ORIGINAL PAPER

Check for updates

Consciousness Ain't All That

Neil Levy

Received: 18 December 2023 / Accepted: 16 April 2024 © The Author(s) 2024

Abstract Most philosophers think that phenomenal consciousness underlies, or at any rate makes a large contribution, to moral considerability. This paper argues that many such accounts invoke question-begging arguments. Moreover, they're unable to explain apparent differences in moral status across and within different species. In the light of these problems, I argue that we ought to take very seriously a view according to which moral considerability is grounded in functional properties. Phenomenal consciousness may be sufficient for having a moral value, but it may not be necessary, and it may contribute relatively little to our overall considerability.

Keywords Phenomenal consciousness · Access consciousness · Prudential value · Moral value · Nonhuman animals

The consciousness boom shows little sign of slowing down. From a topic that was best avoided because engaging was a "career-limiting move" [26], consciousness is now at the centre of philosophical

N. Levy (⊠)

Department of Philosophy, Macquarie University, Sydney, NSW 2109, Australia

e-mail: neil.levy@philosophy.ox.ac.uk

Published online: 24 April 2024

N. Levy

Uehiro Centre for Practical Ethics, University of Oxford, Oxford OX1 1PT, UK

discussion, as well as a major topic in the scientific literature. Part of the reason for this fascination with consciousness is that it's widely, if somewhat inchoately, felt to be important: perhaps the most important thing there is. Consciousness *matters*, whether or not it's material. It's what makes life worth living, if anything does. Or so the thought goes. It's also apparently mysterious, which only adds to the fascination.

There are many senses of consciousness. When I say consciousness matters and is mysterious, I have a particular sense of 'consciousness' in mind. I mean, of course, phenomenal consciousness. Phenomenal consciousness is (supposedly) the sort of consciousness such that (uniquely) there is something that it's like to be conscious in that way [23]. There is, allegedly, something that it's like to hear the opening notes of the prelude to Tristan und Isolde, or to feel sunshine on your cheek, and there's something that it's like to be a kitten chasing a laser pointer, but there's nothing that it's like to be the laser pointer. The laser pointer lacks qualia. This isn't a definition of phenomenal consciousness - no one has succeeded in giving one. Rather, it's a reminder of something we all supposedly know for ourselves, from within.

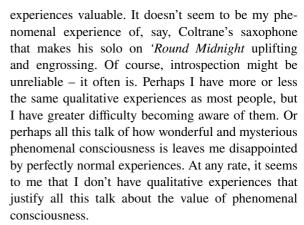
One reason we may have to rest content with ostensive definitions of phenomenal consciousness is that it, alone among the senses of consciousness, is neither reducible to more basic properties nor functionally specifiable. Other senses of consciousness can be cashed out in functional terms, and ascribed to organisms, and even artifacts, in virtue of how they



are constituted. A machine might be self-conscious: it might be set up such that information about itself is in some way tagged and processed as information about itself. It might be access conscious: it might be set up in such a way that the outputs of information-processing subsystems are globally broadcast, such that a wide array of such systems have available the results of the processing of the other systems. Phenomenal consciousness is (intuitively) not functionalizable: not only do we not have a functional account of what's-itlikeness that explains how it arises or its properties, it seems to many that we cannot even imagine how such an account could be possible. It's this fact that makes phenomenal consciousness so apparently mysterious. It also helps account for some of the fascination: it's this ineffable kind of consciousness that underwrites important kinds of moral and prudential value. We are valuable in certain distinctive ways, that is, only because we're phenomenally conscious.

21

While I share some of the intuitions that underlie debates over consciousness, I am somewhat puzzled by much of the fascination with consciousness. I am not – here I want to insist – a zombie. At least, it seems to me that I have phenomenally conscious experiences. But if introspection and the reports of others are a guide, I seem to have weaker and more recessive phenomenal consciousness than others. Pain I get: it seems to me that pain is bad at least in large part because it seems a certain way, and my pain experience can be intense. But pain is one of the rare phenomenally conscious experiences I have that seems intense. Other phenomenally conscious experiences have qualitative properties, it seems to me, but, well, they ain't all that. They're neither especially intense nor especially rich. Nor do they seem to be the properties that make enjoyable or even sublime



Suppose I'm right. Does it matter for me if my qualitative experiences are impoverished, relative to yours? Should you pity me? More pressingly, perhaps, what does my weak qualitative experience entail for my moral status? Am I less morally considerable, given that many philosophers believe that moral considerability depends on phenomenal consciousness? Might it be less wrong to kill me than you, or perhaps even less wrong to kill me than that kitten chasing the laser pointer? That would be bad news for me. But perhaps it wouldn't be all that bad for me, since – after all – my life might matter less for me in virtue of its impoverished phenomenology. On the other hand, perhaps moral value is grounded in functionalizable properties, or perhaps my impoverished phenomenal consciousness is enough to get me in the game – and now that I'm in the game, functionalizable properties ensure that I'm more valuable than a kitten (yay!).

It is this set of issues I aim to explore. In section one, I'll briefly set out the arguments that philosophers have advanced for the claims that prudential value, and a certain kind of moral value, depend on phenomenal consciousness. In section two, I will mention some problems that these views seem to run into, before advancing what to me seems a more serious worry: that accounts that ground a special kind of moral considerability in phenomenal consciousness are unable to explain moral differences between different species.² I also argue that many of the arguments that have been given for the specialness of



¹ From this point on, I will say my phenomenal consciousness appears to be impoverished relative to the consciousness of others, so far as I can tell. This impoverishment might be manifested in more than one way. It might be a property of its content or the intensity of that content, and these properties might fractionate in turn (my conscious experiences might have fewer elements, or those elements might not be unified in the way they are for others, and so on). Perhaps these distinctions make a difference to the value of phenomenal experiences, but the very impoverishment of my experience, coupled with the inaccessibility of the experiences of others, suggests to me that attempting to isolate the respects in which my phenomenal experiences are impoverished is unlikely to yield reliable findings.

² Kammerer [17] offers a very different argument for the conclusion that phenomenal consciousness is unable to ground differences in moral considerability across species. He argues that the truth of materialism entails illusion of at least a weak sort, and weak illusionism entails that it is impossible in principle to

Neuroethics (2024) 17:21 Page 3 of 14 21

phenomenal consciousness are question-begging. I conclude that the case for grounding moral considerability in cognitive properties that phenomenally non-conscious beings might possess is much stronger than is typically recognized.

Prudential and deontological value

A number of philosophers have argued that phenomenal consciousness is necessary for what I am calling prudential value; that is, an organism can only have a welfare that matters to itself if it is phenomenally conscious. Organisms and artifacts might be said to have a welfare in the absence of phenomenal consciousness. A plant may flourish or wither; a machine may degrade or function well, and in virtue of these properties we may ascribe to it a welfare. But it doesn't have a welfare for itself. It doesn't matter to the machine or the plant what its level of welfare is (perhaps it matters only to us, who are phenomenally conscious, or perhaps its welfare is an objective property, regardless of whether anyone appreciates it).

The case for the claim that prudential value depends on phenomenal consciousness has largely been prosecuted by thought experiment; in particular, by reference to the phenolectomy thought experiment introduced by Siewert [33] Variations of Siewert's thought experiment continue to be central to discussion of the topic. I will use a variation due to Uriah Kriegel, which itself adapts a later version of Siewert's thought experiment (from [32]). Kriegel asks us to imagine that God comes to you and makes you an offer. If you accept, your life will go astonishingly well from now on. You will achieve all your desires (perhaps you're going to write a book on the philosophy of consciousness that will revolutionize the field and win you the first Nobel Prize awarded to a philosopher since Jean-Paul Sartre). But there's a downside:

Footnote 2 (continued)

know to what degree different species are phenomenally conscious (it also entails that introspection about consciousness is highly unreliable, rendering it mysterious why we should think that phenomenal consciousness grounds moral considerability at all). While I am sympathetic to (if not wholly convinced by) illusionism, even the weak illusionism Kammerer sketches is a minority view among philosophers and scientists; it is therefore preferable to demonstrate the problems with grounding value in phenomenal consciousness without assuming it.

you'll undergo a phenolectomy. From the moment you accept the offer, you'll lack all phenomenal consciousness. Your cognition will be unchanged, to the extent it's functionalizable, but you'll *feel* nothing.

Kriegel (forthcoming) thinks the prospect of a phenolectomy is dismaying. Indeed, he thinks a phenolectomy would rob his life of all value for him. Of course, it would be a wonderful thing to produce that great work of philosophy, and perhaps its value would be sufficient to justify him accepting the offer of such an achievement, at the expense of phenomenal consciousness. But this wonderful achievement wouldn't be a wonderful achievement that added any value to his life for him. Kriegel thinks the prospect is so bad that he says he would be indifferent between this fate and another, where he is destroyed and replaced with an identical duplicate. He's indifferent between being zombified and being destroyed. In both cases, great achievements might attach to his name, but in neither would they be of value for him.

If the Siewert/Kriegel thought experiment pumps the intuitions they report and those intuitions give us insight into what is actually of value, it's a necessary condition of having a life that matters to one that one is phenomenally conscious. I'm not sure I have the relevant intuitions (perhaps one needs a livelier phenomenal life than I have to have these intuitions). Nevertheless, given that other people report having them (and I'm not certain I don't have them too), they're worth taking seriously. Whether this is bad news for me depends, of course, on the role that phenomenal consciousness plays in prudential value. Perhaps (for example) I might be capable of having a subjective welfare, in virtue of the fact that I am (I suppose) phenomenally conscious, but I might have to work harder than other people to achieve the same level of welfare.

I'll return to this issue. I now turn to the moral value of phenomenal consciousness. As I read Kriegel, Siewert, and also Diane O'Leary [25], they do not suggest that a being must be phenomenally conscious to be morally considerable. Rather, they argue that moral considerability of a certain distinctive kind requires phenomenal consciousness. Only if we are phenomenally conscious are we agents protected by deontological constraints. Diane O'Leary argues that it is in virtue of phenomenal consciousness that doctors owe respect to their patients [25]:



80). For Kriegel [19], phenomenal consciousness grounds the dignity of inviolate beings.

The argument they advance turns on irreplaceability (at least in O'Leary's and Siewert's cases). Each of us has a stream of phenomenal consciousness such that that stream is ours alone. It follows that our phenomenal consciousness is irreplaceable. Because that's the case, we can't be traded off against one another. We are not mere bearers of value, but separate persons: our irreplaceability protects us against utility maximizing trade-offs. Trees and bacteria might be morally considerable, but they are not protected by deontological constraints: their welfare may be traded for ours or even for the welfare of other non-conscious beings. But we are inviolate.

If these twin claims are accurate, phenomenal consciousness is central to the value of each of us. We would still have a welfare if we weren't phenomenally conscious, but we would have no welfare for *ourselves*. We would be morally considerable if we weren't phenomenally conscious, but our well-being could be traded for the well-being of others. Perhaps my impoverished phenomenal consciousness might buy me subjective welfare and protect me against trade-offs; perhaps not. Again, that depends on what role they play, precisely. I turn to that issue next.

The Experience Machine

As Kriegel emphasises, the intuition that our prudential and moral value is underwritten by our phenomenal consciousness appears to conflict with another intuition: that life in the experience machine would have little or no prudential value. The experience machine [24] simulates a wonderful life (perhaps the very life that God offered you after your phenolectomy). Unlike the phenolectomy scenario, in the experience machine you'll have all the experiential riches you could possibly want: every achievement, every sip of wine, every caress will be enjoyed phenomenally. But none of it will be real: it won't be caused by actual wine or actual achievement. Most philosophers are (almost?) as reluctant to enter the experience machine as they would be to agree to a phenolectomy. The phenolectomy intuition is that a life without phenomenal consciousness lacks prudential value - value for the person whose life it is - but the experience machine intuition seems to show that phenomenal consciousness isn't sufficient for a significant degree of prudential value.

If the phenolectomy intuition is correct, then writing Consciousness Finally Explained (no, really this time) is of zero value for my life if I don't experience it. On the other hand, if the experience machine intuition is correct, seeming to experience something is of little (perhaps not zero) value compared to actually experiencing it. If actually writing the book had some value, it would be easy to see why the two combined should be more valuable than either on its own. But it doesn't: it has zero value. But if actually writing it is of zero value, how does combining it with experience add up to more value than experience alone? Where does the extra value come from? What explains the (apparent) fact that different experiences have different degrees of value? It's a commonplace that some of our experiences are profound and important, whereas others are routine. But in both cases, we're phenomenally conscious of actual experience; what explains the differences across cases? The whole thing seems mysterious.

Kriegel considers one possible solution: that experiences matter (more or less) as organic unities. The organic unity of my experiencing a sublime landscape or a wonderful performance of Verklärte Nacht is much greater than the organic unity of my experiencing the taste of a sandwich bought from the trolley on a British train. But Kriegel despairs of finding a way of spelling out this account in even a halfway satisfying way. He prefers an experientialist view, according to which experience somehow 'unlocks' the value that is stored in potential in accomplishments (and in art, and so on). This view would, indeed, yield the right result: the view is consistent with the phenolectomy intuition, since the value is merely potential when it is unexperienced, and consistent with the experience machine intuition since there's no extra value unlocked by simulated experience. But this, too, is mysterious: as Bradford [4] notes, unless we are able to offer some explanation for just how consciousness activates the value stored in achievements, the view seems ad hoc at best.

Kriegel himself admits it's hard to make sense of the idea, but finds it less costly than rejecting the twin intuitions prompted by the experience machine and the phenolectomy. Perhaps he can avail himself of Bradford's defence of the organic unity account, since on her version of the account, the organic unity of



achievement and phenomenal experience "activates" the latent value of the former. She suggests that consciousness *unifies* the actual achievement with the agent, "by bringing it into his experience", and that this unification increases the value, for the person, of both the consciousness and of the achievement. But this is at best question-begging. Why should it be *phenomenal* consciousness that unifies an agent with her achievements, or other welfare constituting goods? Why not, say, access consciousness? Indeed, given that access consciousness is a functional notion, and may be cashed out, precisely, in terms of how it makes information available to the processing systems constitutive of the agent, the latter is a promising candidate for the unifying role.

Given the mysteriousness of these accounts, we have good reason to look elsewhere for a view that might explain the phenolectomy intuition. In the only full-length book on the value of consciousness, Joshua Shepherd [31] presents an account that might achieve the goal.³ He introduces the notion of thick experiences, modelled on the thick concepts of ethicists. A thick experience is a set of experiential properties that is to some degree unified. Most of our experiences are thick: we don't experience yellowness or sweetness by themselves, nor do we experience yellowness and sweetness alongside one another. Rather, we have a unified experience: right now, I have the experience of hearing Red Garland's piano, feeling the pressure of the chair on my back and seeing the black and white words unfold on my laptop (not to mention the complex cognitive phenomenology of struggling to find the right words). Shepherd argues that the value of an experience depends on its determinate properties. For the state to be valuable, some of these properties must be evaluative and some of these evaluative properties must be affective. Together, these properties determine the shape of our experiences.

Shepherd is explicit: phenomenal consciousness, when it's sufficiently rich, is valuable in its own right. It need not be experience of anything real.⁴ Shepherd therefore seems to bite a bullet, rejecting the experience machine intuition. Perhaps this is a cost we should be prepared to pay; unlike Kriegel, I don't find the intuition compelling enough to regard it as worth embracing mystery for. However, all these accounts of the value of phenomenal consciousness face a different, and to my mind much more difficult problem.

The Moral Problem

Kriegel, Siewert and O'Leary are united in thinking that phenomenally conscious beings possess a distinctive kind of moral value that non-conscious organisms cannot possess. We are worthy of respect as ends in ourselves; we have an irreplaceable value. We are protected by deontolgical constraints. But this immediately runs us into serious moral problems.

Like many contemporary moral philosophers, I think contemporary societies accord far too little weight to the interests of non-human animals. For example, I take factory farming to be morally unacceptable. But - in common with most people, I think - I nevertheless think that humans are at least typically more morally considerable than at least most other animals, including most other vertebrates. That is, I think that when we face a choice between human and non-human life, we are permitted and very plausibly required to choose human life, other things being equal. I don't want to suggest any metric of how much more morally considerable we are, beyond saying that this weight is not so great that we are always permitted to prefer human life. We can't value our convenience over their suffering, for instance. However, when we face trade-offs – as we inevitably do – we ought to

⁴ The only discussion of the experience machine in Shepherd's book considers what he calls the part-time experience machine, which is something like a souped-up virtual reality set up that users enter and exit when they feel like it. He reports no hesitation in enjoying the part-time experience machine. Since he holds that the part-time experience machine gives us access to items "of great non-derivative value", I take it that Shepherd is committed to thinking a wholly simulated life is also of significant value.



³ In more recent work, it's worth noting, Shepherd has come to be more sympathetic to the view that phenomenal consciousness is not necessary for direct moral considerability [29, 30]. While he does not have a settled view, he now thinks that it is plausible to ground moral considerability on the capacity of an entity to have "cognitive contact" with items of value, where such contact need not require phenomenal consciousness [29].

prefer human life when, for example, the numbers are balanced.

The problem that any account that grounds deontological constraints in phenomenal consciousness faces is obvious: phenomenal consciousness is almost certainly widely distributed among vertebrates (whether it extends to invertebrates is controversial). It seems to be badly mistaken to think that mice, say, are due to the same sort of respect as we are, or that mice are of irreplaceable value. If it's not mistaken - if mice are due this sort of consideration - then it seems that the trade-offs of their welfare for our own that we make routinely and uncontroversially (unlike, say, factory farming) may well be impermissible. We can't, for example, put down traps or bait to protect grain storage silos from mice; not even if the number of mice that will die is low and the number of humans dependent on that grain significantly higher. That would be to treat the mice as mere containers of value, and that would be impermissible.

The problem is most acute if we adopt a threshold approach: that is, if we hold that a being that possesses phenomenal consciousness is in virtue of that fact entitled to protection by deontological constraints. In that case, mice and humans would count equally. Can we avoid it by a graded notion of moral considerability, where such considerability is nevertheless grounded in phenomenal consciousness? That is, might we hold that we are more morally considerable than them because we have a richer phenomenal consciousness? Such a move might be bad news for me. Whereas on the threshold approach, I count equally alongside you (and the mice), on the graded approach I might count less than you. Perhaps my degree of moral considerability is closer to that of the mouse than to yours. In any case, I'm sceptical the graded approach can come to our rescue here.

The central problem is that it's not at all obvious that we – even *you* – have richer experiences than other animals. They may be less complex along one dimension – we are more cognitively complex than other animals, insofar as we make finer distinctions and broader inferences across a wide range of domains – but for all we know, they may be richer overall. Perhaps their phenomenology is more intense *because* its purer (Shepherd suggests that the purity and wholeheartedness of a child's experiences might make them just as valuable as the more complex and refined experiences of an adult; [31],pp. 80–1).

Moreover (for all we know), non-human animals might have richer phenomenology than us in virtue of a greater number and diversity of sensory inputs. Famously, bats have the sense of echolocation, but other animals have even more exotic senses: electroception (platypuses), polarized vision (octopuses), detection of the Earth's magnetic field (bees), and so on. The apparently simple mantis shrimp has between 12 and 16 different kinds of photoceptor cells (compared to our 3) and some of them possess the ability to tune the sensitivity of their color vision to the local environment [11]. They have more complex visual experiences than us, it appears, and perhaps, therefore, a richer visual phenomenology. If phenomenal consciousness comes in degrees and moral considerability varies continuously as a function of phenomenology, you might find ourselves displaced from the top of the hierarchy to somewhere further down (who knows how low I might fall?).

To my knowledge, none of the philosophers who ground deontological constraints in phenomenal consciousness have offered a response to this kind of concern. But Joshua Shepherd, whose work I just mentioned, has responded to a related concern: that grounding some sort of significance in the richness of phenomenal consciousness might lead to invidious comparisons across human beings. Any theory that commits us to thinking that some humans are more valuable than others is one we should resist. Shepherd denies that his view leads to such a conclusion; perhaps we can adapt his response to our concerns?

On his view, the value of an experience is settled, in part, by the "number, type and subtlety of a subject's affective-evaluative capacities" [31]: 80). Now, these capacities are not themselves phenomenal or wholly dependent on our basic capacity for phenomenality. They are at least partly and perhaps entirely functionalizable. They're certainly trainable. Shepherd compares the experience of a composer and a small child, each listening to a Debussy étude. The composer has a richer and more valuable experience, in virtue of their training. Because they (say) understand the étude's place in musical history and grasp its innovations in texture and harmony, their phenomenology is richer than the phenomenology of the untutored child. All by itself, that suggests differences in the prudential value of the phenomenology each experiences, and therefore of the well-being of the person whose phenomenology it is.



While the value of an experience doesn't depend on our basic capacity for phenomenology, it might well be that I have to work hard to reach the same level of phenomenal value as you. Compare our experience of music. Perhaps you achieve a certain level of value, n, simply in virtue of the richer phenomenology you enjoy, relative to me. I can achieve that same level by acquiring a more sophisticated understanding of music, but I may be at a disadvantage, compared to most people. For me, it's hard work to achieve the same level of value you achieve with ease. One might think that doing this sort of work is itself valuable, and it's no bad thing to be motivated to do it. But each of us has limited time and resources to train our capacities. Perhaps I will sail past n when it comes to my experience of music, but you will achieve a higher level of value across a range of other sorts of experience, and I simply won't have time to match the overall value of your experiences.

Shepherd is strangely unconcerned with the implications of his account for me personally. But he is alive to the worry that it will entail that some people have lower levels of experiential value than others. He dismisses this worry, on the grounds that experiences come in all kinds, and the person (or animal) who experiences less value in this or that domain may be compensated by experiences in another. Continuing with the example of the child and the composer, though the child may have less valuable experiences of Debussy (because she currently lacks the rich set of evaluative capacities of the adult), her experiences of play may be purer and more intense than anything the composer is now capable of. The space of phenomenal value, he writes, is multi-dimensional. It is so diverse and complex that comparisons "will rarely be appropriate" (83).

But the non-comparability claim doesn't seem to follow from the multi-dimensionality claim; not unless it's supplemented with an argument for the incommensurability of thick experiences. Perhaps a more charitable claim would be that experiences can be compared, but what we'll find on comparison is that the experiences of different agents tend to involve more or less the same degree of phenomenal richness. But why should we believe that? As we've seen, there are some reasons to worry that I would have to work hard to instantiate the same degree of phenomenal richness as someone else who doesn't suffer from my deficits in the strength

of phenomenal consciousness. To get the same value out of listening to Debussy as you do, I might have to study music theory intensely, while you can simply relax and enjoy the music. I see no reason to think I'm unique in having this problem. Perhaps we will need to recognize the existence of a class of lesser humans, to which I belong.

The hope that each of us instantiates more or less the same degree of phenomenal richness rests on an assumption that Shepherd never articulates: non-correlation. The non-correlation thesis is the thesis that the different dimensions of the space of phenomenal value do not correlate. That thesis may well be false. If native phenomenal richness enables the acquisition of theoretical capacities, say – if the person with an initially richer experience of music is more motivated or more able to develop a theoretical understanding of it - then we might well see greater divergence over time between those with more and those with less. That seems eminently plausible: finding music richer on early encounters might motivate deeper exploration and may give the person a head start in identifying the components of musical complexity. We may see positive feedback loops that ensure that those who begin with a small initial advantage in phenomenal complexity diverge more and more from those without such an advantage. Worse - because more unjustly – we might see a systematic correlation across people depending on their socio-economic status. Those with more time and more opportunities to develop their capacities will be rewarded and therefore more motivated to develop further. Over time, the haves and have-nots will tend to be haves and have-nots in the value of their experiences.

It's not clear, then, that Shepherd succeeds in demonstrating that different agents won't have significantly different degrees of prudential and moral well-being, in virtue of their thick experiences. Let's leave that aside, though, and return to the moral problem. Assuming his account succeeds, does it offer us a response to the concern that grounding significance in phenomenal consciousness leaves us unable to justify prioritizing human welfare over the welfare of, say, mice? No: if it succeeds, it succeeds too well. Being able to justify differences across phenomenally conscious beings is precisely what is needed to solve this problem. The non-comparability claim entails there are no such (relevant) differences. If we are to hang on to the claim that humans are morally more



considerable than other animals, we're going to have to look elsewhere.⁵

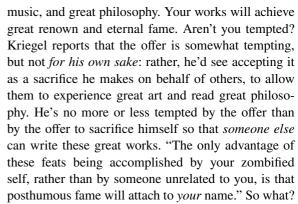
21

A tempting view is that phenomenal consciousness and other aspects of cognition make independent contributions to moral considerability. My apparently rich phenomenal consciousness of pain might be sufficient to grant me protection against its infliction (ceteris paribus, of course), but it might be wrong to frustrate my plans for other reasons. My greater cognitive complexity might more than compensate for my relative paucity of phenomenal richness compared to a mantis shrimp. But if cognitive complexity can underwrite moral status, we should wonder just how far this goes. Is phenomenal consciousness necessary for moral and prudential value at all? Perhaps it's neither necessary nor sufficient, but instead plays a fairly minor role in value. I have defended such a view elsewhere [21], and I will recap that account in a moment. Before turning to it, it's worth pausing to consider how costly it would be to reject the phenolectomy intuition.

As we saw, Kriegel (forthcoming) finds both the experience machine intuition and the phenolectomy intuition so compelling that he would rather pay the cost of accepting a degree of mysteriousness in the account of their source than to give them up. Some philosophers have offered error theories for the experience machine intuition [6],see [27], for a response), and we have seen that Shepherd does not find the thought experiment compelling. I take no stand on the experience machine thought experiment, but I will offer some (albeit indirect) reasons to think that the phenolectomy intuition does not reflect our insight into the grounds of value, but might instead reflect historically variable considerations.

In this context, consider Kriegel's (forthcoming) remarks on the value of the offer God makes in offering the phenolectomy. Accept the deal, and you'll acquire the ability to write great plays, and great

Take no stand on whether we *should* want to hang on to this claim. Some philosophers are now sympathetic to the view that persons are morally more considerable than non-persons, but that personhood extends beyond our species. They will be happy to demote humans from the top of the hierarchy, but they, too, need to give up the idea that phenomenal consciousness is sufficient for rights and perhaps even the claim that moral status varies with phenomenal richness. Personhood and phenomenal richness do not obviously correlate.



I'm no classicist, but as I understand things, it is widely held that for the Greeks of Homeric times, heroic glory (kleos) was the greatest good a person could strive for (see [28], for example). This was a glory that could be won on the battlefield, but also through intellectual achievement, such as writing an epic poem. At least in the *Iliad*, it is a good more important than life itself (the exchange between Achilles and Odysseus in Hades in Book 11 of the Odyssey is sometimes seen as reflecting a lower valuation of kleos).⁶ For the Homeric Greeks, Kriegel's deal might be a no-brainer: of course the value of achieving great renown through wondrous deeds is worth the sacrifice of phenomenal consciousness. This suggests that Kriegel's intuitions, the intuitions he expects us to share, may not have the right status to serve the pivotal role they're called upon to play. They may be too historically specific to give us insight into the nature of the value of consciousness itself.

Of course, there are various ways to push back against this claim. We might argue that the Greeks fully appreciated the incomparable value of consciousness; they just happened also to believe that *kleos* was even more valuable. It's perfectly coherent to believe that losing phenomenal consciousness is akin to death and also to believe that some things



⁶ When Odysseus consoles the shade of Achilles on his death, reminding the latter of the glory he achieved in his life, Achilles responds angrily: "I would rather follow the plow as thrall to another man, one with no land allotted him and not much to live on, than be a king over all the perished dead." *The Odyssey*, Book 11, 489–91. Richmond Lattimore trans.

⁷ Even today, intuitions about the significance of posthumous harms and benefits vary across individuals. See Boonin [3] for a defence of the view that posthumous harms have much greater significance than many believe.

are worthy dying for. Note, however, that accepting a phenolectomy to achieve kleos is a very different prospect to accepting a phenolectomy to achieve social justice: in the second case, the good generated for others makes the sacrifice worthwhile, whereas - we're supposing - it is the good for the person themselves that justifies the phenolectomy. We have good reason to think that Homeric Greeks would see the deal as worthwhile for their own sakes. Another possibility is that the Greeks didn't appreciate the significance of phenomenal consciousness, but that's because they lacked the conceptual competence to focus clearly enough on consciousness to grasp its value.8 I don't have anything like the competence to settle the matter. Rather, my point is simply that the appeal to intuitions here shouldn't be regarded as dispositive without further work.

Do we need to appeal to phenomenal consciousness to ground moral considerability at all?

If it turned out to be very difficult or impossible to explain why we have duties to other people and animals without appealing to phenomenal consciousness, then intuitions be damned: we'd have a strong reason to think that a phenolectomy was a very serious harm; perhaps even the elimination of the intrinsic moral value of the person who undergoes it. As we've seen, some philosophers do think that such an appeal is necessary to ground at least certain, central, kinds of moral value. Diane O'Leary argues that it is in virtue of phenomenal consciousness alone that doctors owe respect to their patients [25]: 80). Kriegel [19] concurs: phenomenal consciousness grounds the dignity of inviolate beings. Siewert [32] reasons along similar lines: phenomenal consciousness is a necessary condition of taking responsibility for one's own thoughts, for respect, and of irreplaceability.

It's important to note that these arguments are meant contrastively: each claims that phenomenal consciousness is required to ground vital components of moral status and that no other kind of cognition could achieve the same end. The claim is not that *only a thinking being* could have these properties: it is that *only a phenomenally conscious being* could have

these properties. O'Leary is explicit that a conscious human being who lacked any other mental states or properties would still be owed the same respect as you and – just possibly – I. But all these arguments beg the question: they appeal to properties or facts that phenomenal consciousness shares with other kinds of cognitive processes or states.

Ned Block [2] famously distinguished two primary senses of 'consciousness.' So far we've focused on phenomenal consciousness. But as Block points out, it's very common - especially outside philosophical circles – to have what he calls *access* consciousness in mind when bandying the term about. Whereas phenomenal consciousness has a qualitative feel, access consciousness is a functional state: information is access conscious just in case it's available at the personal level. The philosophical fascination with phenomenal consciousness arises, in part, precisely from the fact that it is *not* a functional notion: it alone gives rise to the hard problem of consciousness [8], which consists in explaining how this mysterious property can arise. I have no pretentions of chipping away at the hard problem. Rather, my aim is to show that however mysterious and fascinating phenomenal consciousness might be, much less mysterious notions like access consciousness can do all or most of the moral work phenomenal consciousness is called upon to perform.

O'Leary and Kriegel both appeal to the *privacy* of phenomenal consciousness to underwrite central aspects of moral considerability. For O'Leary, its privacy grounds the respect doctors owe to patients: a "zombified patient" would lack "unique authority about her body in the medical setting" (80). She also appeals to the fact that we cannot choose to escape our conscious experiences. Similarly, Kriegel argues that the privacy of consciousness grounds our dignity: "inviolability attaches to conscious creatures precisely in virtue of the fact that the conscious experiences of each conscious creature can only be experienced by them" [18]: 515). Others may know our unconscious mental states better than we do, but only we have direct access to our phenomenal states.

But access consciousness is every bit as private as phenomenal consciousness. Right now (let's suppose) I'm thinking about giraffes. My giraffe thoughts are accessible to a variety of processing mechanisms and globally broadcast, but only I am the recipient of these broadcasts. Perhaps there's something they feel

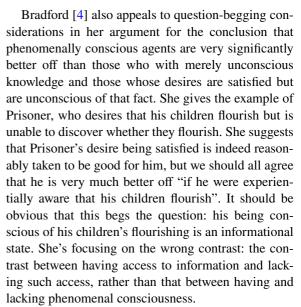


⁸ This, roughly, is how Ned Block [1, 2] suggests we ought to understand the evidence Julian Jaynes [15] cites for a more radical thesis: that the Greeks lacked phenomenal consciousness at all.

like, perhaps not. Either way, you would have no way of finding out about my giraffey thoughts if I hadn't told you. They're just as private as my experience of redness or saltiness (equally, I can't see a way to escape them – not a way that isn't just as available with regard to my phenomenal experiences, at any rate).

Perhaps the thought here is that phenomenal consciousness is essentially private. It's a contingent fact, given the state of neuroscience, that we can't currently discover the contents of individuals' access conscious states unless they tell us. Recent advances promise, or threaten, to end the privacy of such states [34]: soon, we might be able to decode the neural correlates of access conscious states sufficiently well that they won't seem especially private any longer. Kriegel, O'Leary and other philosophers who defend the unique privacy of phenomenal consciousness might have a sense that it, at least, won't yield to such advances. They may rest their conviction on a sense that the hard problem won't soon, perhaps ever, be solved: we may never understand just how brain states give rise to phenomenally conscious states. In the absence of such an understanding, they may think, we can't read off qualia from brain scans. That's a mistake, though: there's no reason to think that it will be harder to uncover the neural correlates of phenomenal states than of access conscious states. We don't need to solve the hard problem to uncover these correlates. We don't need to understand why a certain pattern of brain activation gives rise to, or correlates with, a certain phenomenal experience. We can ascribe phenomenal states to agents just so long as there's a reliable relation between experience and its correlates; there's no in principle reason to think that there's no such relation.

Above, we saw that O'Leary and Siewert appealed to the irreplaceability of our phenomenally conscious states to ground our moral considerability. That, too, is a mistake: there's no sense in which my phenomenal states (poor as they are) are any less replaceable than my access conscious states – so far as I can see, the same goes for yours. They're mine, in just the same way and to just the same extent. The *contents* of my access conscious states can be thought by you. You too can think of giraffes, and under the same aspects if you like. But equally, you can experience phenomenal redness or saltiness.



Siewert [32] gives a different argument for the claim that phenomenal consciousness grounds responsibility and irreplaceability. Siewert notes that agents may express themselves verbally without - consciously - planning what to say. Siewert suggests that it is only if you have the subjective experience of endorsement of such assertions that you make what you say your own and thereby open yourself to normative assessment. I find this deeply puzzling. 'Endorsement' seems to be a functional notion: I endorse my assertions if I have a second-order attitude toward them, or something along those lines (see [12], for defence of a view like this). I don't see what the phenomenal experience of endorsement adds. Consider Donald. Donald suffers from a rare condition: his phenomenal consciousness comes and goes. Donald is otherwise cognitively normal. Donald says something deeply racist. Saying deeply racist things is in character for Donald and Donald is satisfied that this assertion is one that he stands behind. Siewert must maintain that we can't yet know whether Donald is open to normative assessment for his assertion; not unless we know whether he had the normal phenomenal experience of endorsement. That is surely the wrong result. We wouldn't say Donald is responsible for assertions 1, 4, 6, 7 and 8, but not for 2, 3 and 5, because his fluctuating consciousness ensured he did not experience the endorsement of those assertions. Were Donald to be charged with hate speech, I doubt a lawyer could get him off on the grounds that there



is reasonable doubt that he was phenomenally conscious at the crucial moment.

Kahane and Savulescu [16] appeal to the "point of view" (13) of conscious beings to ground aspects of their moral considerability, but point of view has the same problem as endorsement: it's a functional notion. The global workspace, which I have elsewhere argued realizes access consciousness [22], integrates information from a variety of different processing systems, thereby enabling and underwriting a point of view. Point of view might be important for considerability, but it doesn't seem to require phenomenal consciousness.

We have prudential value – and therefore, perhaps, moral considerability – if we are the kinds of beings who have a subjective well-being. We might, therefore, look to accounts of well-being to find the unique role of phenomenal consciousness. Prima facie, at least, the route isn't promising. There are three major theories of well-being: objective list, desire satisfaction, and hedonism [10]. On objective list theories, a good life is one that contains love and friendship, the experience of art, some degree of self-government, achievement, and so on. They don't seem to require phenomenal experience. A zombie seems to have as much opportunity to enjoy most of the items that feature on objective list theories as you do. A zombie also seems capable of satisfying its desires, at least if a desire has as its content that a certain state of affairs obtains (see [17], for an argument that an animal ethics that grounds considerability in desire satisfaction can dispense with phenomenal consciousness).

In recent work, Siewert [32] 28), responding to Levy [21], argues that that's a mistake: his zombie twin can't satisfy *his* desires, because what he desires is to feel a certain way. He reports, for example, that what he wants when he wants to experience the sun on his cheeks is the phenomenology of that feeling. If his twin has the same desires as he has, then his twin is doomed to dissatisfaction. That's sad, but is it any reason to think his zombie twin can't be morally considerable or have a high level of prudential well-being? I think not. Suppose zombie Siewert is doomed to frustration in his desire to phenomenally experience the sun on his cheeks. Well, I am sad to report that I am doomed to frustration in my desire to play football for Arsenal. That's life: we come to

recognize that we can't have some of the things we would like to have and regretfully set them aside.

Perhaps, though, what Siewert means is that our desires are always or typically accompanied – perhaps even partially constituted – by a desire to phenomenally experience certain states of affairs. So construed, this objection has ramifications beyond desire satisfaction theories of well-being. Perhaps some of the items that feature on the objective list of constituents of well-being play their role in underwriting prudential value only, or partly, in virtue of their phenomenology. Perhaps friendship contributes to well-being only when we feel a certain warmth toward our friends. Perhaps aesthetic experience must feel a certain way to contribute to well-being.

I don't rule out the possibility that phenomenology contributes to prudential well-being. However, I do think we need to be cautious in asserting that when we desire something, we desire not only that a certain state of affairs obtains but also a particular associated phenomenology. Our introspective access to the content of our desires is far from perfectly reliable, and we're apt to confabulate in this arena. Even if the satisfaction of our desires depends on our feeling in certain ways, moreover, it may be that these feelings need not be phenomenally conscious. Perhaps, for example, the value of friendship depends in part on a certain feeling of warmth or closeness, but need that feeling be essentially or partly phenomenal? Reflection on Capgras syndrome gives us some reason to think not. Capgras is widely held to arise from a dysfunction that prevents the person experiencing the feeling of familiarity that normally attends meeting someone you know well [9]. Is the experience phenomenal, or conscious in any sense at all? Capgras is widely held to be the mirror image of prosopagnosia, a condition in which people struggle to recognize the faces of familiar individuals. Prosopagnosia is the mirror image of Capgras because in this condition, the affective response is preserved, while the conscious recognition system is dysfunctional [5]. But sufferers from prosopagnosia (who are otherwise cognitively neurotypical) can't use the affect as a guide to recognition. That suggests it isn't conscious: Siewert's zombie twin could experience it in

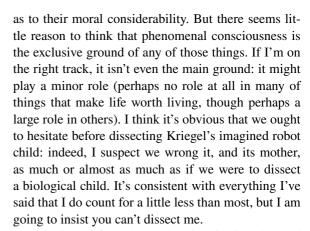


the same way he can. Perhaps something like this is true for the feeling of sunshine on one's cheek too: what we want is very significantly, and perhaps entirely, exhausted by functionalizable states.⁹

It is hedonistic theories that seem to present the biggest obstacle to zombie well-being [18], at least if doing well on such theories depends on the balance of pleasure of pain in one's life and pleasure and pain contribute to well-being at least partly in virtue of their phenomenology. As Kriegel notes, however, as a strategy to vindicate phenomenal consciousness, the appeal to hedonism faces two obstacles. First, hedonism is seen as unattractive to many philosophers due to its apparent inability to handle objections from the experience machine. Second, some philosophers argue that hedonistic well-being does not depend on phenomenology [14]. Assessing these theories is beyond my competence, but it's worth noting that it's not enough to show that phenomenology makes some contribution to well-being to show that zombies lack moral status: it's necessary to show that non-phenomenal properties make no contribution at all. That's a claim that Kriegel. Siewert and O'Leary all embrace. Kriegel asks us to imagine that we discover a planet inhabited by intelligent robots: would it be wrong to dissect a small robot child even as its mother pleads with you to let it go? Kriegel says "one would have to be seriously confused" to think the answer is yes (Kriegel, forthcoming). Surely no one can be entitled to this level of confidence, in the face of the variety of ways in which access consciousness seems to underwrite moral considerability, even if phenomenal consciousness is also important.

Conclusion

Let's return to the really important question: where does all this leave *me*? It's difficult to judge. It's consistent with everything I've said that phenomenology makes an important contribution to well-being, to the value of a life for the person whose life it is, as well



The view defended here has implications beyond its import for me. If phenomenal consciousness is the exclusive ground of value, then non-human animals may well be just as valuable as we are: perhaps, indeed, they're more valuable. Perhaps the complexity of our cognition displaces the centrality of phenomenality for us (think of the difference between a person who understands why the doctor is inflicting pain on her, and that the pain will be brief, and a sentient non-human animal who lacks the capacity to contextualize the pain or to take comfort in the knowledge it will soon be over; the latter might suffer more intensely because it can't understand why it's undergoing pain or that it will soon pass). It's also possible that non-human phenomenality is more intense and lively than ours, perhaps in virtue of their sensory capacities. Perhaps I'm slightly less considerable than you, but octopuses outweigh both of us. Alternatively, if phenomenal consciousness is but one contributor to moral considerability, we might fare better in situations where we're faced with trade-offs between their well-being and ours.

If access consciousness, or something like it, grounds moral status, then we're likely far closer to building morally considerable AIs then if it depends on phenomenal consciousness. It is possible, of course, that phenomenal consciousness comes along for free with certain kinds of complexity (recurrent processing; possession of a world-model, a global workspace, and so on (see [7], for a discussion of these properties and the degree to which large language models might soon instantiate them). If these properties directly ground moral considerability, however, we can avoid the epistemic obstacles that seem inevitably to confront the attribution of phenomenal consciousness. We'll probably have good reason to



⁹ Siewert suggests that phenomenal consciousness underwrites our capacity to recognize what and where things in the world are (6). The well-known neuroscientific work on the role that the dorsal and ventral visual pathways play in perception suggests this is false [13]. The dorsal stream underlies our capacity to navigate the world, but we don't have any sort of conscious experience of its contents.

Neuroethics (2024) 17:21 Page 13 of 14 21

treat AIs as moral beings soon; long before we settle debates on either the neural correlates of phenomenal consciousness or whether such correlates can have non-biological analogues.

The current fascination with consciousness centres around phenomenology and its alleged mysteries. I don't know whether it is really so mysterious. But we needn't think consciousness matters only because of how it feels. It matters regardless of how it feels, or whether it feels like anything at all. Of course I *would* think that, wouldn't I?¹⁰

Funding I am grateful to the European Innovation Council for support of this work, through the Counterfactual Assessment and Valuation for Awareness Architecture grant (CAVAA – 101071178); UKRI Reference Number 10069044.

Declarations

Competing Interests None.

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

References

- Block, N. 1995. Consciousness. In A Companion to the Philosophy of Mind, ed. S. Guttenplan, 210–219. Oxford: Blackwell Publishers.
- Block, N. 1995. On a confusion about a function of consciousness. Behavioral and brain sciences 18: 227–247.

- 3. Boonin, D. 2019. *Dead Wrong: The Ethics of Posthumous Harm.* New York, NY: Oxford University Press.
- Bradford, G. 2023. Consciousness and Welfare Subjectivity. Noûs 57: 905–921. https://doi.org/10.1111/nous. 12434.
- Breen, N., M. Coltheart, and D. Caine. 2001. A two-way window on face recognition. *Trends in Cognitive Sciences* 234–235. https://doi.org/10.1016/S1364-6613(00) 01659-4.
- Brigard, F.D. 2010. If You Like It, Does It Matter If It's Real? *Philosophical Psychology* 23: 43–57. https://doi. org/10.1080/09515080903532290.
- Chalmers, D.J. 2023. Could a Large Language Model Be Conscious? *Boston Review*, 9 August. https://www.bosto nreview.net/articles/could-a-large-language-model-beconscious/.
- Chalmers, D.J. 1995. Facing up to the problem of consciousness. *Journal of consciousness studies* 2: 200–219.
- Coltheart, M., and M. Davies. 2022. What is Capgras delusion? Cognitive Neuropsychiatry 27: 69–82. https:// doi.org/10.1080/13546805.2021.2011185.
- Crisp, R. 2021. Well-Being. The Stanford Encyclopedia of Philosophy, ed. Zalta EN. https://plato.stanford.edu/entri es/well-being/.
- Cronin, T.W., R.L. Caldwell, and J. Marshall. 2001. Tunable colour vision in a mantis shrimp. *Nature* 411: 547–548. https://doi.org/10.1038/35079184.
- Frankfurt, H.G. 1988. The Importance of What We Care About: Philosophical Essays. Cambridge: Cambridge University Press.
- Goodale, M., Milner, D. 2013. Sight Unseen: An Exploration of Conscious and Unconscious Vision. Oxford: Oxford University Press.
- Heathwood, C. 2007. The Reduction of Sensory Pleasure to Desire. *Philosophical Studies* 133: 23–44. https://doi. org/10.1007/s11098-006-9004-9.
- Jaynes, J. 2000. The Origin of Consciousness in the Breakdown of the Bicameral Mind. Boston: Houghton Mifflin.
- Kahane, G., and J. Savulescu. 2009. Brain Damage and the Moral Significance of Consciousness. *Journal of Medicine and Philosophy* 34: 6–26. https://doi.org/10.1093/ jmp/jhn038.
- Kammerer, F., 2022. Ethics Without Sentience: Facing Up to the Probable Insignificance of Phenomenal Consciousness. Journal of Consciousness Studies 29, 180–204. https://doi.org/10.53765/20512201.29.3.180
- 18. Kriegel, U. 2019. The Value of Consciousness. *Analysis* 79: 503–520. https://doi.org/10.1093/analys/anz045.
- Kriegel, U. 2018. Dignity and the Phenomenology of Recognition-Respect. In *Emotional Experience: Ethical and Social Significance*, ed. J.J. Drummond and S. Rinofner-Kreidl, 121–136. London: Rowman & Littlefield International.
- Kriegel, U. forthcoming. The Value of Consciousness to the One Who Has It. *The Importance of Being Conscious*, eds. Lee G, Pautz A. Oxford University Press.
- Levy, N. 2014. The Value of Consciousness. *Journal of Consciousness Studies* 21: 127–138.
- Levy, N. 2014b. Consciousness and Moral Responsibility. Oxford: Oxford University Press.



¹⁰ I am grateful to two anonymous reviewers for Neuroethics for encouraging and helpful comments. I am also grateful to the audience at the conference Consciousness: An Interdisciplinary Perspective, held at the University of Oxford in April 2023, and to Mattia Cecchinato for useful feedback. This work was supported by a grant from the European Research Council for the project Counterfactual Assessment and Valuation for Awareness Architecture, UKRI Reference Number 10069044.

21 Page 14 of 14 Neuroethics (2024) 17:21

23. Nagel, T. 1974. What is It Like to Be a Bat? *Philosophical Review* 83: 435–450. https://doi.org/10.2307/2183914.

- Nozick, R. 1981. Philosophical Explanations. Harvard University Press.
- O'Leary, D. 2021. The Value of Consciousness in Medicine. In Oxford Studies in Philosophy of Mind, vol. 1, ed. U. Kriegel, 65–85. Oxford: Oxford University Press.
- Rose, S. 2002. Keep these in mind. New Scientist 2364, 12 October.
- Rowland, R., 2017. Our Intuitions About the Experience Machine. Journal of Ethics and Social Philosophy 12, 110–117. https://doi.org/10.26556/jesp.v12i1.216
- 28. Segal, C. 1983. Kleos and Its Ironies in the Odyssey. L'Antiquité Classique 52: 22–47.
- Shepherd, J. 2024. Sentience, Vulcans, and zombies: The value of phenomenal consciousness. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01835-6.
- Shepherd, J. 2023. Non-Human Moral Status: Problems with Phenomenal Consciousness. *American Journal of Bioethics Neuroscience* 14: 148–157. https://doi.org/10.1080/21507740.2022.2148770.

- 31. Shepherd, J. 2018. *Consciousness and Moral Status*. Routledge.
- Siewert, C. 2021. Consciousness: Value, Concern, Respect. In Oxford Studies in Philosophy of Mind, vol. 1, ed. U. Kriegel, 3–40. Oxford: Oxford University Press.
- 33. Siewert, C. 1998. *The Significance of Consciousness*. Princeton University Press.
- Takagi, Y., and S. Nishimoto. 2022. High-resolution image reconstruction with latent diffusion models from human brain activity. https://doi.org/10.1101/2022.11.18. 517004.

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

