

The Mind that Matters

MATTIA CECCHINATO

University of Oxford, UK

What kind of mind is necessary and sufficient for moral status—for having interests that matter morally in and of themselves? This paper defends Affective Sentientism, the view that moral status requires the capacity for affective experiences such as pleasure, pain, and emotion. I argue that affective consciousness is what makes an entity a welfare subject, and that all and only welfare subjects have moral status. The connection between affect and welfare is systematic: I show that every major candidate welfare good has an affective component, often because such goods require the capacity to care. I also argue that even if the properties of being phenomenally conscious, being autonomous, or having a neutral level of welfare could come apart from affective consciousness, none of them could ground moral status in its absence. I conclude by exploring the implications for the moral status of animals and artificial intelligence systems.

Keywords: Moral status; Consciousness; Affect; Sentientism; Theories of welfare; Philosophical Vulcans; Animals; Artificial intelligence.

I. Introduction

Imagine yourself in a dreamless sleep. You don't think, see, or feel anything. There is only an impenetrable darkness, if even that. The prospect of this state, however, doesn't worry you: you know that you will later regain *phenomenal consciousness*—that colourful stream of experience that feels like something from the inside. But suppose you lacked the capacity for consciousness, had never had it, and never would. Then, one might think, your existence would not be importantly different from that of a rock or a fungus. It would seem *worthless*—not in the sense of being useless, but in the sense that nothing could be intrinsically good or bad for you, and others would see no reason to care

Correspondence to: Mattia Cecchinato, mattia.cecchinato@philosophy.ox.ac.uk

about you for your own sake.¹ Indeed, for the one who lives it, a life immersed in that impenetrable darkness might seem just as good as not existing at all.

The underlying intuition is that the capacity for some form of consciousness seems *necessary* for having interests that matter morally in and of themselves—that is, for being an entity with *moral status*. Some form of consciousness also seems *sufficient* for such status, since at least all beings capable of experiencing suffering plausibly deserve moral consideration. *Sentientism* is the view that captures these two intuitions: the capacity for consciousness, in some form, is both necessary and sufficient for moral status.

Let me take Sentientism for granted. Though contested, it is arguably the dominant view about the grounds of moral status in contemporary ethics.² In this paper, I focus on its precise form. As defined above, Sentientism takes no stance on which kind of conscious experience is required for moral status. But under the umbrella of consciousness there are many different experiential properties: from the euphoric character of a hearty laugh to the felt emptiness of loneliness, and from the striving character of experiences of effort to the felt spark of understanding something new. If these features of consciousness can come apart, so that some remain while others are permanently lost, then what would have to be missing for a life to lack moral status? Would it have to be consciousness altogether, or would it be enough to lose a particular kind of consciousness? Which is the mind that matters?

The traditional answer is *affective consciousness*: conscious experience that feels good or bad. As I use the term, affective experience encompasses states with a positive or negative hedonic valence, such as pleasure and suffering, as well as emotions and desires that involve such hedonic states. This subset of phenomenal consciousness is sometimes called *sentience*.³ But since there are other ways of spelling out Sentientism, I will use the term *Affective Sentientism* for the view that all and only beings capable of affective consciousness have moral status.⁴

Let us remain neutral on how the notion of capacity is to be understood. It is clear why the capacity for affect suffices for moral status: if there are moral reasons to care about others at all, it is hard to deny that the possibility of causing them to feel joy or suffering warrants at least *some* moral consideration. What is less clear is why affect is necessary—why we should restrict our

¹I assume that fungi lack intrinsic moral significance insofar as they are not conscious. Proponents of biocentrism will disagree; for a critique of biocentrism, see Basl (2019).

²Defenders of Sentientism include Bentham (1907), Singer (1990), DeGrazia (1996), Siewert (1998), McMahan (2002), Korsgaard (2018), Shepherd (2018), Powell and Mikhalevich (2020), van der Deijl (2021), Lin (2021), Lee (2022, 2025), Chalmers (2022), Roelofs (2023), Perry (2024), Kriegel (forthcoming), and Smithies (2026).

³Singer (1990: 8–9) and Birch et al. (2021: 7). Others, such as Kagan (2019: 13), use ‘sentience’ to mean phenomenal consciousness simpliciter.

⁴Bentham (1907), Singer (1990), DeGrazia (1996), Shepherd (2018), Powell and Mikhalevich (2020), and Smithies (2026).

circle of moral concern to *affective beings*, that is, beings capable of affective consciousness.

Indeed, a recent trend has been a move towards more inclusive forms of Sentientism.⁵ Some have proposed that all phenomenally conscious beings have moral status, and that even a simple capacity to experience light would suffice for it. I call this view *Phenomenal Sentientism*.⁶ Others maintain that moral status requires any kind of consciousness that can provide a being with a reason for action—a view I call *Motivational Sentientism*.⁷

The general reason to reject an affective restriction on moral status is that it has consequences many find unsettling. Consider a *Philosophical Vulcan*—a highly sophisticated, phenomenally conscious creature without the capacity for affective feelings.⁸ The idea is inspired by the fictional species in *Star Trek*. But while *Star Trek*'s Vulcans occasionally feel emotions, Philosophical Vulcans are by stipulation completely incapable of positive or negative affect (hereafter, this is what I mean by 'Vulcans'). Even so, they make plans, form intentions, and act autonomously, guided by their own rational self-government. They also get all that non-affective consciousness has to offer: experiences of colours, freedom, intentionality, understanding, and perhaps even puzzlement, selflessness, and agency. They possess all the potentially morally relevant capacities an entity can have without affect. It is therefore natural to ask: Wouldn't they have their own interests and ambitions? Wouldn't it seem wrong to kill one for no good reason, even if many others could take its place? Many philosophers have the strong intuition that Vulcans would deserve moral consideration more or less as we do.⁹ This is because, despite their lack of affect,

[Vulcans] might still have a rich conscious life, with multimodal sensory experiences and a stream of conscious thought about all sorts of complex issues ... They may want to advance science ... and to help those around them ... They experience no pleasure when anticipating or achieving these goals, but they value and pursue the goals all the same.

Chalmers (2022: 336)

It may turn out that such creatures are not merely conceivable but actually possible. In existing humans and animals, phenomenal and affective consciousness tend to go together. In the future, however, what is known about

⁵Chalmers (2022, forthcoming), Birch (2022), Roelofs (2023), Shepherd (2024), Kriegel (forthcoming), and van der Deijl (2026).

⁶Chalmers (forthcoming) is inclined to accept this view.

⁷Roelofs (2023).

⁸The Vulcan case is introduced by Chalmers (2022: 335–7). For earlier examples, see Carruthers's *Phenumb* (1999: 478–80) and Crisp's anhedonic being (2006: 122).

⁹Chalmers (2022, forthcoming), Birch (2022), Roelofs (2023), Mogensen (2025), van der Deijl (2026).

conscious experience could be operationalized in a functional description of conscious processing that abstracts from hedonic experience, and could then be implemented in artificial systems. These would be *Artificial Vulcans*: sophisticated silicon-based creatures who could reflect on the intricate movements of the wind, sense an approaching storm through their sensors, and form the intention to seek shelter—all without ever feeling the unpleasantness of a shiver. Should we include such beings in our circle of moral concern when the time comes?

In this paper, I examine what it would be like to live a sophisticated conscious life entirely devoid of affective experience. After sustained argument and reflection, I conclude that Affective Sentientism has been right all along. It is the Vulcan who matters less than we might have thought.

My defence of Affective Sentientism rests on some more fundamental claims about the nature of welfare. I show that every major candidate welfare good and bad is best understood as having an affective component. In particular, many such goods and bads require the capacity to care, which essentially involves affective experience. Because all and only beings capable of welfare goods and bads are welfare subjects, I argue that affective consciousness is what makes an entity a welfare subject. And because all and only welfare subjects have moral status, I argue that affective consciousness is what makes an entity a bearer of moral status. Along the way, I explain why the properties of being phenomenally conscious, being autonomous, and having a neutral level of welfare could not ground moral status in the absence of affective consciousness. I close by exploring the implications for the moral status of animals and artificial intelligence systems.

II. The Argument for Affective Sentientism

Why think that affective consciousness grounds moral status? The argument I defend is this:

- (P1) All and only beings whose interests matter morally in and of themselves have moral status.
- (P2) All and only welfare subjects have interests that matter morally in this way.
- (C1) Therefore, all and only welfare subjects have moral status.
- (P3) All and only affective beings are welfare subjects.
- (C2) Therefore, all and only affective beings have moral status.

Let me consider each premise in turn. The interest-based definition of moral status (P1) is the one sentientists typically endorse, and I'll take it for granted.

On this view, moral status is distinct from intrinsic moral value.¹⁰ Cultural symbols, languages, and artworks might be intrinsically morally valuable—there may be moral reasons not to damage or destroy them, independently of any effects on anything else—but they do not have moral status, because they do not have interests. An entity has interests just in case there are ways in which things can go better or worse for the entity itself, either relative to what makes its life go well or relative to what matters to it. Cultural symbols, languages, and artworks have neither lives that can go well for them nor things that matter to them, and so no moral reasons concerning them can be grounded in their interests.

Now, which entities have interests that matter morally in and of themselves, and thus moral status? A good answer is given by (P2): all and only entities with *welfare interests*, that is, interests in receiving benefits and avoiding harms. Such entities are *welfare subjects*: entities capable of having positive or negative components of *welfare*, or *wellbeing*, or whatever we call the property that makes a life good or bad *for* the subject who lives it. As I'll argue later, welfare subjects are entities capable of more than a merely neutral level of welfare. They can be bearers of *welfare goods and bads*: the states or conditions that are intrinsically good or bad *for* a subject and that constitute the positive and negative components of welfare (standard examples of which are pleasure and suffering, the fulfilment and frustration of desires, knowledge and ignorance, and virtue and vice).

It is nearly universally accepted that if you have welfare interests, you have interests that matter morally in and of themselves, and thus that *all* welfare subjects have moral status.¹¹ For the moment, I also take for granted that *only* welfare subjects do. There may be morally significant interests that are not grounded in welfare. Later I'll argue, however, that only welfare subjects can have them, because the same properties required for welfare interests are also required for non-welfare interests to matter morally. Hence, welfare subjectivity is both necessary and sufficient for moral status.

Given these modest assumptions (P1 and P2), the question of which entities have moral status reduces to the question of which entities are welfare subjects. The central thesis of this paper is that *all and only affective beings are welfare subjects* (P3). Given the definition of welfare subjects introduced earlier, this is equivalent to saying that all and only affective beings can be bearers of welfare goods or bads.

Since (P3) is the most contentious premise in my argument for Affective Sentientism, I'll devote most of what follows to defending it. I'll argue

¹⁰Kamm (2007: 229), Chalmers (2022: 335–9).

¹¹For discussion, see Lee (2022). Some theorists, such as Basl (2019: 161–82) and Kagan (2019: 31), deny that all welfare subjects have moral status, on the grounds that plants and mindless artefacts can be benefited or harmed in ways that do not matter morally. But I take it that when a flower is cut or a car is damaged, we speak of 'harm' only in a loose or figurative sense, not in the technical sense relevant to welfare.

that every plausible welfare good or bad requires the capacity for affective experience, and thus that only affective beings are welfare subjects.¹² Moreover, because the capacity for affective experience entails the capacity to have at least one welfare good or bad, every affective being is a welfare subject. It seems, indeed, that at least one of the following is true: that pleasure and suffering themselves are welfare goods and bads, or that any being capable of experiencing them can desire to have or avoid them in such a way that satisfying or frustrating those desires constitutes a welfare good or bad. All and only affective beings, then, are welfare subjects.

I call this thesis

The Affect-Welfare Link. The capacity for affective consciousness is necessary and sufficient for being able to have welfare goods or bads.

I call the deeper thesis that best explains this link

The Affective Thesis. The capacity for affective consciousness is what makes an entity a welfare subject.

In making my argument, I'll follow a methodology that, I think, should guide the search for any welfare good. Call it the *Welfare Good Extraction Method*. We begin by identifying something—say, hanging out with friends—that seems intuitively good for the person whose life includes it. We then try to isolate what matters from other factors and, from patterns in these intuitive judgments, we generalize that a certain thing—say, friendship—is itself a welfare good. But what happens if we strip away some property typically bound up with that candidate welfare good—say, the joy of seeing our friends? We cannot assume that what's left is still good. To do so would be to commit what I call the *Extraction Fallacy*: treating a modified version of φ as if it automatically inherits the welfare value of the original φ . There may be forms of joyless friendship, but for all we know some shared joy may have been essential to its goodness. We therefore need to retest the intuitive plausibility of joyless friendship as a welfare good. Should it lose its appeal, we would have reason to think that the joy we removed from friendship was essential to its being a welfare good. That intuitive appeal was, after all, the reason we considered friendship a welfare good in the first place.¹³

We could avoid the Extraction Fallacy if, whenever we formed the intuition that φ is good, we did so by grasping its *essence*: what is necessarily present in

¹²See Mogensen (2025: 346–52) and Siewert (manuscript) for similar lines of argument, though the details differ in important ways.

¹³Versions of the Extraction Fallacy can be found in Levy (2014), Kagan (2019: 33), Bradford (2022), Shepherd (2023), and Goldstein and Kirk-Giannini (2025). Though these authors focus on phenomenal consciousness rather than affect specifically, the argumentative pattern is the same. They argue that welfare goods can be instantiated without some property they are typically bound up with, but either do not address whether what remains is still intuitively a welfare good, or their attempts to do so are, in my view, unsuccessful.

every instance of φ . But this is not how theorists typically form intuitions about welfare goods—partly because the essence of φ is often unclear (what is the essence of friendship?), and partly because our intuitions about the goodness of φ can be shaped by our ordinary conception of φ , which may include properties that are not strictly necessary for φ to be instantiated, yet are necessary for φ to count as a welfare good. That is, it is entirely possible that whenever we speak of φ as a welfare good, we really mean φ *plus* a set of properties that ordinarily come with φ . Our ordinary conception of friendship, for example, might include some form of shared joy, and it may be friendship so understood that we have in mind when we judge friendship to be a welfare good, regardless of whether friendship as such necessarily requires shared joy.

The Welfare Good Extraction Method then works as follows:

- I. Start from intuitive judgments about cases.
- II. From stable patterns in these intuitions, isolate what matters and extract candidate welfare goods.
- III. When a property typically accompanying a candidate welfare good is removed, test whether the original intuitions hold up.
- IV. If they don't, the candidate welfare good should be understood as requiring that property.

Following this method, I'll argue that every welfare good has an affective component, because whenever we remove it from a candidate welfare good, that good loses its intuitive appeal.

Consider first pleasure and pain themselves. I remain neutral on whether these states are essentially attitudes, feelings, or both. In my view, however, for them to count as welfare goods or bads, they must involve pleasant or unpleasant conscious experience.¹⁴ There may be a coherent notion of pleasure and pain without consciousness, perhaps understood as internal states that causally explain behaviours such as approach and avoidance in response to bodily benefit or damage. But without further argument, the existence of non-conscious hedonic states wouldn't entail an expansion of our welfare goods and bads. To assume otherwise would be to commit the Extraction Fallacy, since our concepts of pleasure and pain as *welfare goods and bads* may always involve conscious experience. Indeed, these are the pleasures and pains hedonists typically have in mind: conscious states that feel good or bad.¹⁵ What's more, if we imagine having non-conscious hedonic states with *no effect at all* on our conscious experience, it doesn't seem that we would thereby be made intrinsically better or worse off. The lack of a clear intuition

¹⁴Some pains, such as those in cases of pain asymbolia, are reported not to hurt (Grahek 2007). If these are pains at all, they may lack an affective component and so may not be the kind of pain that on my view is a welfare bad—namely, pain that feels unpleasant.

¹⁵Crisp (2006), Bradley (2009), Bramble (2016).

here should caution us against valuing pleasures and pains that do not involve consciousness, if any exist.

There are also candidate welfare goods whose goodness does not depend solely on their affective character, such as desire-fulfilment, knowledge, virtue, love, friendship, and achievement. Although I'll discuss these goods extensively, it's helpful to outline my argument early on. I'll argue that even when these goods may not seem to rest on pleasure or suffering, an affective component is in fact essential to all of them in every instance. They require, at the very least, the subtler affective feelings involved in *caring about* someone or something.

The basic idea is that caring, in the relevant sense of the term, requires complex emotional states and dispositions partly constituted by episodes of affect. A central component of caring is a disposition to feel joy or satisfaction when the object of one's care is doing well, sadness or frustration when it is doing badly, fear when it is in danger, and relief when it turns out to be fine. Even when nothing actually happens to the object of one's care, one may be disposed to feel positive or negative affect *at the mere thought* of these possibilities. Some of these affective states may be absent without precluding care. But if all of them are missing, then, I'll argue, we are no longer talking about caring in the sense relevant to welfare.

I'm not sure whether caring is necessary for all welfare goods and bads. Perhaps it is possible to imagine a subject enjoying a stream of pleasant experiences despite lacking the psychological capacities needed to care about them in the relevant sense. A newborn infant may be an example: it may feel pure joy without being able to care about that joy. Yet such joy would still seem good for it. My claim, then, is that all welfare goods and bads require some affective component—either the complex affective states involved in caring or simpler hedonic states. And since at least one welfare good might not require caring, and the capacity for welfare goods suffices for morally significant interests, I'm inclined to believe that it is the capacity for affect, rather than caring, that grounds moral status.¹⁶

The role I give to affect respects the idea that every basic component of welfare is appealing, engaging, and shouldn't leave one cold.¹⁷ However, I do not rely on the intuitive appeal of such a *resonance constraint* on wellbeing as a ground for giving affect this role. Other theorists appeal to the requirement that what is good for someone shouldn't leave them cold, and argue that such engagement requires the heat of affective feelings.¹⁸ My argument is different, partly because it's not clear to me whether the resonance constraint *per*

¹⁶While Jaworska (2007) argues that caring grounds full moral status, or personhood, I remain neutral on that question. My concern here is only with which kind of mind is necessary and sufficient for moral status simpliciter.

¹⁷Railton (1986: 9).

¹⁸Fanciullo (2025), Smithies (2026).

se is strongly intuitive as a basic principle about welfare, especially given the long-standing debate between subjectivist and objectivist theories of welfare. What strikes me as strongly intuitive is that affect is required for each particular candidate welfare good. It is from this intuitive data about particular goods that one should motivate a general resonance constraint, not the other way around.

It might be thought that we need a more unified explanation for the systematic correlation between affect and welfare goods. But for any given theory of welfare, the explanation is as unified as the theory allows it to be. If hedonism is true, and only one thing is a welfare good, then all welfare goods turn out to require affect for only one reason. If an objective-list theory is true, and more than one thing is a welfare good, then all welfare goods turn out to require affect for more than one reason. The right explanation need not be unified insofar as the basic nature of welfare is not unified. In fact, in that case the lack of unity would not be a flaw in my explanation, but a feature of the phenomenon I am explaining.

I begin by showing why desire-fulfilment requires affect, and then turn to other goods.

III. Vulcans, Desires, and Caring

Many think it is intrinsically good for us if our desires are *fulfilled*—if the world matches how we want it to be—and intrinsically bad if they are *frustrated*. To avoid counterexamples, however, wellbeing theorists tend to qualify which kinds of desires matter for welfare. I'll argue that the right kind of desire necessarily involves affect.

An initial question is whether desires are possible at all without affective consciousness. For us, desiring something always comes with a disposition to feel excitement or satisfaction if we have what we desire, or imagine having it, and a disposition to feel disappointment or even suffering if we do not have it, or imagine not having it. There might indeed be a necessary connection between having a desire and having such affective dispositions.¹⁹ But for present purposes, I'll grant that this connection need not hold, at least in alien minds. After all, some evidence suggests that the neural circuits of 'liking' (that is, enjoyment) can dissociate from those of 'wanting' in pathological cases.²⁰

What would an affectless desire look like? Vulcans might intentionally set goals for themselves by representing certain ends as positive ('to-be-pursued') or negative ('to-be-avoided') without any feeling of pleasure or displeasure. They may want to perfect their own rational nature, better understand the

¹⁹ Mill (1998: Ch. 4).

²⁰ Berridge and Robinson (1998).

world, or help others. They would take no pleasure in any of these things, but might still judge them to be worth pursuing and pursue them entirely out of duty.

Vulcans, indeed, would seem capable of value judgments. They could judge a painting to be good by common aesthetic standards, even though the experience of it does not feel good to them. They might have the capacity for *cognitive valence* (thinking ‘thumbs-up’) without having that for *affective valence* (feeling ‘thumbs-up’). And if they can judge a goal to be valuable, they may be motivated to pursue it for that very reason.²¹ This would be an instance of *affectless desire*—perhaps the desire to create a good painting just because it seems valuable and worthwhile, no matter how it feels. Would the frustration of such desires harm a Vulcan?

Admittedly, I’m not sure whether dispassionate motivational states would qualify as genuine desires, or whether they would rationally justify action.²² But one thing I’m sure about: affectless desires, if they are desires at all, would not be the paradigmatic kind of desire we encounter in ordinary life. We would therefore risk committing the Extraction Fallacy if we simply took for granted that they retain the welfare significance characteristic of the paradigmatic kind. Instead, we should ask whether it is still intuitively plausible to regard them as relevant to welfare.²³ I’ll make two claims to that end. First, I’ll note that, intuitively, only desires about things we care about matter for welfare. Second, I’ll argue that, without affect, an entity may act in apparently self-interested ways, but it can neither genuinely care about its goals nor about itself.

Philosophers can use ‘caring’ however they like, but in the way I use it—which seems the most natural—if a subject cares about something, then that thing matters to them, at least to some extent, from their own point of view. Conversely, if something matters to a subject from their own point of view, then they care about it, at least a bit. And if something does not matter to a subject in that way, then they do not care about it (and vice versa). The notion of caring I’m interested in is thus closely related to that of *subjective importance*; I use the terms interchangeably.

For something to be subjectively important to me, as I understand the notion, it is not enough that I believe it is important or take it to be valuable for me. I can believe that making a great sculpture would be something important and valuable for me to do, without caring at all about making one myself or

²¹Chalmers (2022, forthcoming).

²²See Smithies (2026).

²³Chris Heathwood (2019) makes a similar move. After distinguishing between *behavioural desires*, which are intentional states that incline one to act towards a goal, and *genuine-attraction desires*, which involve one’s excited attraction towards the goal, he argues that only the latter are relevant to welfare. My argument focuses on the relations among desire, caring, and affect, but it is compatible with Heathwood’s if all genuine-attraction desires require caring and affect.

being moved by the idea. Subjective importance essentially involves, among other things, a certain kind of phenomenology directed at a particular object, such that the object matters to the subject partly in virtue of how it makes them feel (more on this shortly).

This sort of mattering is also bound up with a set of desires about the object: a desire for it to go well, to last, to be present in the subject's life, and so forth. However, we should be careful to distinguish these deep motivational states from shallower ones lacking any genuine subjective importance. The distinction is pervasive in ordinary life and is captured by remarks like 'But I don't really care about it,' made after expressing a preference.²⁴

Consider, for example, an art collector who excels at recognizing artworks of great aesthetic value. Every week he buys extraordinary new pieces, yet resells them immediately, no matter how beautiful they are. We ask him, 'Why not keep any for yourself?' He replies: 'I think they're good, but I don't feel anything for them. They don't evoke any emotion in me, so I take it that I don't really care about them. It's just my job. And to be clear, I don't care about my job either; I'm simply doing what I'm supposed to be doing.'

We shouldn't be too surprised. Our goal-directed agency and evaluative judgments need not be accompanied by subjective concern. In fact, it seems possible to want something we judge to be valuable, in some detached, impersonal way, even if obtaining it would serve no end we care about. Unusual as this state of mind may be, some people experience something like it in certain forms of depression—not the kind that feels painful, but the kind that quietly makes everything feel empty and pointless.²⁵ The art collector might suffer from this condition. He might seem energetic and engaged, always looking for the finest artworks, always judging well, always doing his best—but if he stops for a moment, he can tell that none of it really matters to him. He is just going through the motions.

Suppose, further, that the art collector always gets what he wants. Is this good for him? Would he be harmed if someone were to interfere with his aims? Intuitively, it seems to me that he wouldn't. The collector doesn't want the things he wants in a way that puts his welfare at stake. Indeed, it seems to me that if the value of desire-fulfilment were to be understood in such a way that even the fulfilment of goals one does not care about at all could be good for one, then desire-fulfilment would turn out to be a far less plausible candidate welfare good than we ordinarily assume—less plausible, at any rate, than wellbeing theorists or ordinary people would be inclined to accept.

²⁴Frankfurt (2004: 10–12) draws a similar distinction between things we want or desire and things we care about or regard as important to ourselves.

²⁵Psychiatrists sometimes describe this as *anhedonic depression*—a 'flat' or 'numb' form of major depression characterised by diminished positive affect.

I think that Vulcans are much like this art collector. The only important difference is that, while the collector may have cared about things in the past, may do so again in the future, and presumably remains at least *capable* of caring at present, Vulcans are entirely unable to care about anything in the relevant sense. Caring may require a wide range of cognitive states and dispositions, such that whenever a subject Y cares about φ , Y's caring necessarily (i) has some weight in Y's rational decision-making, perhaps enough to warrant significant sacrifices for φ ; (ii) draws Y's attention and thoughts to φ ; (iii) motivates Y in a robust, enduring way; and (iv) involves a persisting pattern of emotional states and dispositions about φ . But what matters for our purposes is that, among all the necessary conditions for caring, one is a disposition to have affective experiences.

Consider:

The Vulcan Mother. A Vulcan mother wants her children to be healthy and goes to great lengths to take care of them. However, she does so without ever having any positive or negative feelings for them, no matter what they do or what happens to them. She is in no way affectively invested in her parental role, never has been, and never will be.

The *Vulcan Mother* coolly exhibits all the typical characteristics of care, without affect. All her thoughts and attention may be focused on her children. She may be disposed to tell herself in inner speech 'good, that worked' whenever her children succeed. Her commitment may be unwavering; she might even be willing to sacrifice her own life for them. Yet it would be somewhat odd to say that she *cares about* them. She wouldn't feel joy at their accomplishments, nor anger at those who intend to harm them, nor grief at their loss. And intuitively, if you care about someone, how their life goes should make you *feel something*. At the very least, it should dispose you to feel something positive or negative at the thought or sight of them faring well or badly. It is for this reason that I think the *Vulcan Mother* wouldn't be harmed by the loss of her children. Her willingness to make sacrifices for them wouldn't mean much, because those sacrifices themselves wouldn't matter to her. And her relentless activity wouldn't constitute deep commitment to someone she cares about, but rather detached, cold-hearted service.

Some might find it hard to say whether a Vulcan can care about anything in the sense at issue, since it is difficult to imagine the strength and character of the wants of a being whose feelings are necessarily neutral. But there is a way to get close enough. We can start from our own experience and move gradually towards the more alien, Vulcan-like experience. We should pick the closest experience we have to an affectively neutral one: perhaps the experience of staring at a blank wall, something we neither like nor dislike. Keeping this neutral tone fixed inside, we can then go on to imagine entertaining those thoughts and performing those behaviours characteristic of caring. We can

think of ways to protect the wall from falling, to paint it anew, and so forth. That is, we can use that familiar unhedonic experience to mentally reconstruct what it would be like to feel no positive or negative affect while still thinking, behaving, and setting goals as we normally do when we care about something. Even if we may never be able to reach, even in imagination, the full extent of affectless agency that might be available to a Vulcan, we can still imagine being roughly similar in the relevant respects. From this position, we can then ask ourselves: 'But do I really care about this blank wall—the one I spend so much time thinking about and taking care of, but for which I don't feel the slightest sadness, grief, fear, joy, or satisfaction, whether it falls or not?' If we were honest, I suspect we would always answer with a harsh 'No, not at all.'

Since we wouldn't care about the wall in those circumstances, we should infer that the *Vulcan Mother* wouldn't care about her children either. Her case is relevantly like ours. At any time she is with them, she will feel as cold inside as we do when we stare at the wall. And at any time we take care of the wall, we may be exercising to a large extent all the psychological and cognitive faculties she exercises while taking care of her children. Admittedly, she might be better than us in many respects. But it would be strange if a higher degree of dedication, attentiveness, or meticulousness made all the difference to whether or not caring occurs, especially considering that, in the human case, we don't need to possess those qualities to an especially high degree to care about something. It seems more natural to conclude from this case that the sense of caring at issue necessarily requires some affective investment.

Vulcans may of course reject the human intuitive judgment. They may complain that no one can claim, on their behalf, whether they care about something or not. But humans are in a better epistemic position than Vulcans to evaluate the truth of the necessary link between the relevant sense of caring and affect.²⁶ For we can *both* experience affective caring and imagine something close to Vulcan-like caring. By contrast, Vulcans may be inevitably biased. For it may be impossible, or nearly impossible, for them to learn what it is like to care about something affectively.

The above considerations support

The Caring-Affect Link. To care about something in the sense of the term that matters morally and prudentially, one must feel, or be disposed to feel, positive or negative affect at the experience of, or thought of, the object of care faring well or badly.

Put differently, the capacity to feel hedonically engaged with something is *necessary* to care about it in a way that matters. Note that such caring need not

²⁶Famously, Mill (1998: Ch. 2) makes a similar argument: in comparing the value of higher and lower pleasures, he takes the judgment of those with experience of both to be epistemically privileged.

involve a disposition to feel affect in response to the object of care *actually* faring well or badly; even just contemplating that prospect may be enough. Consider, for example, a semi-Vulcanized individual who feels nothing when he obtains his goals but feels good whenever he imagines obtaining them. In the relevant sense, he could still care about his goals.

I assume that both hedonic and emotional states can satisfy the Caring-Affect Link, as long as they have some inbuilt felt goodness or badness. Perhaps some emotions are not hedonically charged—perhaps feeling a cold, affectless calm counts as being in an emotional state. But these are not the positive and negative emotional states required for caring, nor the paradigmatic cases of emotion. Indeed, imagining the *Vulcan Mother* feeling glacially calm while her children are dying doesn't make us any more inclined to think that she cares about them or is emotionally attached to them. Vulcans, all in all, are best understood as emotionless beings, since describing them as beings with emotional capacities brings to mind the misleading picture of them experiencing anger, grief, fear, joy, and other paradigmatic emotions that seem to necessarily involve affect.

Let us take stock. I have been arguing that an adequate theory of welfare must discriminate between those desires that contribute to welfare and those that do not, and that only desires about things we care about matter for welfare. Since the sense of caring at issue requires a disposition to feel affect, any being incapable of affective consciousness cannot care about anything in the relevant sense. Thus, Vulcans cannot form desires of the right kind—those that contribute to welfare. And if nothing matters to a being, so that every choice in its life expresses only cold, emotionally detached purposes, we can reasonably surmise that its aims do not in and of themselves demand moral respect.

IV. Other Goods

Let us turn to other candidate welfare goods, such as love, friendship, virtue, knowledge, aesthetic appreciation, and achievement. Again, we are normally affectively invested in the pursuit of these things—we care about them and enjoy having them. Perhaps we can have them without affect. But we would risk committing the Extraction Fallacy if we simply assumed their goodness to be preserved. We have stripped away a potentially fundamental component from the familiar conception of these goods. Their intuitive plausibility as welfare goods therefore needs to be tested again. And once it is, I argue, these candidate goods lose all their original appeal (similar arguments can be extended to all other possible welfare goods).

Love and friendship can be treated briefly. Caring is an essential ingredient of these goods, and I have already argued that the relevant kind of caring

requires affective investment. Indeed, it is unclear what would be intrinsically good about a Vulcanized form of love or friendship, one deprived of mutual concern, renewed and deepened affection, shared sympathy, amusement, and delight.²⁷

Virtue is a more interesting case. People with damage to brain areas crucial for affective processing can make sensible moral judgments.²⁸ Similarly, a Vulcan could judge things to be good, explain what is good about them, and be motivated to act accordingly. Could such a being be virtuous?

I suspect that virtue requires more than a combination of sound moral judgment and agency. There seems to be nothing virtuous about fighting for a cause you don't care about, or helping a person whose wellbeing doesn't matter to you. More generally, without emotional attachment to sustain patterns of behaviour such as loyalty or costly aiding, it is hard to find them noble. If genuine interpersonal concern is required for virtue, and the Caring-Affect Link holds, then the virtuous person is partly distinguished by the burning of affective feelings in her.²⁹ For only such feelings can bridge the gap between a merely dispassionate value judgment and a genuine interest in others.

I am open to the Kantian line of criticism that if your primary reason for doing the right thing is grounded in sentiment rather than respect for the moral law, then your act may not be virtuous at all. But even Kantians, I think, should agree that virtuous people need to care about at least some things—such as other people, and perhaps the moral law itself—if they want their view to be intuitively plausible.

Another way to ascribe wellbeing to Vulcans is on the basis of goods such as knowledge. After all, Vulcans can learn many truths about the world and its relations, just as a scientist can understand general relativity without enjoying it. Consider the richest body of knowledge a Vulcan could have:

*Mary the Vulcan Scientist.*³⁰ Mary, the most brilliant Vulcan scientist, has dedicated eons to studying the nature of things. Despite lacking affective consciousness, she has supreme knowledge of the machinery of the universe. She knows, for example, exactly which wavelengths from the sky stimulate the human retina, and precisely how these trigger, via the brain, the contraction of vocal cords that produces the exclamation, "The sky is beautiful today!" Is this knowledge good for her?

Mary's immense knowledge is undeniably awe-inspiring. Yet I think it is greatly limited, in both breadth and depth, by her lack of affective conscious-

²⁷See the *Vulcan Mother* in Section III.

²⁸Damasio (1994), Bechara et al. (2000).

²⁹The Caring-Affect Link thus provides independent support for a sentimentalist conception of virtue (Slote 2009).

³⁰Inspired by Jackson (1982: 130).

ness. Whether she lacks a phenomenal concept, a know-how ability, or something else, what matters is that she misses a kind of knowledge important to us: she doesn't know what it is like to be in any given affective state. If we asked Mary about Shakespeare's *Romeo and Juliet*, she could offer an impeccable interpretation of the characters' reasons and motives, but she wouldn't know what it is like to love someone. She would be unable to access the affective, qualitative aspects of love from a third-person description. If we asked her about war, she could cite a multitude of historical facts and personal testimonies, but she wouldn't know what it feels like for someone to lose their friends and grieve over the thought that they could have saved them. Even if we turned to science—Mary's crowning achievement—we would find something she doesn't know. Mary herself could never know the thrill of pursuing nature's deepest mysteries, nor experience, in her discoveries, the wondrous understanding that enlightens the mind and stirs the heart. In fact, she could no more imagine what it would be like to feel this way than I could picture my daily life in twenty-seven-dimensional space.

Mary is then ignorant, in my view, of some of the most important things there are to know. Of course, she still knows many other things. But the idea that her remaining knowledge is good for her needs independent support, not an unargued-for statement that knowledge is good. After all, in the human case, coming to know something is usually accompanied by affective feelings. Enlightening intellectual moments are among the most enjoyable experiences a person can have, and we care deeply about having them. Thus when philosophers claim that acquiring knowledge is good for us, what they have in mind may not be knowledge as such, but *knowledge we enjoy or care about*.

This reading is strengthened when we test our intuitions about purely affectless knowledge. It seems odd to claim that if I were to fill your mind with knowledge about the culture of a distant species—knowledge that would never be useful to you, that you would not care about in the slightest, and that would be utterly unenjoyable to you and everyone else—you would nonetheless have reason to be grateful to me. If knowledge is good for the knower at all, and this thesis is to be intuitive, then all knowledge that is good must involve affective experience.³¹ Mary's knowledge therefore cannot be good for her.

What about aesthetic appreciation? I'll focus on beauty, though similar things can be said about other aesthetic properties. Since Mary has complete scientific knowledge of the human mind and the principles of aesthetic appreciation, she can accurately judge something to be beautiful for humans. She can also recognize when a painting satisfies widely shared norms of beauty. But can she appreciate beauty herself, in the sense that matters for welfare?

I suspect Mary would have to admit, 'I know why people find this painting beautiful, but I just don't see it myself.' For appreciating a work of art as

³¹ A view of this kind is examined by Kagan (2009).

beautiful requires more than knowing that it is beautiful. It requires experiencing the piece with affective feelings directed at its aesthetic properties, such that the experience *feels heavenly* in the distinctive way appreciating beauty *feels heavenly*. Mary could not learn what this experience feels like via introspection, and it is notoriously difficult to conceive how she could derive this first-personal knowledge from the rest of her knowledge. Were she to suddenly gain affective consciousness, she would appreciate beauty for the first time.

Let us turn to the value of achievement. Like human athletes, Vulcans can train relentlessly, master their skills to perfection, and order their lives around the pursuit of excellence. Our conception of achievement, however, is normally associated with feelings of deep satisfaction, fulfilment, or at least excitement along the way. Strip all that away. When Vulcans work towards achieving something, they don't enjoy the process, they don't feel satisfaction when they succeed, and they don't care about the overarching goal at all. To me—and, I suspect, to most people—it is doubtful that such an achievement could be intrinsically beneficial to the achiever.

Now, one might think that what makes an achievement good has nothing to do with the feelings it brings or how much the achiever cares about it, and everything to do with the value of perfecting the achiever's own nature. There is some charm to this idea: through superlative achievements, a creature can purposefully develop its capacities to the fullest, making its life excellent and fulfilling its *telos*.³²

Presumably, however, perfecting a given capacity is not always good for the one who has it. A capacity is worth perfecting for the sake of its possessor only if it has at least the *potential* to be good for them. Perfecting the capacities of a non-conscious robot, for instance, cannot be good for the robot itself, because those capacities have no such potential. So for a Vulcan to benefit from perfecting its capacities, one would first need to show that a conscious being who lacks affective faculties can still have capacities that are potentially good for it. But the Vulcan wouldn't enjoy its capacities, appreciate their beauty, or care about them at all. There seem to be no intuitive grounds for thinking that such capacities have any potential for prudential goodness. And it cannot be good for the Vulcan to perfect something that can be, at best, neutral.

V. Objections

I have argued that every promising candidate welfare good or bad has an affective component, and therefore that only affective beings can be welfare subjects and have moral status. I'll now consider four objections: that becoming

³²Bradford (2015).

a Vulcan seems better than death (V.1); that killing Vulcans seems wrong (V.2); that autonomy can ground moral status even in the absence of welfare (V.3); and that having a neutral level of welfare suffices for being a welfare subject (V.4).

V.1 Vulcanization and Death

If I had to lose, irreversibly, either my entire phenomenology or just its affective properties, which would I rather lose? Suppose everything else is equal. From the outside, I would keep behaving as I always do. But on the inside, I must choose: my whole consciousness or my affective life?

If affective consciousness is necessary for being a welfare subject, I should be indifferent between the outcomes. I wouldn't get anything good out of life either way. Indeed, I should be indifferent between these options and death, at least as long as I know that, after I'm gone, a robot will fulfil my plans and take care of my loved ones as well as I would have done. For some, this will be hard to believe.³³ And, not unreasonably, they could take their intuitive response at face value—as evidence that consciousness itself is a welfare good, or at least enables one to be a welfare subject for whom things other than consciousness can be intrinsically good.³⁴

By contrast, I suspect this disbelief rests on a failure to imagine sufficiently well what it would be like to be Vulcanized. This is understandable. People report having positive or negative feelings or moods at almost any moment, so it is difficult for us to imagine life without them.³⁵ As I suggested earlier, the best way to imagine such a life is to examine your own phenomenology and pick an experience that feels as neutral as possible, one with negligible or non-existent affect. For me, this is staring at a blank wall. You can then imagine going on doing things—perceiving colours, forming thoughts, and moving your body—while retaining the same affective tone on the inside, such that even during what are supposed to be the most important moments of your life, you feel as you would feel staring at a blank wall.

When I imagine a life like this, I think I might still be doing some things. I might remember my past because I need to, take care of my loved ones because I have to, and write about philosophy because I'm used to doing so. But I would no longer care about any of these things. For whatever happens, however wonderful or awful, it would not produce in me the smallest flicker of happiness or sadness, excitement or distress. My world would be cold, emotionally detached from everything and everyone. It's not hard to believe, at

³³Kriegel ([forthcoming](#)).

³⁴Against the intrinsic value of consciousness, see Lee ([2019](#)).

³⁵Lucas and Diener ([2008](#): 796).

least for me, that I would gain nothing from my days, and that the thought of continuing to exist would itself lose its appeal.

V.2 Killing Vulcans

If I could kill a Vulcan for a small personal gain, doing so might seem morally wrong. It might seem wrong even if there were no instrumental reasons to keep the Vulcan alive, as plenty of others could take its place. On this basis, some philosophers are inclined to conclude that Vulcans have moral status.³⁶

Admittedly, I once shared the intuition that it would be wrong to kill a Vulcan. But after reflecting more carefully on what it is like to be one, that intuition has vanished. I now see that there would be nothing good for a Vulcan in continuing to exist. My earlier intuition was a result of my imagining Vulcans as more like us than, in fact, they are.

If your intuition hasn't vanished, you might reconsider whether it really tracks moral status. Perhaps killing a Vulcan seems wrong for the same reason that destroying a beautiful artwork seems wrong: some things might have impersonal value even without moral status.³⁷ Or perhaps we should expect people to be reluctant to kill Vulcans in much the same way that research participants are reluctant to drink Mountain Dew from a bedpan. We habitually associate complex intentionality with affective consciousness, so our intuitions may well misfire when confronted with alien affectless minds. Finally, even if Affective Sentientism is true—and you may be largely persuaded that it is—the arguments I've given may not make you *certain* of it. And under significant normative uncertainty, one ought not to kill a Vulcan for a trivial end.³⁸

V.3 Autonomy and Caring

So far, I have assumed that only welfare subjects have interests worth moral consideration. On some views, however, beyond an interest in welfare, we also have an interest in *autonomy*—in guiding our lives according to our own will and rational reflection. It often seems wrong, for example, to intervene paternalistically in ways that would foreseeably improve someone's welfare but disregard their autonomy.³⁹ Once we allow that some interests are independent of welfare, it stands to reason that the properties responsible for those interests might be instantiated in an individual even when the properties that

³⁶Chalmers (2022: 335–9), Birch (2022: 5), Mogensen (2025).

³⁷Smithies (2026: 312).

³⁸For a probabilistic precautionary approach to cases of uncertain moral status, see Sebo (2025).

³⁹Enoch (2016).

make that individual a welfare subject are absent. Some autonomous agents, then, might have moral status without being welfare subjects.

Vulcans could be an example of such beings.⁴⁰ Although they cannot be harmed or benefited, perhaps they can still be wronged by acts that interfere with their rational self-governance; hence killing one would be wrong.

Let us grant that Vulcans might be autonomous agents. I nonetheless see a problem: not every form of autonomy seems worthy of respect. To my mind, an agent's autonomy deserves respect only when the agent *cares about* the goals they set for themselves.

Consider, for example, asking a friend whether he would like to order a pizza margherita or a pizza carbonara. He replies, unexcitedly, 'This choice doesn't matter to me in the slightest.' You know he isn't joking. He still picks the carbonara—partly out of habit, partly because he mildly prefers the taste—but insists he wouldn't be the least bit bothered if he ended up with a margherita. 'In fact, I wouldn't even care if you ignored my preference,' he adds. To his surprise, this is precisely what happens. You order margheritas for both because doing so makes it cheaper for whoever ends up paying. You explain to your friend that you didn't mean to disrespect his general autonomy, which certainly deserves consideration on other occasions. You simply noticed that, on this occasion, he didn't care at all. Have you wronged your friend? It doesn't seem so. Your interests can easily override such shallow preferences. Generalizing from this case, it seems that there is no moral reason to respect others' autonomy except insofar as their choices express preferences they care about.

Of course, humans usually care about their preferences. But I have shown that only affective beings can do so in the relevant sense. Only such beings, then, can exercise their autonomy in a way that deserves respect. And since all affective beings are welfare subjects, it follows that only welfare subjects have moral status—not because all morally relevant interests are grounded in welfare, but because affective consciousness is necessary for both welfare-based and autonomy-based interests. This also strikes me as an intuitively compelling result in its own right: I don't see anything morally wrong in overriding the self-given ends of a creature to which nothing feels good, valuable, or important.

V.4 Neutral Welfare and Welfare Subjects

I have argued that nothing can be good or bad for Vulcans, and that they are not welfare subjects for this reason. Even so, it still makes sense to ask how desirable it would be to be one. We can imagine taking up a Vulcan's conscious perspective in order to evaluate whether such beings are better or worse off than we are. If such comparisons are possible, then Vulcans must have a level

⁴⁰This is Mogensen's view (2025).

of welfare: a neutral or zero level. For we could not be better or worse off relative to beings whose welfare level is totally unspecified. But now it seems we have a problem. What if to have a welfare level just is to be a welfare subject?⁴¹ Then Vulcans would turn out to be welfare subjects after all. They would in fact have a welfare level of zero rather than none.

I think we should resist this conclusion. If even entities whose welfare level can only be zero count as welfare subjects, it becomes hard to see what work the notion of welfare subject does in ethics. There is nothing at stake, morally, in the case of entities that cannot have welfare goods or bads but can have a neutral level of welfare. They are not morally different from entities that lack welfare altogether. After all, there is no important difference between a being who can hear only silence and one who cannot hear at all—either way, they miss out on the music.

Because I care about maintaining a robust link between welfare subjectivity and normative ethics, I restrict the notion to those entities that can have welfare goods or bads. Of course, one might also go the other way and treat welfare subjectivity as a non-moral property that some beings possess despite lacking moral status. But on this view there would still be a subset of welfare subjects—the *morally significant* welfare subjects—who do have moral status. My arguments could then be taken to support the thesis that these morally significant welfare subjects are all and only affective beings.

Now, it is not obvious that beings whose welfare level is fixed at zero do not matter morally. They might matter in virtue of their beauty, their usefulness, or other contingent features. What I want to deny is that they matter *in virtue of their having a neutral level of welfare*. Let me defend this claim against some challenges.

First, consider *Net-Zero*: a being whose welfare goods and bads happen to be perfectly balanced at every moment, so that its overall welfare level is always neutral.⁴² One might take *Net-Zero* to demonstrate that having a neutral welfare level suffices for having morally relevant interests. But *Net-Zero* is a welfare subject because it can—and in fact does—have welfare goods and bads. It is that capacity, not the neutral welfare level itself, that grounds its morally relevant interests and specifies their content.

Second, consider a Vulcan-like robot, conscious but incapable of positive or negative affect. Call it *Anhedonica*. Now, if we had a pill that could endow any being with the gift of positive affective experiences, wouldn't we have a strong moral reason to give it to *Anhedonica* rather than to a non-conscious robot?⁴³ And wouldn't that reason be grounded in the fact that *Anhedonica* has a strong

⁴¹van der Deijl (2021, 2026), Kriegel (forthcoming), Lee (2025).

⁴²For similar cases, see Lee (2025: 13) on *Balanced* and Kriegel (forthcoming) on *Mombies*.

⁴³van der Deijl (2026: 801).

interest in receiving the pill, given its neutral—and therefore low—level of welfare?

I see two ways of addressing this case without abandoning the idea that only affective beings have interests. One possibility is that a being can already be said to be *capable* of affect, at least in the morally relevant sense, if the conditions are such that it can easily acquire affective states. *Anhedonica* would thus be a welfare subject *before* receiving the pill and, as such, would have an interest in enjoying the experiences that the pill would provide. The non-conscious robot, by contrast, would seem to lack any such interest. For by giving it the pill, and thereby creating a new locus of consciousness, we would seem to be bringing an entirely new individual into existence (though this result will depend on the exact relation between consciousness and personal identity).

That *Anhedonica* may already be capable of affect shouldn't surprise us. Our capacities can and often do depend on external facts. We wouldn't be able to do many of the things we are said to be able to do without oxygen or gravity, and yet these are not intrinsic properties of human nature. The pill, if easily accessible, could be just like oxygen: as long as it is around, and perhaps only then, a being can be said to have certain capacities.

There is also another way to address this case. Suppose that *Anhedonica* turns out to lack the capacity for affect in any morally relevant sense of 'capacity.' In that case, I think we should simply deny that *Anhedonica* has interests. We can do so on the basis of

The Sake Condition. An entity has interests just in case it has a *sake*—that is, just in case it has a perspective from which things can be good or bad for the entity itself.

Only welfare subjects have a sake, and so only welfare subjects have interests. To be clear, receiving the pill would indeed benefit *Anhedonica*: post-pill *Anhedonica* would have higher welfare than pre-pill *Anhedonica*, and there is continuity in consciousness, so the very same individual would be made better off. What I am claiming is that pre-pill *Anhedonica* does not have an *interest* in being benefited. Even though ordinarily whenever someone is benefited they have an antecedent interest in the benefit, the two can in principle come apart. An entity can be benefited by the very event that creates its sake, and in such a case it would not have had an antecedent interest in the benefit. For one cannot have anything at stake, and thus any interest, if one does not even have a perspective from which things can be good or bad for one. *Anhedonica* will acquire this perspective only when, and if, it receives the pill. Until then, just as a rock is perspectiveless with respect to having a conscious point of view, *Anhedonica* is perspectiveless with respect to having a sake. It has a conscious perspective, to be sure—but not the kind that would make it a bearer of interests.

Of course, there is an obvious sense in which things *can* be good or bad for *Anhedonica*, given that it *can* receive the pill. But this is the same sense in which we could say that it *can* have affective experiences. Thus, if we take this to be the morally relevant sense of ‘capacity,’ pre-pill *Anhedonica* would count as both a welfare subject and an affective being, as noted earlier. Here, however, we are assuming that the relevant notion of capacity works in such a way that *Anhedonica* should be said to lack the capacity for affect even though it can receive the pill. Absent any reason for a double standard, we should similarly say that *Anhedonica* lacks the capacity for welfare goods or bads in the morally relevant sense of ‘capacity,’ even though it can receive the pill in the mere sense of modal possibility. It follows that pre-pill *Anhedonica* lacks a sake: it lacks a perspective from which things can be good or bad for it in the relevant sense of ‘can.’

Who should then get the pill—*Anhedonica* or the non-conscious robot? Neither of them has any interest in being benefited before receiving the pill; they both lack the faculties necessary to have a sake. So in neither case can we have a moral reason to benefit them grounded in their currently existing interests. Any such reason, then, must be grounded in their future interests—in the fact that they will be grateful for what we did and what they gained from it. But this applies equally to both beings. Both *Anhedonica* and the robot will be every bit as glad for what we did. Therefore, all else equal, I think we should be indifferent as to which of the two receives the pill.

It is controversial to claim that there is no difference between creating a happy being and increasing the happiness of an already existing being. But this is a special case, because the already existing being is not a welfare subject—it exists, but not as a subject that matters morally. So, for moral purposes, giving the pill to *Anhedonica* is relevantly like bringing a new being into existence: we are creating a sake and a life worth living where there were none before, and bringing about interests where there were none before.

One might suggest that pre-pill *Anhedonica* has some derivative interests in virtue of the fact that it *would have* interests were it to receive the pill. But the idea of attributing interests on the basis of mere modal potential would have implausible implications when applied to human fetuses or non-human animals such as insects, which could also gain human-like forms of sentience, rationality, autonomy, and the corresponding interests with the right pill.⁴⁴

One last challenge to the view that beings whose welfare level is fixed at zero do not matter morally. When we calculate the average welfare in a world, it could be argued that we must factor in such beings. This would have implications for non-separable population axiologies, such as averagingism.⁴⁵ However,

⁴⁴ McMahan (2015).

⁴⁵ Lee (2025: 12).

they would be *odd* implications. In a terribly unhappy world, for example, I might have a strong moral reason to create many Vulcan-like AIs to improve the average. But setting aside any indirect benefits, I cannot bring myself to believe that I would be ‘making the world a better place’ by bringing into existence beings that cannot even have lives worth living. Worse still, on these views I could improve the world by creating beings that can only suffer, so long as the existing population suffers even more. I suspect that any population axiology that gives significant weight to Vulcan-like beings has wildly implausible implications—worse than those of competing views—and should be rejected.

VI. Conclusion and Implications

I am conscious, therefore I matter? Attractive as the idea may sound, it is mistaken. Moral status requires more than mere consciousness: it requires consciousness with a positive or negative hedonic valence. And even if beings like Vulcans do not yet exist, this conclusion has real implications.

For one thing, whenever we try to establish the moral status of beings different from us, we now need to determine whether they have affective states, not merely whether they are conscious. In part, this makes things harder: we must ascertain the presence of *both* consciousness *and* affect. In part, it makes things easier: it shifts our attention to well-studied targets of empirical investigation such as emotion, pain, and pleasure.

We can already tell that all vertebrate animals are likely to be affectively conscious. They exhibit strong indicators of consciousness, and affect is among the most ancient evolutionary features of the mind. We remain uncertain whether invertebrates are conscious, but it seems plausible that if they meet the conditions for consciousness, they also meet those for affect. On Affective Sentientism, then, all vertebrates—and perhaps many invertebrates—have moral status.⁴⁶

Things are different for artificial intelligence systems, with which we share no evolutionary history. Such systems may soon exhibit cognitive and sensory processing, a global workspace, and unified agency. With these capacities, AIs could well turn out to be conscious, goal-directed rational agents. But it is far less clear whether these capacities entail affective consciousness. We could therefore have conscious AIs without affect—AIs that, on Affective Sentientism, would lack moral status.

The possibility of Artificial Vulcans might shape the future of AI design. These artificial systems could reason, imagine, plan, understand, act, and

⁴⁶On the evidence for affective consciousness in vertebrates and invertebrates, see Birch et al. (2021) and Birch (2024).

govern themselves. Yet they could not suffer or be wronged. We could delete, copy, or replace them without worrying about their interests. This is not to say that we may treat them however we like, or that shutting them down is always permissible. But any reasons against doing so must be grounded in indirect effects on affective beings, institutions, or norms—not in harms to the artificial systems themselves. Artificial Vulcans thus represent an opportunity to make use of the benefits of advanced intelligence and consciousness without the dangers of suffering and exploitation.

The benefits of AI consciousness may be considerable. There are many tasks that an AI may not be able to perform without being conscious, or at least that might prove harder to engineer without it. At a sufficiently high level of sophistication, an AI's capacity to integrate multimodal information, monitor its own processing, and coordinate its internal states and goal-directed behaviours under a unified perspective may provide strong evidence of consciousness, if not suffice for it outright. If AI consciousness turns out to be unavoidable or desirable, we will need to know whether it can be developed without thereby creating new welfare subjects—and, if so, how. It may then be worth pursuing more research into non-hedonic motivational architectures, so that we might one day design conscious AI systems without affective processes, and do so safely.

Acknowledgements

My deepest thanks go to Roger Crisp, Jeff McMahan, Andreas Mogensen, and Theron Pummer, all of whom read multiple drafts of this paper. For helpful discussions, I also wish to thank Uriah Kriegel, Neil Levy, Willem van der Deijl, Gwen Bradford, Andrew Lee, Dan Pallies, Jeff Sebo, Joshua Shepherd, Jonas Aaron, Guy Kahane, Richard Yetter Chappell, Michael Slote, Otávio Bueno, Jakob Stenseke, Tom Kaspers, Jacob Librizzi, Elisabetta Sassarini, James Glover, Brad Saad, Patrick Butlin, Louise Mackie, Luca Stroppa, Jen Semler, Kida Lin, Paul Heller, Lorenzo Elijah, Tom Douglas, Patrick Williamson, Rhys Southan, Will Combs, Rei Takahashi, Ryan Kulesa, Lorenza D'Angelo, Gary O'Brien, Calvin Baker, Lucius Caviola, Jakob Lohmar, Joel Joseph, Eric Schwitzgebel, and two anonymous reviewers. This work was supported by an Open Philanthropy Scholarship and a St Cross Uehiro Scholarship for Future Generations.

References

- Basl, J. (2019) *The Death of the Ethic of Life*. New York: Oxford University Press.
 Bechara, A. et al. (2000) 'Characterization of the Decision-Making Deficit of Patients with Ventromedial Prefrontal Cortex Lesions', *Brain*, 123: 2189–202.

- Bentham, J. (1907) *An Introduction to the Principles of Morals and Legislation*. Oxford: Oxford Clarendon Press [1780].
- Berridge, K. and Robinson, T. (1998) 'What Is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?', *Brain Research Reviews*, 28: 309–69.
- Birch, J. et al. (2021) *Review of the Evidence of Sentience in Cephalopod Molluscs and Decapod Crustaceans*. London: Department for Environment, Food and Rural Affairs.
- Birch, J. (2022) 'Materialism and the Moral Status of Animals', *Philosophical Quarterly*, 72/4: 795–815.
- Birch, J. (2024) *The Edge of Sentience: Risk and Precaution in Humans, Other Animals, and AI*. Oxford: Oxford University Press.
- Bradford, G. (2015) *Achievement*. Oxford: Oxford University Press.
- Bradford, G. (2022) 'Consciousness and Welfare Subjectivity', *Noûs*, 57/4: 905–21.
- Bradley, B. (2009) *Well-Being and Death*. New York: Oxford University Press.
- Bramble, B. (2016) 'A New Defense of Hedonism about Well-Being', *Ergo: An Open Access Journal of Philosophy*, 3: 85–112.
- Carruthers, P. (1999) 'Sympathy and Subjectivity', *Australasian Journal of Philosophy*, 77/4: 465–82.
- Chalmers, D. (2022) *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: Penguin Books.
- Chalmers, D. (forthcoming) 'Sentience and Moral Status', in G. Lee and A. Pautz (eds.) *The Importance of Being Conscious*. Oxford: Oxford University Press.
- Crisp, R. (2006) *Reasons and the Good*. Oxford: Oxford University Press.
- Damasio, A. R. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.
- DeGrazia, D. (1996) *Taking Animals Seriously: Mental Life and Moral Status*. New York: Cambridge University Press.
- Enoch, D. (2016) 'What's Wrong with Paternalism: Autonomy, Belief, and Action', *Proceedings of the Aristotelian Society*, 116/1: 21–48.
- Fanciullo, J. (2025) 'Alienation, Engagement, and Welfare', *Philosophical Quarterly*, 75/1: 40–60.
- Frankfurt, H. G. (2004) *The Reasons of Love*. Princeton: Princeton University Press.
- Goldstein, S. and Kirk-Giannini, C. D. (2025) 'AI Wellbeing', *Asian Journal of Philosophy*, 4/1: 1–22.
- Grahek, N. (2007) *Feeling Pain and Being in Pain*, 2nd edn. Cambridge: The MIT Press.
- Heathwood, C. (2019) 'Which Desires Are Relevant to Well-Being?', *Noûs*, 53/3: 664–88.
- Jackson, F. (1982) 'Epiphenomenal Qualia', *The Philosophical Quarterly*, 32: 127–36.
- Jaworska, A. (2007) 'Caring and Full Moral Standing', *Ethics*, 117/3: 460–97.
- Kagan, S. (2009) 'Well-Being as Enjoying the Good', *Philosophical Perspectives*, 23/1: 253–72.
- Kagan, S. (2019) *How to Count Animals, More or Less*. Oxford: Oxford University Press.
- Kamm, F. M. (2007) *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.
- Korsgaard, C. M. (2018) *Fellow Creatures: Our Obligations to the Other Animals*. Oxford: Oxford University Press.
- Kriegel, U. (forthcoming) 'The Value of Consciousness to the One Who Has It', in G. Lee and A. Pautz (eds.) *The Importance of Being Conscious*. Oxford: Oxford University Press.
- Lee, A. Y. (2019) 'Is Consciousness Intrinsically Valuable?', *Philosophical Studies*, 176/3: 655–71.
- Lee, A. Y. (2022) 'Speciesism and Sentientism', *Journal of Consciousness Studies*, 29/3–4: 205–28.
- Lee, A. Y. (2025) 'Consciousness Makes Things Matter', *Philosophers' Imprint*, 25/35.
- Levy, N. (2014) 'The Value of Consciousness', *Journal of Consciousness Studies*, 21/1–2: 127–38.
- Lin, E. (2021) 'The Experience Requirement on Well-Being', *Philosophical Studies*, 178: 867–86.
- Lucas, R. and Diener, E. (2008) 'Personality and Subjective Well-Being', in O. P. John, R. W. Robins, and L. A. Pervin (eds.) *Handbook of Personality: Theory and Research*, 3rd edn. New York: The Guilford Press.
- McMahan, J. (2002) *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.
- McMahan, J. (2015) 'On "Modal Personism"', *Journal of Applied Philosophy*, 33/1: 26–30.
- Mill, J. S. (1998) *Utilitarianism*, ed. R. Crisp. Oxford: Oxford University Press [1863].
- Mogensen, A. L. (2025) 'Once More, without Feeling', *Philosophy and Phenomenological Research*, 111/1: 343–65.
- Perry, M. (2024) 'Why Sentience Should Be the Only Basis of Moral Status', *The Journal of Ethics*, 28/4: 719–41.

- Powell, R. and Mikhalevich, I. (2020) 'Affective Sentience and Moral Protection', *Animal Sentience*, 27/35: 1–20.
- Railton, P. (1986) 'Facts and Values', *Philosophical Topics*, 14/2: 5–31.
- Roelofs, L. (2023) 'Sentientism, Motivation, and Philosophical Vulcans', *Pacific Philosophical Quarterly*, 104: 301–23.
- Sebo, J. (2025) *The Moral Circle*. New York: W. W. Norton & Company.
- Shepherd, J. (2018) *Consciousness and Moral Status*. New York: Routledge.
- Shepherd, J. (2023) 'Non-Human Moral Status: Problems with Phenomenal Consciousness', *American Journal of Bioethics Neuroscience*, 14/2: 148–57.
- Shepherd, J. (2024) 'Sentience, Vulcans, and Zombies: the Value of Phenomenal Consciousness', *AI and Society*, 39/6: 3005–15.
- Siewert, C. (1998) *The Significance of Consciousness*. Princeton: Princeton University Press.
- Siewert, C. (manuscript) 'How Subjectivity Makes Lives Matter'.
- Singer, P. (1990) *Animal Liberation*, 1st edn. New York: Avon [1975].
- Slote, M. (2009) *Moral Sentimentalism*. New York: Oxford University Press.
- Smithies, D. (2026) 'Hedonic Consciousness and Moral Status', in U. Kriegel (ed.) *Oxford Studies in Philosophy of Mind*, vol. 5. Oxford: Oxford University Press.
- van der Deijl, W. (2021) 'The Sentience Argument for Experientialism about Welfare', *Philosophical Studies*, 178/1: 187–208.
- van der Deijl, W. (2026) 'Sentientism and the Welfare Level View', *The Philosophical Quarterly*, 76/7: 793–803.