending on the different ways of mattering, we may be able to preserve many ordinary practical moral judgements.

Suppose that this is correct. Suppose further that humans are maximally non-fungible and electrons are maximally fungible. Then, our reasons to attend to human interests may be so pervasive that electrons' interests can be ignored for all practical purposes.

Plainly, it would take some argumentation to defend the claim that (non-)fungibility works as we've suggested. We can't pursue that here. Instead, let's briefly consider some candidate bases for non-fungibility. If we can find bases that humans don't share with electrons, then it's worth trying to do the work of developing the views. But if we can't, then the fungibility story probably isn't worth exploring in any detail.

Perhaps the most plausible basis for non-fungibility is consciousness: we have distinct vantage points on the world and, for that reason, we're marked off as 'individuals' in some morally significant sense of that term. But given RP, electrons are conscious too. So, that basis doesn't offer a way to get differences in moral status.

An alternative is to say that *self*-consciousness is the basis for non-fungibility, but that's no better. First, this view probably implies that infants are fungible, as they aren't yet self-conscious. That conclusion is hard to accept.

Second, given RP, electrons could be self-conscious too. Proponents of RP will tell us that micro-level entities like electrons possess the simplest form of phenomenal consciousness possible. But the whole question is what the 'simplest form' looks like. And according to many theorists, self-consciousness is a *constitutive* feature of phenomenal consciousness, at least at the macro-level.[16] If that's right, there is no phenomenal consciousness without self-consciousness at the macro-level. But if there is to be a transparent explanation of macro-level consciousness (at least in part) in terms of micro-level consciousness, then it seems as if the friend of RP might need to admit that electrons possess self-consciousness. Yet if *that's* right, then in the current proposal, electrons are non-fungible too.

Notice that, despite broadly similar motivations, the issue here is subtly different from that encountered when discussing VST and electron pain. On the view under consideration, unlike pain, self-consciousness is not a *determinate* of phenomenal consciousness. Instead, every determinate of phenomenal consciousness—having an olfactory experience of garlic, a visual experience of a uniform grey expanse, or a slight pain experience—includes, as a constitutive feature, self-consciousness. Put differently: self-consciousness is a base ingredient of phenomenal consciousness.[17] Proponents of RP may deny this, but they cannot do so on the grounds that it would make electrons non-fungible.

Of course, there's uncertainty around all this, so we can't say anything definitive. But on the face of it, RP will be hard-pressed to find a fungibility-based justification for electrons having very low moral status.

4.2 Electrons and Capacity for Welfare

Perhaps, then, we should consider the second strategy for accommodating pro-human intuitions: namely, appealing to differences in capacity for welfare.[18] What might explain such differences?

[16] This view is popular with those in, and influenced by, the phenomenological tradition. It is also endorsed by many proponents of higher-order approaches to consciousness including, e.g., Rosenthal (2005) whose HOTs refer to the self, and Kriegel's (2009) self-representationalism, which emphasises the so-called 'for me-ness' of experience. Of course, here infants *would* be self-conscious if they are phenomenally conscious, making the first worry inapt.

[17] By saying that self-consciousness is constitutive of phenomenal consciousness, we aren't automatically undercutting the anti-reductionist (i.e., reduction to the physical) impulses that often motivate RP. After all, it's the *sense* of self that would need to be reduced.

[18] In principle, electrons could differ in their capacities for welfare, just as there could be intra-species capacity for welfare differences. But we don't know why electrons would vary this way.

The details depend on the theory of welfare, but the structure is the same. On hedonism, for instance, all and only positive experiences are good for you and all and only negative experiences are bad for you. Then, there are two routes to differences in capacity for welfare. First, if individuals can have different welfare ranges they can differ in the strengths of the positive and negative experiences they can have, which means that they can realize different amounts of welfare at a time. Second, suppose that some individuals have longer lifespans than others. Then, given hedonism, individuals who live longer will have greater capacity for welfare: they have longer to accrue welfare than those who live for shorter periods.

The same two routes to variation are available given an objective list theory. However, the explanation for variation in individuals' welfare ranges will depend on the items on the objective list. Suppose that knowledge is on there. Then, if individuals can differ in how much or the kind of knowledge they can acquire, that difference seems to be a plausible basis for differences in how much welfare these individuals can realize.[19]

As soon as we make these observations about welfare goods, however, we quickly see that rather than electrons having a tiny capacity for welfare relative to humans, electrons have an *enormous* capacity for welfare—indeed, almost incomprehensibly large. This is because, however impoverished their experiential lives may be, they exist for the lifespan of the universe (Agnostini *et al.* 2015). Moreover, Lewtas (2013: 53) argues that electrons have their experiences *continuously* since fundamental entities have all their non-relational properties continuously (for example, the mass of an electron doesn't change). In the aggregate, then, if electrons can realize any welfare at all, then they can realize wildly more than humans. Indeed, electrons may have greater capacity for welfare than any other kind of welfare subject.

5. Reasons

Where are we? In brief, if RP is true, then it's difficult to avoid the conclusion that electrons have moral standing. Moreover, if electrons have moral standing, then it isn't easy to defend the conclusion that they're morally trivial. These are incredulous-stare-inducing results. Can they be tempered?

Suppose that we can't affect how things go for electrons. Then, we wouldn't have reason to act differently upon learning that RP is true. If our reasons for acting are essentially unaffected by RP, then while it might be a strange truth (if it's true), it wouldn't be a consequential one. So, let's consider whether this position is open to RP's proponents. To begin, let's be clear that humans *can* affect electrons. Indeed, everything humans do involves altering the trajectories of countless such particles. However, it doesn't follow from our ability to alter their trajectories that we can affect their *welfare*.

Can we affect their welfare? Maybe. Consider a process called 'pair annihilation', where an electron colliding with a positron (its antiparticle) produces two photons —a result of experiments using particle accelerators. Of course, we have no idea whether this process affects the valenced states of electrons, if they have any. For all

[19] Note that objective list theorists assume that different kinds of individuals will have different abilities (or no ability, in many cases) to realize the goods on the objective list. Objective list theorists don't deny that slugs are welfare subjects because they can't achieve the relevant form of knowledge.

we know, after all, pair annihilation occurs too rapidly (instantaneously, at the limit) to affect those states. However, we might not need to sort out such details to assess whether pair annihilation affects electrons experientially, as we can understand it as causing electrons' *deaths*. Then, if we assume a deprivationist theory of death (Timmerman 2019), according to which death's goodness or badness is determined by the future that you would otherwise have had, we can say that pair annihilation is good or bad for an electron in so far as it prevents the negative or positive experiences they would otherwise have had.

There are at least three problems for this defence of humans' impacting electrons' welfare. One of them is that deprivationism could be false: there are interesting arguments in favour of Epicureanism, for instance, which is the view that your death is neither good nor bad for you.

A second problem is that RP isn't committed to any particular fundamental ontology. We've been assuming smallism and that electrons are fundamental, but it could turn out that, on some other smallist fundamental ontology, the fundamental particles are materially harder to affect welfare-wise, or perhaps impossible to affect at all. We just don't know.[20]

The third problem, however, is the most serious. Suppose that RP is true, that electrons have valenced states, and that deprivationism is correct. Then, we can conclude that pair annihilation is either good or bad for electrons. But since we have no idea whether electrons are faring well or poorly on net, we don't know whether pair annihilation spares electrons from bad lives or deprives them of good lives. That is, since we don't know what their futures would have held, we don't know whether death harms or benefits electrons. So, the issue isn't whether we can perform actions that can affect electrons' welfare; it's whether we can affect their welfare in a way that's action-relevant.

It's important to be clear about this third problem. Roughly, the idea is that we just don't know what's good and bad for electrons: we can't harm or benefit them *in expectation* (even if we can harm or benefit them in fact). However, there are two ways of understanding this idea. The first is value theoretic, insisting that we don't even know what *would* be good or bad for electrons. So understood, the claim is that we can't even answer that question at an axiological level.

The value theoretic version can be swiftly dismissed. As Lin (2018) argues, fundamental goods are good for every welfare subject, even if, for whatever reason, that good is inaccessible to that subject. For example, if an objective list theory of welfare is true, then knowledge might be a welfare good. Granted, some welfare subjects don't have access to knowledge (McMahan 1996: 7). Nevertheless, *if* one of those welfare subjects *were* to acquire knowledge, their welfare would increase, all else equal. In short, then, our theory of welfare tells us what's ultimately good and bad for all welfare subjects,

[20] This issue is perhaps more pressing if we reject smallism in favor of *priority monism,* where what's fundamental is the universe itself, with facts about the 'little things' being grounded in facts about the universe. The panpsychist version of this is *cosmopsychism* (e.g., see Shani 2015; Nagasawa and Wager 2016; Goff 2017), where cosmic-level consciousness facts ground—perhaps 'by subsumption', as Goff (2017: 220) puts it—all other facts, including facts about ordinary conscious subjects. This view seems compatible with both affirming and denying the claim that electrons are conscious, but by saying that the *universe* is conscious, the views discussed earlier suggest that the universe is a welfare subject.

whether eagles or electrons. Some will be able to access a wide range of these goods and bads (if there are many kinds of goods and bads); some won't. Regardless, the goods and bads are the goods and bads.

The interesting version of the problem, then, is that in so far as we can affect electrons, we can't predict the sign of the impact (that is, whether it's positive or negative on net). By contrast, we don't just know what's good for many animals *in principle*; we know what's good for them *in practice*. True, we might not know what it's like to be a bat—there's a sense in which the experiential lives of bats are alien to us. Nevertheless, we have some sense of bats' interests: if we cut off their wings without analgesics, they would suffer and that suffering would be bad for them. However, even if we know that pain is bad for electrons, we have no idea *what pains them*.

One way to appreciate this is to try to imagine what it would take to get evidence that a particular event is good or bad for an electron. Like all other conscious beings, we can't access their phenomenal states directly; we must make inferences from their behaviour. Unfortunately, we have no idea what to infer about their preferences, if any they have, from such observations. For instance, electrons repel other electrons and are attracted to protons. Does this mean that electrons find it *aversive* to be near other electrons and *like* being near protons—which is the inference we might make if electrons were eels? In the case of eels, there are evolutionary and physiological reasons to accept such an inference. In the case of electrons, there aren't. So, while we might know what's good and bad for electrons in the abstract (if we know which theory of welfare is true), electrons may be too alien for us to know *when* things are going well or badly for them. And if that's right, then we don't have any reasons to modify our behaviour because of potential impacts on them.

This problem may be unique among the more expansive takes on the distribution of consciousness. Suppose that plants are conscious and that the PT is true—then, they are welfare subjects. Presumably, we're in a much better position to assess whether we are harming or benefiting plants. If we can predict the impacts of our actions on plants, we may have reasons—even if very weak ones—to act differently in light of their interests. A parallel conditional applies to electrons, but the antecedent doesn't seem to be true: we can't predict the impacts of our actions on them. If RP is correct, then electrons may be like sentient aliens outside our light cone: fitting objects of moral concern, but nonetheless irrelevant to our deliberations.

Someone might object as follows:

> Granted, if RP is true, then we have no reason to alter our many actions that affect electrons. Still, we may have strong reason to try to *study* electron welfare. That is, if we have sufficient reason to believe that things can go well or badly for electrons, then given their incredible numbers, we ought to devote huge resources to learning about what's good and bad for them. So, RP may have radical results nonetheless.

By way of reply, it's difficult to run this objection without some kind of maximising decision theory, which is obviously controversial. After all, it seems plausible that we're in a situation whether the probability of advancing our knowledge of electron welfare is vanishingly low. We have no idea how we would even *begin* to study it. Again, electrons could be like sentient aliens outside our light cone: it would be *fitting* to try to learn about their welfare, but given the incredibly low odds of success, there's no moral mandate to do so.

6. Coda

Thus far, we've argued:

(1) Given RP, it's plausible that electrons have moral standing.
(2) If electrons have moral standing, then it is not easy to avoid the conclusion that individual electrons matter to a non-trivial degree.

And, despite (1) and (2):

(3) It isn't clear that RP has any practical consequences.

In closing, we'll defend:

(4) Given (3), proponents of RP can reasonably accept (1) and (2).

In other words, given (3), (1) and (2) do *not* provide proponents of RP with decisive reason to abandon RP.

The case for this last claim emerges from reflections on the process of achieving equilibrium between various commitments—in this case, between ethical theories, specific ethical judgements, the metaphysics of consciousness, and the epistemology of consciousness.[21] How should panpsychists react to the possibility that electrons have moral standing? Should they accept the possibility that we have reason not to build antimatter engines because they destroy electrons? Or should they take this as a *reductio* of RP? Plainly, some panpsychists will get off the boat earlier. They will *not* accept some of the philosophical theses that bring these questions to the fore. However, we're interested in the in-principle question. How bad is the situation for the proponent of RP who accepts (1), (2), and (3)?

In our view, not that bad. Consider, for instance, the incredulous stare in response to Lewis's modal realism. As Lewis tells us, the stare isn't an argument. We already know that modal realism is in tension with common sense. So, what's needed is an overall cost-benefit analysis for the view, which includes an assessment of the *severity* of modal realism's disagreement with moral common sense. Lewis thought that it was easy to overstate the severity of the disagreement. Granted, modal realism implies that there are talking donkeys *in some world or other*. At the same time, Lewis agrees that there *aren't* any talking donkeys since there aren't *actually* any talking donkeys. The disagreement is real, but for Lewis, it's worth the price.

This conclusion is hard to accept if modal realism has dramatic practical implications. If, as Heller (2003) argues, modal realism implies that it's morally permissible to let children drown even when we could save them, that seems bad indeed. But as we've argued, RP isn't like modal realism in this respect.[22] Yes, RP disagrees with common sense (or, at least, *Western* common sense). But it's also the case that it's easy to overstate the disagreement. RP implies that electrons are conscious, but it doesn't imply that each electron has a mental life like yours. Yes, RP may well imply that electrons have moral standing, but it may not imply that you have any reason

[21] Many thanks to Luke Roelofs for raising this point.
[22] Fischer (2017) contends that Lewis can dodge this problem.

to act differently as a result. So, the disagreement could be real without being severe. And if *that's* right—and the theoretical benefits of RP are real—then the price may be right.

Acknowledgements

Many thanks to Luke Roelofs, Jason Schukraft, two anonymous referees for AJP, and an audience at Texas Tech University for very helpful comments and discussion.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

Joseph Gottlieb http://orcid.org/0000-0001-9014-8487
Bob Fischer http://orcid.org/0000-0001-9605-393X

References

Agostini, Matteo *et al.* (Borexino Collaboration) (2015) 'Test of Electric Charge Conservation with Borexino', *Physical Review Letters* **115**. doi:10.1103/PhysRevLett.115.231802.
Barron, Andrew and Colin Klein (2016) 'What Insects can tell us about the Origins of Consciousness', *Proceedings of the National Academy of Sciences* **113**: 4900–4908. doi:10.1073/pnas.1520084113.
Brunet, Tyler and Marta Halina (2020) 'Minds, Machines, and Molecules', *Philosophical Topics* **48**: 221–41. doi:10.5840/philtopics202048111.
Chalmers, David (2015) 'Panpsychism and Panprotopsychism', in Torin Alter and Yujin Nagasawa, eds., *Consciousness in the Physical World: Perspectives on Russellian Monism*: 246–76. Oxford University Press.
Chalmers, David (2022) *Reality+: Virtual Worlds and the Problems of Philosophy*. W.W. Norton.
Coleman, Sam (2006) 'Being Realistic: Why Physicalism May Entail Panexperientialism', *Journal of Consciousness Studies* **13**: 40–52.
Coleman, Sam (2015) 'Neuro-Cosmology', in Paul Coates and Sam Coleman, eds., *Phenomenal Qualities, Sense, Perception, and Consciousness*: 66–102. Oxford University Press.
Crummett, Dustin (2022) 'What if We Contain Multiple Morally Relevant Subjects?', *Utilitas* **34**: 317–34. doi:10.1017/S0953820822000115.
Duncan, Matt (2021) 'Experience is Knowledge', in Uriah Kreigel, ed., *Oxford Studies in Philosophy of Mind*, vol. **1**: 106–29. Oxford University Press.
Fischer, Bob (2017) 'The Moral Objection to Modal Realism', *Erkenntnis* **82**: 1015–30. doi:10.1007/s10670-016-9856-0.
Fischer, Bob, ed. (forthcoming) *Weighing Animal Welfare: Comparing Welfare Across Species*. Oxford University Press.
Gennaro, Rocco (2012) *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. MIT Press.
Goff, Phillip (2017) *Consciousness and Fundamental Reality*. Oxford University Press.
Goff, Phillip (2019) *Galileo's Error: Foundations for a New Science of Consciousness*. Pantheon Books.
Heathwood, Chris (2007) 'The Reduction of Sensory Pleasure to Desire', *Philosophical Studies* **133**: 23–44. doi:10.1007/s11098-006-9004-9.
Heller, Mark (2003) 'The Immorality of Modal Realism, or: How I learned to Stop Worrying and let the Children Drown', *Philosophical Studies* **114**: 1–22. doi:10.1023/A:1024471800380.
Kagan, Shelly (2019) *How to Count Animals, More or Less*. Oxford University Press.
Kammerer, François (2019) 'The Normative Challenge for Illusionist Views of Consciousness', *Ergo* **32:** 891–924. doi:10.3998/ergo.12405314.0006.032.
Klein, Colin (2015) *What the Body Commands: The Imperative Theory of Pain*. MIT Press.

Kriegel, Uriah (2009) *Subjective Consciousness: A Self-Representationalist Theory*. Oxford University Press.
Lee, Andrew (ms) 'Consciousness Makes Things Matter', https://www.andrewyuanlee.com/_files/ugd/2dfbfe_8f899234be99498ba1c10a09fc25f1da.pdf.
Lewtas, Pat (2013) 'What it is Like to be a Quark', *Journal of Consciousness Studies* **20**: 39–64.
Lin, Eden (2018) 'Welfare Invariabilism', *Ethics* **128**: 320–45. doi:10.1086/694272.
Lin, Eden (2021) 'The Experience Requirement on Well-Being', *Philosophical Studies* **178**: 867–86. doi:10.1007/s11098-020-01463-6.
Matthews, Freya (2003) *For the Love of Matter: A Contemporary Panpsychism*. SUNY Press.
Maher, Chauncey (2017) *Plant Minds: A Philosophical Defense*. Routledge.
McMahan, Jeff (1996) 'Cognitive Disability, Misfortune, and Justice', *Philosophy and Public Affairs* **25**: 3–35. doi:10.1111/j.1088-4963.1996.tb00074.x.
Mørch, Hedda Hassel (2014) *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem*, DPhil thesis, University of Oslo.
Nagasawa, Yujin and Khai Wager (2016) 'Panpsychism and Priority Cosmopsychism', in Godehard Bruntrup and Ludwig Jaskolla, eds., *Panpsychism: Contemporary Perspectives*: 113–29. Oxford University Press.
Nagel, Thomas (1974) 'What is it Like to be a Bat?', *Philosophical Review* **83**: 435–50. doi:10.2307/2183914.
Roelofs, Luke (2014) 'Phenomenal Blending and the Palette Problem', *Thought* **3**: 59–70. doi:10.1002/tht3.113.
Roelofs, Luke and Jeff Buchanan (2019) 'Panpsychism, Intuitions, and the Great Chain of Being', *Philosophical Studies* **176**: 2991–3017. doi:10.1007/s11098-018-1160-1.
Roelofs, Luke (2019) *Combining Minds: How to Think about Composite Subjectivity*. Oxford University Press.
Roelofs, Luke (2022) 'Sentientism, Motivation, and Philosophical Vulcans', *Pacific Philosophical Quarterly*. doi:10.1111/papq.12420.
Rosenthal, David (2005) *Consciousness and Mind*. Clarendon Press.
Sebo, Jeff (forthcoming) 'Moral Circle Explosion', in David Copp, Tina Rulli and Connie Rosati, eds., *The Oxford Handbook of Normative Ethics*. Oxford University Press.
Shani, Itay (2015) 'Cosmopsychism: A Holistic Approach to the Metaphysics of Experience', *Philosophical Papers* **44**: 389–417. doi:10.1080/05568641.2015.1106709.
Shani, Itay (2022) 'Eden Benumbed: A Critique of Panqualityism and the Disclosure View of Consciousness', *Philosophia* **50**: 23–256. doi:10.1007/s11406-021-00377-9.
Siewert, Charles (1998) *The Significance of Consciousness*. Princeton University Press.
Singer, Peter (1979) *Practical Ethics*. Cambridge University Press.
Skrbina, David (2013) 'Ethics, Eco-Philosophy, and Universal Sympathy', *Dialogue and Universalism* **23**: 59–74. doi:10.5840/du201323449.
Timmerman, Travis (2019) 'A Dilemma for Epicureanism', *Philosophical Studies* **176**: 241–57. doi:10.1007/s11098-017-1014-2.
Tomasik, Brian (2014) 'Is There Suffering in Fundamental Physics?', https://reducing-suffering.org/is-there-suffering-in-fundamental-physics/.
Tononi, Guilio (2012) 'Integrated Information Theory of Consciousness: An Updated Account', *Archives Italiennes de Biologie* **150**: 293–329. doi:10.4449/aib.v149i5.1388.
Tononi, Guilio and Christof Koch (2015) 'Consciousness: Here, There and Everywhere?', *Philosophical Transactions of the Royal Society B* **370**: 1–18. doi:10.1098/rstb.2014.0167.
van der Deijl, Willem (2021) 'The Sentience Argument for Experientialism about Welfare', *Philosophical Studies* **178**: 187–208. doi:10.1007/s11098-020-01427-w.