

Hedonic Consciousness and Moral Status

Declan Smithies

Abstract. Which beings have moral status? I argue that moral status requires some capacity for hedonic feelings of pleasure or displeasure. David Chalmers rejects this view on the grounds that it denies moral status to Vulcans, which are defined as conscious creatures with no capacity for hedonic feelings. On his more inclusive view, all conscious beings have moral status. We agree that only conscious beings have moral status, but we disagree about how to explain this. I argue that we cannot explain why consciousness is necessary for moral status without denying that it is sufficient for moral status. Hence, we cannot exclude unconscious zombies without excluding conscious Vulcans too.

1. Varieties of Sentientism

Which beings have moral status? Although this was not always true in the past, it is now widely acknowledged that all humans have moral status. Moreover, there is increasing recognition that many non-human animals have moral status too. Even so, there is little temptation to extend moral status to all living organisms, including plants, fungi, and bacteria, since these simple life forms don't have the right kinds of minds. This raises the question: what kind of mind do you need for moral status?

According to *sentientism*, moral status requires the capacity for consciousness.¹ The pertinent concept of consciousness is the phenomenal concept, according to which consciousness is subjective experience: a mental state is conscious when and only when there is “something it is like” for the subject to experience having that mental state. Sentientism, as defined here, takes no stand on what kind of subjective experience you need for moral status.

David Chalmers (2022) argues for an *inclusive* version of sentientism, according to which the capacity for consciousness is not only necessary but also sufficient for moral status:

Inclusive Sentientism: A being has moral status if and only if it is a conscious being: it has the capacity for consciousness.

On this view, all conscious beings and only conscious beings have moral status. Any conscious being has moral status, no matter how simple it is, although its degree of moral status may vary depending on its psychological complexity. Meanwhile, no unconscious being has any degree of moral status no

¹ Proponents of sentientism about either moral status or welfare include Singer (1975), Rosati (2009), Shepherd (2018), van der Deijl (2020), Lin (2021), Siewert (2021), and Chalmers (2022).

matter how complex its life might be in other ways. Not even *zombies* – that is, unconscious beings that function just like conscious beings – have moral status.

I agree with Chalmers that consciousness is necessary for moral status, but I disagree with his claim that it is also sufficient for moral status. In this paper, I will defend the view that moral status depends on the capacity for hedonic consciousness:

Hedonic Sentientism: A being has moral status if and only if it has the capacity for hedonic consciousness: that is, feelings of pleasure or displeasure.²

Chalmers rejects this view on the grounds that it denies moral status to Vulcans, which he defines as conscious beings with no capacity for pleasure or displeasure. My preliminary goal in this paper is to answer this objection. In the course of answering it, however, I hope to achieve the more ambitious goal of explaining why consciousness is required for having moral status at all. I'll argue that we cannot explain why unconscious zombies lack moral status without denying that conscious Vulcans have moral status either. Hence, we cannot maintain that consciousness is a necessary condition for moral status unless we also deny that it is a sufficient condition.

2. The Moral Status of Vulcans

Vulcans are defined as conscious beings with no capacity to experience hedonic feelings of pleasure or displeasure. As Chalmers writes, “A *Vulcan* is a conscious creature who experiences no happiness, suffering, pleasure, pain, or any other positive or negative affective states” (2022: 343). Our question is whether Vulcans, so defined, have any degree of moral status: inclusive sentientism implies they do, whereas hedonic sentientism implies otherwise. Hence, Vulcans provide a useful test case for deciding between competing versions of sentientism.

There need be no incoherence in this definition of a Vulcan so long as we make the plausible assumption that not all conscious experience is hedonically valenced. Vulcans may lack any capacity to feel pleasure or displeasure, but they can nevertheless experience many other sensations, perceptions, and conscious thoughts. This assumption is what generates the debate between inclusive and hedonic versions of sentientism. Without it, the debate simply collapses.

² Chalmers (2022: 342) attributes this view to Jeremy Bentham and Peter Singer, but it's worth explicitly noting that hedonic sentientism can be decoupled from their broader commitments to hedonistic utilitarianism.

When we imagine Vulcans, there are certain pitfalls we must be careful to avoid. One of them is neglecting the distinction between *sensory pleasure* and *attitudinal pleasure*. All pleasure is attitudinal pleasure, since you always feel pleasure *about* something: the feeling of pleasure is an attitude towards some content that captures what your pleasure is about.³ Sensory pleasure is just a special case of attitudinal pleasure that is directed towards your sensations or their contents. On a *higher-order* attitudinal theory, you feel sensory pleasure when you experience a sensation that you feel pleased about for its own sake. On a *first-order* attitudinal theory, in contrast, you feel sensory pleasure when you experience a sensation that represents something else that you feel pleased about for its own sake. Either way, sensory pleasure is reducible to attitudinal pleasure directed towards sensations themselves or whatever those sensations represent. Moreover, sensory displeasure can be explained in much the same way, including the unpleasant feeling of sensory pain.⁴

Crucially, however, sensory pleasure is not the only species of attitudinal pleasure. You can feel pleased about something without feeling any sensory pleasure, since we sometimes experience feelings of intellectual pleasure as well as sensory pleasure. To adapt an example from Fred Feldman (1988), consider an injured motorcyclist who feels pleased that he survived his accident, although he has taken a powerful anesthetic that prevents him from feeling any sensations at all. He can feel pleasure in the thought that he survived the accident even while all his senses are numbed.⁵ So, when we imagine Vulcans, it is not enough to imagine creatures who are incapable of sensory pleasure and displeasure. We must imagine creatures who are incapable of feeling pleased or displeased about anything at all. Otherwise, we artificially inflate the plausibility of the claim that they have moral status.

A second pitfall is failing to recognize that feelings of emotion and desire are hedonic feelings. The hedonic dimension of these feelings is essential to explaining their positive or negative valence. You can, of course, make the conscious value judgment that something is good or bad without thereby feeling any pleasure or displeasure. Our feelings of emotion and desire, however, are positively or negatively valenced in a psychologically more primitive way. Positively valenced emotions feel pleasant because they make us feel pleased about things, whereas negatively valenced emotions feel unpleasant

³ We're not forced to choose between attitudinal and phenomenological theories of pleasure, since there is an attitude of feeling pleased that *p*, which has its own distinctive phenomenology. Lin (2020) in defense of a hybrid theory.

⁴ Feldman (1988) and Heathwood (2007) defend higher-order theories of pleasure, whereas Jacobson (2019) defends a first-order theory of why pain feels unpleasant: it represents a bodily condition towards which we feel aversive desire.

⁵ Feldman makes the questionable assumption that all feelings are sensory feelings, but we needn't assume that phenomenal consciousness is restricted to sensory experience. See the essays in Bayne and Montague (2011) for discussion.

because they make us feel displeased about things. When you feel happy or amused, for example, you feel pleasure in whatever it is that you feel happy or amused about. When you feel angry or afraid, in contrast, you feel displeasure in whatever it is that you feel angry about or afraid of. Similarly, when you feel the desire for something, you feel pleasure in the thought of having what you desire, or displeasure in the thought of not having it.⁶ When we imagine Vulcans, we must imagine beings who are incapable of any hedonically valenced feelings, including feelings of desire and emotion.

A third pitfall is imagining Vulcans as having some *diminished* capacity for hedonic feelings, rather than no such capacity at all. This is why Chalmers is careful to distinguish his Vulcans from the Star Trek characters that inspired them: Spock may be less emotional than his human counterparts, but he still has some minimal capacity to feel pleasure and pain. In contrast, Vulcans are defined as having no capacity for hedonic experience. So, when we imagine Vulcans, we must be careful to avoid cheating by imagining human beings who suffer from real-life psychological disorders, such as depression or anhedonia, which diminish your capacity for hedonic experience without eliminating the capacity altogether.

With these clarifications in mind, let's now consider whether Vulcans have moral status. Chalmers' argumentative strategy is to pump intuitions about moral status by using variations on the classic trolley problem. First, he argues that *zombies* lack moral status:

You're at the wheel of a runaway trolley. If you do nothing, it will a single conscious human, who is on the tracks in front of you. If you switch tracks, it will kill five nonconscious zombies. What should you do? (2022: 340)

Chalmers' intuition, and mine too, is that you should kill five zombies to save one conscious human. He even feels some temptation to say that you should destroy a whole planet of humanoid zombies in order to save one conscious chicken, although he admits to wavering on this point. When we substitute Vulcans for zombies, however, his intuitions become much clearer:

Would it be morally acceptable to kill a planet of philosophical Vulcans to save one human with ordinary affective consciousness? I think the answer is obviously no. More simply, suppose you're faced with a situation in which you can kill a Vulcan in order to save an hour

⁶ See Smithies (MS) for further discussion and defense of the hedonic theory of desire.

on the way to work. It would obviously be morally wrong to kill the Vulcan. In fact, it would be monstrous. (2022: 344)

What makes all the difference between these cases, according to Chalmers, is that Vulcans have moral status, whereas zombies have none.

While I feel the force of these moral intuitions, I deny that they provide any reliable evidence for the moral status of Vulcans. My complaint is that Chalmers's intuitions about the trolley problem conflate an important distinction between *moral status* and *moral significance*. A being has moral status only if it has *intrinsic* moral significance: it matters morally for its own sake. Other things have moral significance that is not intrinsic but *derived*: they matter morally only because they affect other things that matter morally for their own sake. Hence, not everything with moral significance has moral status, since its moral significance may be derived, rather than intrinsic.

According to Chalmers, "A being has moral status when . . . it's a being whose welfare we need to take into account in our moral deliberations" (2022: 339). The suggestion here is that what gives you moral status is being a welfare subject, since your welfare is what makes you matter morally for your own sake. This needn't imply that maximization of welfare is the only relevant consideration in moral deliberation, since there may be various deontological constraints on how welfare may be permissibly distributed. Any such deontological constraints apply only to welfare subjects, however, since other beings matter morally only by affecting the welfare of beings with moral status. Following Chalmers, then, I'll assume that all and only welfare subjects have moral status.⁷

Our intuitions about trolley problems conflate this distinction between moral status and moral significance: although only welfare subjects have moral status, other things have derived moral significance when they impact the welfare of beings with moral status. Hence, we needn't suppose that Vulcans have moral status to explain why it would be monstrous to destroy them merely to save time on your morning commute. After all, it would be monstrous to destroy a beautiful rainforest or valuable artwork just for your own convenience. And yet rainforests and artworks don't have moral status, since they are not welfare subjects: they don't satisfy the minimal requirement for consciousness imposed by even the most inclusive version of sentientism.

Indeed, we can use this distinction to explain why Chalmers finds his own intuitions wavering when he considers extreme versions of the zombie trolley problem. According to inclusive

⁷ Anyone who rejects this assumption can simply recast the whole debate in terms of welfare rather than moral status.

sentientism, conscious chickens have welfare status, whereas unconscious zombies have none. And yet Chalmers is reluctant to embrace the conclusion that we should destroy a whole planet of unconscious zombies to save just one conscious chicken. We can resolve any apparent tension here by recognizing that zombies – like rainforests and artworks – may have some moral significance without having any moral status. If conscious beings value them highly enough, then it may sometimes be morally permissible to countenance some loss of conscious life in order to preserve them. And if this is true of zombies, then all the more so for Vulcans. Vulcans may have moral significance that is derived from their impact on welfare subjects, but it doesn't follow that they are welfare subjects with moral status themselves.

In short, trolley problems are an unreliable guide to which beings have moral status. When faced with a choice between destroying beings that have moral status and beings that lack moral status, it doesn't always follow that you should take the second option. Sometimes, you should preserve things that lack moral status because of their impact on the welfare of beings that have moral status. Having undermined Chalmers's positive argument that Vulcans have moral status, I'll now turn to my negative argument that Vulcans lack moral status.

3. Against Inclusion

Rather than simply trading intuitions about the moral status of Vulcans, my goal is to move the debate forward by confronting inclusive sentientism with an explanatory question. What could explain why Vulcans, unlike zombies, are welfare subjects with moral status? There are two things to explain here:

- (1) **Inclusion:** All conscious beings are welfare subjects.
- (2) **Sentientism:** No unconscious beings are welfare subjects.

In this section, I'll argue that Chalmers cannot plausibly explain both claims without equivocating between distinct concepts of value. Once we have this distinction in clear focus, I'll argue in later sections that we cannot plausibly explain sentientism without abandoning inclusion.

The materials for Chalmers's explanation of inclusive sentientism are contained within the following passage:

Consciousness is the ground of all value. Whenever anything is good or bad for someone, it's because of their consciousness. Consciousness has value, what a conscious creature values has

value, and relations between conscious creatures have value. If a creature has no capacity for consciousness, nothing can be good or bad for it from its own perspective. And it's natural to conclude that if nothing can be good or bad for a creature, then the creature has no moral status. (2022: 341-2)

This passage is compressed enough that it needs fleshing out, but we can reconstruct an explanation of inclusive sentientism by extracting two central claims:

- (1) **Intrinsic Value:** All consciousness has intrinsic value.
- (2) **Grounding:** All other value is grounded in the intrinsic value of consciousness.

The first claim is designed to explain *inclusion*. The idea is that all consciousness has value – not just instrumentally but intrinsically – for any conscious being. On this view, consciousness is an intrinsic welfare good that always makes your life better for you to some extent. Consequently, all conscious beings are welfare subjects, since they can be harmed by depriving them of an intrinsic welfare good – namely, consciousness. Meanwhile, the second claim is designed to explain *sentientism*. Since all value – including welfare value – is grounded in consciousness, no unconscious beings are welfare subjects. Zombies may have value for conscious beings like us, but their lives have no value for them.

As elegant as it is, this explanation trades on equivocation. We cannot plausibly maintain both claims without shifting between distinct concepts of value: namely, *intrinsic value* and *welfare value*. Admittedly, some welfare values are intrinsic values, but not all intrinsic values are welfare values. To illustrate the distinction: many epistemologists accept that knowledge has intrinsic epistemic value in addition to its instrumental welfare value in promoting the satisfaction of my desires, but this needn't commit them to the further claim knowledge has intrinsic welfare value. Indeed, this further claim seems dubious, since knowledge of random trivia is not even *pro tanto* good for me unless it satisfies some curiosity or other desire of mine. In any case, there is no incoherence in this combination of views so long as we recognize that knowledge can have intrinsic *epistemic* value without having any intrinsic *welfare* value.

With this distinction in mind, let's consider Grounding: is it true that all value is grounded in consciousness? Without any further restriction, this claim seems doubtful. Following Moore, we can contrast two worlds devoid of consciousness, one full of beauty and the other replete with filth:

Let us imagine one world exceedingly beautiful. Imagine it as beautiful as you can; put into it whatever in this world you most admire – mountains, rivers, the sea; and sunsets, stars and moon . . . And then imagine the ugliest world you can possibly conceive. Imagine it simply one heap of filth . . . The only thing we are not entitled to imagine is that any human being ever has or ever, by any possibility, *can*, live in either, can ever see and enjoy the beauty of the one or hate the foulness of the other. (1903: 84)

As Moore observed, the first world seems better than the second because it is more beautiful. After all, beauty is a kind of aesthetic value. Moreover, the beauty of the first world must be intrinsically valuable, rather than merely extrinsically valuable for conscious beings, since it contains no conscious beings to enjoy its beauty. Perhaps it will be said that what makes this world beautiful is our disposition to enjoy it when we imagine it. But ugly things may be disposed to elicit aesthetic pleasure too. What makes something beautiful is not the merely descriptive fact that it is disposed to elicit aesthetic pleasure, but rather the evaluative fact that it *merits* aesthetic pleasure. And yet it remains quite obscure how to ground this evaluative fact about an unconscious world in the value of consciousness. The issue needs more extended discussion, of course, but we cannot simply assume without argument that consciousness is the only intrinsic value from which all other values are derived.

Perhaps Chalmers can avoid this problem, since he is primarily concerned with welfare value, and its connection with moral value, rather than aesthetic value.⁸ A beautiful world devoid of consciousness might be good in itself, but it is not good *for* anyone in the sense that is relevant to welfare and morality. So, perhaps what Chalmers intends is a restricted version of Grounding, according to which all moral and welfare value is grounded in consciousness. If so, then we are in full agreement. In a world without consciousness, nothing has value in this restricted sense: nothing is morally good or bad, since nothing is good or bad *for* anyone.

Once we impose this restriction, however, we undercut the plausibility of Intrinsic Value. This is a more serious difficulty for Chalmers, since he needs the restricted version of Intrinsic Value in order to explain why all conscious beings are welfare subjects. I have no dispute with the claim that consciousness has intrinsic value in something like the way that unconscious life does. Perhaps it has

⁸ Chalmers (2022: 315) says he is mainly concerned with personal or prudential value, which I'm calling "welfare value". Although he connects this with moral value, he mentions aesthetic value only to set it aside, and doesn't refer back to it when he introduces his proposal that "all value arises, one way or another, from consciousness" (2022: 329).

aesthetic value or perhaps intrinsic value of some other kind. Other things being equal, a conscious world seems better to me than a zombie world just as a world teeming with unconscious life seems better to me than a world comprising only rocks and dust. And yet these intuitions about intrinsic value cannot be used to support more restricted claims about welfare value. After all, we can accept that life has intrinsic value without supposing it is a welfare good for any living organism. Similarly, we can accept that consciousness has intrinsic value without supposing it is a welfare good for any conscious being.

To further illustrate the point, let's consider an example from Andrew Lee:

Consider again two worlds that are empty save for a single creature inhabiting the universe. In the first world, the creature has a painful experience, and the character of its experience is fully exhausted by this pain phenomenology . . . In the second world, the creature is not conscious at all. (2018: 664)

Is the painful world better or worse than the zombie world with no consciousness at all? My own reaction is one of ambivalence. On the one hand, how awesome it is that consciousness emerged in this world! The painful world contains a source of intrinsic value that is entirely missing from a zombie world. On the other hand, how unfortunate for the creature in question! It would be better for this creature not to be conscious or not to exist at all. There need be no conflict between these evaluations, since they use distinct and perhaps incommensurable evaluative concepts. Consciousness can have intrinsic value without being a welfare value for all conscious beings. We cannot draw conclusions about welfare or moral status solely from premises about the intrinsic value of consciousness.⁹

Chalmers might insist that consciousness always has positive welfare value, although it can sometimes be outweighed by the negative welfare value of specific ways of being conscious, including unpleasant pain. However, we can make trouble for this view by varying the intensity of displeasure in the painful world (cf. Lee 2018: 664-67). If consciousness has non-trivial welfare value, then this can in principle outweigh the disvalue of pain when its unpleasantness is mild enough. And yet it

⁹ Lee himself doesn't use this example to distinguish between intrinsic value and welfare value. Indeed, he seems to conflate them in much the same way as Chalmers. While Chalmers assumes that consciousness has welfare value because it has intrinsic value, Lee seems in effect to argue in the opposite direction: that consciousness has no intrinsic value because it has no welfare value. In my view, however, both inferences are fallacious.

seems implausible that being conscious is good overall for a creature who experiences nothing but mildly unpleasant pain. Arguably, this is because the determinable property of being conscious – as distinct from any more determinate way of being conscious – has no positive welfare value at all.

My main objection, however, is more straightforward. We can imagine conscious beings who are not welfare subjects at all. Consider *the Blob*, an alien life form superficially resembling a blobfish that drifts around on the bottom of the ocean. Blobs are not agents who are capable of self-movement. Instead, they are passive observers who experience an impoverished stream of conscious sensations, perceptions, and thoughts about themselves and their environment. Crucially, however, they are incapable of hedonic experience of any kind. When Blobs see something red, they might think, “That is red,” or even, “That looks red to me,” but they feel indifferent to the presence or absence of red things. Indeed, we may even suppose that they feel indifferent to their own continued existence. Blobs are constitutionally incapable of feeling positively or negatively about anything at all.

I find it highly implausible to suppose that Blobs are welfare subjects. A welfare subject is the kind of being who can possess welfare goods or bads, but Blobs do not appear to satisfy this criterion. After all, what could make the lives of Blobs go better or worse for them? We cannot benefit them by moving red objects into their vicinity, or harm them by taking them away, since they feel indifferent to the presence or absence of red things. And we cannot harm them by destroying them, since they feel indifferent to their own continued existence. We might talk about what’s good for Blobs in the same way that we talk about what’s good for plants and machines, but not all relational goods are welfare goods (Rosati 2009). Like plants and machines, Blobs can be damaged or destroyed, but they cannot be harmed or wronged.

The obvious rejoinder is that we can harm Blobs by depriving them of consciousness. To my mind, however, this reply loses much of its plausibility once we recognize that not all intrinsic values are welfare values. We can acknowledge that knowledge has intrinsic epistemic value, for example, without supposing that we harm Blobs by depriving them of the capacity to know about the location of red things. Similarly, we can recognize that consciousness has intrinsic value without supposing that we harm Blobs by depriving them of the capacity to experience red things.

Someone may protest: isn’t better for you to be a conscious Blob than an unconscious zombie? As Charles Siewert (2021) observes, many of us would decline a “radical pheno-ectomy” that surgically removes all our conscious experience while leaving our behavioral capacities intact. We prefer to remain conscious, rather than zombified, because we value conscious experience for its own sake and not just because of its behavioral consequences. Even so, it is doubtful that we intrinsically value the

determinable property of consciousness itself, rather than some of its more specific determinates. Being zombified would deprive me of ways of being conscious that I value for their own sake, including the pleasures of drinking wine or thinking about philosophy. Without any conscious perception or conscious thought, these things would be impossible. Speaking for myself, however, I can see no welfare value in conscious thought or perception *per se*.

Suppose an accident will render me permanently comatose: I have no preference for a future in which I experience only a monotonous earworm over one in which I remain unconscious. Indeed, I cannot see how this impoverished stream of consciousness would make my life any better for me than the alternative. Perhaps others have different preferences: a conscious future may be better for them insofar as it would satisfy their current desires. It doesn't follow, however, that consciousness has any welfare value for creatures who are completely indifferent towards it. Intuitively, for instance, there is nothing good or bad for the Blob about being conscious.

If so, then inclusion is false: not all conscious beings are welfare subjects with moral status. Inclusion can seem plausible when we imagine humanoid Vulcans whose lives are otherwise similar to our own, but it is much less plausible when we imagine Blobs. According to Chalmers' official definition, however, Blobs *are* Vulcans: they are conscious creatures with no capacity for hedonic feelings. Their conscious life may have intrinsic value, but it has no welfare value *for them*.

4. Explaining Sentientism

The previous section cast doubt on inclusion by appealing to intuitions about Blobs. Even so, I'd prefer not to rely solely on intuitive verdicts about cases. My main argument against inclusion is an explanatory one: we must reject inclusion in order to explain sentientism. Although Chalmers and I agree that sentientism is true, we disagree about how best to explain it. In fact, I'll argue that Chalmers doesn't really explain it at all, but merely presupposes what needs to be explained. Moreover, I'll argue that we cannot adequately explain sentientism except in a way that undermines inclusion.

The challenge is to explain why all welfare subjects must be conscious. Willem van der Deijl (2020) argues that the only adequate explanation of sentientism is *experientialism*: the view that all intrinsic welfare goods are ways of being conscious. One version of experientialism is *hedonism*, according to which the only intrinsic welfare good is feeling pleasure, but there are also more liberal versions of experientialism that include some non-hedonic feelings as intrinsic welfare goods, such as the experience of novelty, self-understanding, or compassion (van der Deijl 2019). To explain inclusion, of course, we need a maximally liberal version of experientialism that includes every way of

being conscious as an intrinsic welfare good. Otherwise, we cannot explain why all conscious beings are welfare subjects. What we need to explain sentientism, in contrast, is the claim that all intrinsic welfare goods are ways of being conscious. On this view, all welfare subjects must be conscious because only conscious experience can make an intrinsic contribution to welfare.

The main problem with this explanation of sentientism is that experientialism has the consequence that welfare supervenes on experience alone. As Robert Nozick's (1974) experience machine illustrates, however, experiential duplicates can diverge in welfare. Intuitively, your welfare depends not on your experience alone, but on whether your experiences correspond with objective reality. Like many philosophers, I'm inclined to regard this as a compelling – perhaps even decisive – objection to experientialism. If experientialism is the only way to explain sentientism, then so much the worse for sentientism.

Is there any other way to explain sentientism? Proponents of sentientism who reject experientialism confront a more difficult explanatory challenge. If it's not the case that all welfare goods are ways of being conscious, then why should consciousness be a prerequisite for possessing any welfare goods at all? Suppose there are some non-experiential welfare goods, such as friendship or achievement, in addition to experiential welfare goods, such as pleasure. By definition, zombies cannot feel pleasure, but why can't they have friends or achievements or any other non-experiential welfare goods? We need some unifying explanation that makes it intelligible why none of these welfare values can be possessed by any unconscious being.¹⁰

To explain sentientism, Chalmers invokes his restricted version of Grounding: all welfare values are grounded in consciousness in such a way that only conscious beings can possess them. And yet this explanation only pushes the crucial question back one stage further. What principled reason is there to suppose that any welfare values must be grounded in consciousness in such a way that only a conscious being can possess them? Simply to maintain this without any further argument is tantamount to endorsing sentientism without explaining it.

Some proponents of sentientism may reject any demand for explanation. Quietism is a risky strategy, however, since sentientism may be regarded with suspicion in the absence of any principled explanation. For instance, Gwen Bradford (2023) rejects sentientism on the grounds that it has no principled and plausible explanation, while Uriah Kriegel (forthcoming) comes dangerously close to

¹⁰ To my mind, it is not enough to argue on a case-by-case basis, as Lin (2021: 880) does, that consciousness is required for each basic welfare good considered by itself, since this strategy is too disjunctive to provide any unifying explanation.

the same conclusion. This paper aims to develop an explanation that neither Bradford nor Kriegel considers in any detail, although this explanation of sentientism will undermine inclusion.

My explanation of sentientism appeals to a normative version of the Resonance Constraint, according to which your welfare is constrained by your intrinsic desires and other pro-attitudes that give you justifying reasons for action. Crucially, however, your intrinsic desires and pro-attitudes cannot give you justifying reasons for action except by disposing you to feel intrinsic pleasure or displeasure in their contents. Therefore, you cannot be a welfare subject at all unless you have some disposition to experience hedonic feelings of pleasure or displeasure. This criterion not only excludes unconscious zombies, since they cannot feel anything at all, but also conscious Vulcans, since they cannot feel pleasure or displeasure.

My position bears comparison with Luke Roelofs' (2023) *motivational sentientism*, according to which a being is a welfare subject with moral status only if it has the right kind of consciousness to rationally motivate action. Roelofs defends sentientism on the grounds that only consciousness can rationally motivate action. At the same time, however, he officially remains agnostic about which kinds of consciousness are required to rationally motivate action. This affords him some flexibility in accommodating the moral status of Vulcans: it depends on whether or not they can have the right kind of consciousness to rationally motivate action. In contrast, I take the hard line that Vulcans lack moral status because only hedonic consciousness can rationally motivate action in the way that is required to satisfy the Resonance Constraint.

5. The Resonance Constraint

My initial objection to inclusion is that Blobs are not welfare subjects. But what exactly is wrong with supposing that the Blob's welfare consists in staying alive, or being conscious, or being acquainted with red things? We can explain what's wrong with this idea by appealing to the Resonance Constraint, according to which your welfare must "resonate" with you in some way. Here is Peter Railton's influential statement of this idea:

What is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. It would be an intolerably alienated conception of someone's good to imagine that it might fail in any such way to engage him. (1986: 9)

We violate the Resonance Constraint if we suppose that it's good for the Blob to be alive, or conscious, or acquainted with red things in spite of the Blob's complete indifference. These conceptions of welfare are flawed because they imply that what is allegedly good for the Blob fails to resonate with him in any way. As Peter Railton observes, it is implausible that anyone can be alienated in this way from their own good.

The Resonance Constraint imposes a necessary condition on welfare but without entailing any specific theory of welfare. It is sometimes used as a premise in arguing for subjectivist theories, according to which welfare is determined by subjective pro-attitudes, such as pleasure or desire. However, it may also be combined with objectivist theories, which deny that welfare is determined by subjective attitudes alone. For example, Derek Parfit (1984) proposes a hybrid theory, which says that something is good for you only if it resonates with your subjective attitudes, although it is better for you if it is objectively worth desiring. Meanwhile, Guy Fletcher (2013) proposes an objective list theory, according to which any objective welfare value on the list is one that essentially resonates with your subjective attitudes. Friendship and achievement, for example, are objective welfare goods that you cannot plausibly possess while remaining wholly indifferent towards them. The Resonance Constraint excludes only the more extreme view that there are objective welfare goods, such as consciousness, which are good for you even if you are completely indifferent to them and to everything else. This view is implausible insofar as it countenances welfare subjects, including Blobs, who are alienated from their own welfare.

The Resonance Constraint is a recurring theme in the literature on welfare, but there is considerable variation in how it is formulated.¹¹ According to one standard formulation, your welfare is constrained by your *desires* in such a way that something is good for you only if it would satisfy some desire of yours. One serious problem with this formulation is that it's not clear that desire is the only pro-attitude that constrains welfare: primitive creatures with no desires might feel pleasure or displeasure in ways that impact their welfare. Some philosophers endorse desire-based theories of pleasure, which attempt to reduce pleasure to desire, but it is arguable that pleasure is a more basic pro-attitude than desire.¹² We can avoid this debate for now by formulating the Resonance Constraint as the idea that your welfare is constrained by your *pro-attitudes*, including pleasure as well as desire.

¹¹ See Railton (1986), Rosati (1996), and Velleman (1998) for some of the most important landmarks.

¹² Heathwood (2006) argues that being pleased that *p* is desiring that *p* while believing that *p*, whereas Smithies (MS) argues that desiring that *p* is disposed to feel pleasure in what you imagine when you imagine that *p*.

But this only raises the question: what is the common feature that unifies the class of pro-attitudes that constrain welfare?

One popular answer is that the pro-attitudes that constrain welfare are unified by their causal role in *motivation*: they motivate you to act in ways that you believe will satisfy your pro-attitudes. On this interpretation of the Resonance Constraint, your welfare is constrained by your motivational states in such a way that something is good for you only if it is capable of motivating you. As Connie Rosati observes, “Many have thought it is a necessary condition on something being good for someone that she be capable of being motivated by it” (2010: 227). This formulation, however, faces much the same problem as the previous one: primitive creatures with no capacity for motivation might still feel pleasure or displeasure in ways that impact their welfare. For example, Galen Strawson (1994) argues that “Weather Watchers” have desires about the weather that they are not at all motivated to act upon, although their welfare may be affected by the apparent satisfaction or frustration of these desires. Similarly, David Lewis (1980) argues that a “madman” need not be motivated to alleviate his own feelings of pain, but this surely doesn’t imply that his pain is not bad for him. Hence, it is not obvious that you must have any motivation to pursue your own good.

A better answer is that the pro-attitudes that constrain welfare are unified by their normative role in providing us with *justifying reasons* for action. We need not combine this with the further claim that you are always motivated to act upon your justifying reasons for action. Instead, the idea is that your pro-attitudes are *apt* to motivate action when they give you justifying reasons for action.

Not all your pro-attitudes play this normative role. You might desire something only because you believe it is a means to something else you desire for its own sake, but your instrumental desires cannot provide justifying reasons for action beyond those already provided by your intrinsic desires. Nor can the real or apparent satisfaction of your instrumental desires boost your welfare except through the real or apparent satisfaction of your intrinsic desires. When you desire something for its own sake, however, this gives you some defeasible reason to do what you have reason to believe will satisfy your intrinsic desire. This explains why so acting can be instrumentally rational. Similarly, when you are intrinsically pleased or displeased about something for its own sake, you thereby have some reason to do whatever you have reason to believe will prolong or prevent what you are pleased or displeased about. And the same goes for any other intrinsic pro-attitudes that constrain welfare.

Accordingly, I propose the following normative interpretation of the Resonance Constraint:

The Resonance Constraint on Welfare Goods: Something is good for you only if it resonates with your intrinsic desires or other pro-attitudes that provide you with justifying reasons for action.

This formulation raises questions about how to understand *resonance*: what relation must hold between your welfare goods and your intrinsic desires or pro-attitudes in order for them to “resonate” in the right way? For current purposes, however, we can finesse such questions by reformulating the Resonance Constraint as a constraint on *welfare subjects*, rather than *welfare goods*:

The Resonance Constraint on Welfare Subjects: A being is a welfare subject only if it has intrinsic desires or other pro-attitudes that provide it with justifying reasons for action.

Perhaps this constraint on welfare subjects ultimately derives from a corresponding constraint on welfare goods. Even so, it is independently plausible on its own terms. Indeed, the best argument for any corresponding constraint on welfare goods is an abductive one: namely, we need it to give a principled explanation of this independently plausible constraint on welfare subjects.

We find similar reasoning in Peter Railton’s seminal discussion of the Resonance Constraint. He articulates the core insight as follows:

It seems to me that notions like good and bad have a place in the scheme of things only in virtue of facts about what matters, or could matter, to beings for whom it is possible that something matter. Good and bad would have no place within a universe consisting only of stones, for nothing matters to stones. (1986: 9)

Inanimate objects like rocks and stones are not welfare subjects because they have no intrinsic desires or pro-attitudes that justify action. And what goes for these inanimate objects goes equally for Blobs. Although they are conscious, they have no intrinsic desires or pro-attitudes that justify action. In short, nothing matters *for a Blob* because nothing matters *to a Blob*.

6. Can Vulcans Have Desires?

Our next question is whether Vulcans can satisfy the Resonance Constraint by having desires or other pro-attitudes of the right kind. Chalmers assumes that they can:

Vulcans' lives may be literally joyless, without the pursuit of pleasure or happiness to motivate them . . . But they nevertheless have serious intellectual and moral goals. They may want to advance science, for example, and to help those around them. They might even want to build a family or make money. They experience no pleasure when anticipating or achieving these goals, but they value and pursue the goals all the same. (2022: 343-4)

If so, then perhaps Vulcans are welfare subjects in virtue of having desires, rather than conscious experiences. This is less inclusive than Chalmers' official view, since it excludes conscious Blobs from being welfare subjects, although it has the compensating advantage of complying with the Resonance Constraint. Moreover, it threatens to generate counterexamples to hedonic sentientism by including Vulcans as welfare subjects with moral status when they have desires.

But why assume that Vulcans can have desires? Chalmers' main consideration is that we can imagine Vulcans who are motivated to pursue various outcomes, such as advancing science, building families, or making money. Indeed, we can even imagine Vulcans who are motivated to act in all the same ways that we are. These humanoid Vulcans are *hedonic zombies* in whom unconscious motivations combine with conscious thoughts and perceptions to play the causal roles normally occupied by our hedonic feelings.¹³

According to the *motivational theory* of desire, the essence of desire is to combine with belief in the motivation of action. Here is Robert Stalnaker's classic formulation:

The Motivational Theory of Desire: To desire that p is to be disposed to act in ways that would tend to bring it about that p in a world in which one's beliefs, whatever they are, were true. (1984: 15)

Stalnaker combines the motivational theory of desire with a corresponding theory of belief:

¹³ See Smithies and Weiss (2019: 30) on "affective zombies". Compare Strawson's Aldebaranians: "They are not capable of any affect states at all, but they are capable of entering into . . . motivating states that are functionally very similar to states that we normally think of as desire states . . . in respect of the way in which they interact with a being's informational states to cause it to move in apparently goal-directed ways" (1994: 281).

The Motivational Theory of Belief: To believe that p is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in a world in which p (together with one's other beliefs) were true. (1984: 15)

The idea is that belief and desire are inter-defined by their causal role in combining to motivate action. This is perhaps the default option for any functionalist theory of mind, according to which mental states can be functionally defined by their causal role.

The motivational theory implies that Vulcans desire to achieve their goals just so long as they are motivated to pursue them: they have these desires not because of their conscious experience, but rather because of their motivational dispositions. The problem is that the motivational theory implies that *zombies* can have desires for much the same reason. After all, they can be motivated to act in just the same way that Vulcans can. Hence, the motivational theory includes Vulcans within the class of welfare subjects only by including zombies. This undermines sentientism, the view that only conscious creatures are welfare subjects with moral status.

Instead of abandoning sentientism, I'll argue that we should abandon the motivational theory of desire. Here is my argument in a nutshell. Our desires not only *motivate* us to act but also provide *justifying reasons* for action. After all, our actions can be instrumentally reasonable, rational, or justified when they are based on the justifying reasons provided by our beliefs and desires. And yet a motivational disposition to do something is not enough for having any justifying to do it. Hence, motivation is not sufficient for desire. This not only blocks the argument that zombies have desires, but also undermines Chalmers' assumption that Vulcans have desires.

We can illustrate the point by considering examples of "Vulcanized" desire in which we excise all hedonic feelings associated with desire, while leaving their motivational dispositions intact. Here, for instance, is Warren Quinn's example of radioman:

Suppose I am in a strange functional state that disposes me to turn on radios that I see to be turned off. Given the perception that a radio in my vicinity is turned off, I try, all other things being equal, to get it turned on. Does this state rationalize my choices? Told nothing more than this, one may certainly doubt that it does. But in the case I am imagining, this is all there is to the state. I do not turn on the radios in order to hear music or get news. It is not that I have an inordinate appetite for entertainment or information. Indeed, I do not turn them on in order to *hear* anything. (1993: 236)

Radioman is motivated to act in ways that he believes will bring it about that radios are turned on. So, for example, he will press a button when he believes that this will turn on a radio. Nevertheless, I'll argue, he has no desire to turn on radios.¹⁴

Here are my assumptions about Quinn's example. First, radioman's actions are *intentional*, rather than mere reflexes, since his behavior is mediated by his beliefs and desires. He won't press the button when he believes it is wired up to explosives, for example, since he doesn't want to cause an explosion. Second, he has no *instrumental* desire, reason, or motivation to turn on radios. In particular, he is not disposed to derive any pleasure from turning them on or displeasure from seeing them turned off. And, finally, although he is disposed to turn on radios, he is not disposed to feel any *desire* to turn them on. When he imagines turning on radios, he feels no anticipatory pleasure in what he is imagining. As Quinn says, the motivational disposition is "all there is to the state".

The motivational theory implies that radioman has a desire to turn on radios, since he is motivated to act in ways that he believes will turn them on. And, since he has no instrumental desire to turn on radios, it must be an intrinsic desire. And yet radioman violates the following normative constraint on intrinsic desire:

The Normative Constraint: All intrinsic desires provide justifying reasons for action.¹⁵

This normative constraint explains why you always have some instrumental reason to do what you have reason to believe will satisfy your intrinsic desires. If you prefer the taste of gin rather than vodka, for example, then you have some reason to order it from the bar. Of course, these desire-based reasons can be rebutted or undercut – say, by evidence that the gin is mixed with gasoline. Nevertheless, your intrinsic desires give you defeasible reasons that can justify action in the absence of any such defeating considerations. When your evidence is misleading – perhaps you have no inkling that the gin is mixed with gasoline – your intrinsic desires can motivate actions that are *bad* for you. Even so, your intrinsic

¹⁴ Smithies and Weiss (2019) provide supporting arguments that appeal to higher-level requirements on justifying reasons.

¹⁵ This normative constraint doesn't imply the Humean theory of reasons, according to which all justifying reasons for action are provided by intrinsic desires. It is compatible, for example, with the Kantian thesis that morality provides reasons for action that need not align with your desires.

desires give you justifying reasons that can make it instrumentally reasonable, rational, or justified to act in ways that might sometimes turn out badly for you.¹⁶

Radioman's motivational disposition violates this normative constraint on intrinsic desire. As Quinn observes, merely being motivated to turn on radios gives him no reason whatsoever to do so:

I cannot see how this bizarre functional state in itself gives me even a *prima facie* reason to turn on radios, even those I can see to be available for cost-free on turning. It may help explain, causally, why I turn on a particular radio, but it does not make the act sensible. (1993: 237)

It is not that radioman has reasons to turn on radios that are defeated by countervailing considerations, since we can stipulate that there is no cost in turning them on. Even so, the act of turning on radios seems utterly pointless when considered from his own perspective.

A related point is that there is nothing good for radioman about turning on radios despite his motivation to turn them on. This strengthens the case for formulating the Resonance Constraint in terms of justifying reasons for action, rather than motivation. The mere fact that you are motivated to pursue an outcome doesn't make it even *pro tanto* good for you, since it must resonate with intrinsic desires or other pro-attitudes that give you justifying reasons for action.

In summary, here is my argument against the motivational theory of desire. We can stipulate that radioman has no instrumental desire to turn on radios: if he has any such desire, it must be an intrinsic one. But all intrinsic desires provide justifying reasons for action. And yet radioman has no such reason to turn on radios. Therefore, he has no desire to turn on radios. Even so, he is motivated to do whatever he believes will bring it about that radios are turned on. Hence, the motivational theory is false: motivation is not sufficient for desire.

7. How Desires Justify Action

Our next question is what radioman needs to have an intrinsic desire that gives him justifying reasons to turn on radios. What is the missing ingredient? In this section, I'll argue that radioman has no

¹⁶ In this context, I am using the ambiguous or context-sensitive term 'reason' in what is often called the *subjective* sense, rather than the *objective* sense (e.g. Schroeder 2008). It counts in favor of an action by making it reasonable, rational, or justified in light of your subjective evidence, rather than good in light of the objective facts.

justifying reason to turn on radios because he has no disposition to feel any pleasure in the thought of turning them on. More generally, I'll argue for the following hedonic constraint on intrinsic desire:

The Hedonic Constraint: All intrinsic desires and pro-attitudes that provide justifying reasons for action do so by disposing you to feel intrinsic pleasure in their contents.

On this view, the common feature that unifies pro-attitudes that provide justifying reasons for action is not their functional disposition to motivate action, but rather their phenomenal disposition to make you feel pleased about something for its own sake.¹⁷

My argument for the hedonic constraint proceeds in several steps. First, when you desire something for its own sake, you're disposed to *feel the desire* for it when you imagine it. The feeling of desire is something we all know from our own personal experience. Everyone knows what it's like to feel the desire for a slice of cake, a short nap, or a career change. Moreover, these feelings of desire are not just brute sensations with no intentionality. On the contrary, they are valenced orientations towards an intentional content: they are ways of feeling positively about what we desire.

Second, we can explain this valenced intentionality by noting that feelings of desire are hedonic feelings: you *feel pleasure* in what you desire when you imagine it. Smithies and Weiss (2019: 44) raise the objection that feelings of desire tend to *feel unpleasant* when they are frustrated – for example, in cases of unrequited love. There is no conflict here, however, since you can feel pleasure in imagining that your desire is satisfied while also feeling displeasure in the knowledge that your desire is frustrated. Indeed, this hedonic ambivalence explains why the experience of unrequited love is so bittersweet.

Third, when you desire something for its own sake, you're disposed to feel pleasure in what you desire for its own sake and not for any further reason. Suppose you want to take a long walk in the hot sunshine only so that you can buy some beer to drink. In that case, you're not disposed to feel pleasure in the thought of taking the long walk for its own sake, although you might feel pleasure in the thought of doing so in order to get beer. In contrast, you're disposed to feel pleasure in the thought of drinking beer for its own sake and not for any further reason. More generally, when you intrinsically desire something, you're disposed to feel *intrinsic pleasure* in what you desire.

¹⁷ Others who highlight hedonic or affective dimensions of desire include Vadas (1984), Davis (1986), Strawson (1994), Chang (2004), Heathwood (2019), Smithies and Weiss (2019), Siewert (2021), and Smithies (MS).

Fourth, feeling – or being disposed to feel – intrinsically pleased about something for its own sake gives you some justifying reason to do what you have reason to believe will bring it about. If you're disposed to feel intrinsic pleasure in the thought of drinking beer, for example, then this gives you some justifying reason to take the long walk in the sun when you have reason to believe that this is the only way to get beer. Of course, these justifying reasons for action can be rebutted or undercut by evidence that there is nothing of value to be gained from drinking beer.

This doesn't imply here that we are selfishly motivated by our own pleasure whenever we act on our intrinsic desires.¹⁸ After all, our intrinsic desires are not always concerned with our own pleasure. Most parents, for example, intrinsically desire their children's welfare as well as their own. This disposes us to feel intrinsic pleasure in imagining our children's welfare, which gives us justifying reasons to do what we have reason to believe will promote this outcome. When we act on this desire, however, we need not be acting for the reason that it will be pleasurable to do so. Parents can feel pleasure in caring for their children without thereby acting for their own sake, rather than solely for the sake of their children. Even so, feeling pleasure in the thought of our children's welfare gives us justifying reasons to promote this outcome by representing it in a positive way.

Putting all this together, the suggestion is that our intrinsic desires and pro-attitudes provide us with justifying reasons for action by disposing us to feel intrinsic pleasure in what we desire. What I'm suggesting, in effect, is a specific version of the proposal that our intrinsic desires provide us with justifying reasons for action by presenting what we desire under "the guise of the good".¹⁹

Quinn defends the related proposal that your desires give you justifying reasons for action only if you believe there is some *value* in so acting. To have any such reason, he says, "I need the *thought* that the direction in which I am psychologically pointed leads to something good" (1993: 242). Arguably, however, this is an over-intellectualization. Children and animals can act rationally on their desires without any capacity for evaluative thought, judgment, or belief. Indeed, even mature adults can act for justifying reasons provided by desires that conflict with their evaluative beliefs. As Nomy

¹⁸ Feinberg (1971: 185-7) makes essentially the same point, although Quinn (1993: 243) seems to miss the point in his critique of hedonic theories of desire.

¹⁹ This slogan is defended in various forms by Stampe (1987), Oddie (2005), Tenenbaum (2007), Schapiro (2009), and Schafer (2013), although none of them defend it by appealing to a hedonic theory of desire. I'm not claiming that we always *act* under the guide of the good, since our actions may be motivated by our beliefs, rather than our desires. Perhaps, as Velleman (1992) argues, we can make sometimes perverse decisions to do things we don't want to do precisely because we believe there is nothing good about them. And yet such actions are not justified by our desires.

Arpaly (2002) argues, Huck Finn acts rationally and morally in helping Jim to escape slavery, despite believing that he ought to turn Jim over to the authorities. If so, Quinn's proposal is false: evaluative belief is not *necessary* for your desires to provide you with justifying reasons for action.

More importantly, evaluative belief is not *sufficient* to give you justifying reasons for action. Suppose radioman believes for no particular reason that there is some value in turning on radios. This unjustified belief gives him no justifying reason to turn them on. After all, beliefs about value – like beliefs about anything else – are justified only if they are based on justifying reasons. If your beliefs are not based on justifying reasons, then they cannot provide justifying reasons for action, since only justified beliefs can justify action. If your beliefs are based on justifying reasons, however, then they cannot provide justifying reasons for action, since they merely transmit justification from the reasons on which they are based. Either way, beliefs about value cannot provide justifying reasons for action.

In contrast, intrinsic desires can provide justifying reasons for action because they dispose you to feel intrinsic pleasure in what you desire. When you feel intrinsic pleasure in what you desire, you feel pleased about what you desire for no further reason. And feeling pleased about something for no further reason gives you some reason to do what you have reason to believe will bring it about. Thus, intrinsic desires can block the regress of justifying reasons in a way that evaluative beliefs cannot.

Even so, my proposal captures a grain of truth in Quinn's proposal that our desires provide reasons for action by representing what we desire under the guise of the good. Desire represents value in a psychologically more primitive way than belief: belief represents value in virtue of its *content*, whereas desire represents value in virtue of its *attitude-type*. More specifically, we represent what we desire as good by virtue of our hedonic reactions to what we desire: we're disposed to feel pleasure in what we desire when we imagine it. This is why our desires give us justifying reasons for action.

8. Hedonic Sentientism

We now have all the ingredients in place to argue against Chalmers' inclusive version of sentientism and in favor of a more restrictive version of hedonic sentientism, according to which the capacity for hedonic feelings of pleasure and displeasure is required to be a welfare subject with moral status.

The argument begins with the Resonance Constraint, according to which something is good for you only if it resonates with your intrinsic desires and other pro-attitudes, including the attitude of intrinsic pleasure. What unifies these pro-attitudes is not their causal role in motivating action but their normative role in providing justifying reasons for action. This generates the following normative constraint on welfare subjects:

The Resonance Constraint on Welfare Subjects: A being is a welfare subject only if it has intrinsic desires or other pro-attitudes that provide it with justifying reasons for action.

We now combine this with the further premise that what unifies all intrinsic pro-attitudes that provide justifying reasons for action is not their functional disposition to motivate action, but rather their phenomenal disposition to make you feel pleasure:

The Hedonic Constraint: All intrinsic desires and pro-attitudes that provide justifying reasons for action do so by disposing you to feel intrinsic pleasure in their contents.

These two premises imply the conclusion that distinguishes hedonic sentientism from inclusive sentientism: a being is a welfare subject only if it has some disposition to feel intrinsic pleasure.

Of course, no argument is stronger than the premises it relies upon. Even so, these premises are plausible because they explain why there is nothing good for radioman about turning on radios despite his motivation to produce this outcome. Turning on radios is not good for him because it doesn't resonate with any of his dispositions to experience positively valenced feelings of pleasure. What goes for radioman, however, goes equally for Vulcans. They may be motivated to build families, make money, or advance science, but they are not disposed to experience any pleasure or displeasure either in producing these outcomes or in imagining them. As a result, these outcomes are not good for them, and they have no self-interested reason to pursue them. Nothing is good or bad for them because they are not disposed to feel pleasure or displeasure in anything.

If this argument is sound, then it is a mistake to suppose that Vulcans are welfare subjects with moral status. Their lives may have some intrinsic value, since they are alive and conscious, but their lives have no intrinsic *welfare* value for them. As a result, their lives may have some derived moral significance for welfare subjects, but they have no intrinsic moral significance of their own. Perhaps intuitions to the contrary rely too heavily on conflating humanoid Vulcans with the human beings they superficially resemble.

How might a Vulcan rights activist respond to this argument? One strategy is to argue that Vulcans are welfare subjects because their motivations have a degree of functional complexity that radioman fails to exhibit. My immediate reaction is that it makes no difference how much functional complexity we add to radioman's motivations: he has no justifying reason to turn on radios so long as

he has no disposition to feel any pleasure in the thought of doing so. In any case, Chalmers cannot adopt this strategy, since it abandons *sentientism*: it includes conscious Vulcans within the moral circle only by including unconscious zombies too.

A different strategy is to argue that Vulcans are distinguished from zombies as welfare subjects because they can make conscious judgments about their own welfare. This strategy upholds sentientism, since only conscious beings can make conscious judgments. At the same time, it avoids any obvious commitment to hedonic sentientism, since you need not feel any pleasure or displeasure about what you judge to be good or bad for you. It is rather less clear, however, that you can make judgments about your own welfare without any capacity to feel pleasure or displeasure at all. My main objection is that value judgments are not pro-attitudes of the right kind to constrain welfare: like instrumental desires, they cannot provide justifying reasons for action, since they merely transmit the justifying reasons on which they are based. The only pro-attitudes that play this normative role are those that dispose you to feel intrinsic pleasure in their contents. Since Vulcans cannot feel pleasure, they have no welfare to make conscious judgments about. In any case, Chalmers cannot adopt this strategy, since it abandons *inclusion*: it includes some Vulcans within the moral circle only by excluding others. After all, not all conscious beings are conceptually sophisticated enough to make value judgments. As a result, this strategy is unlikely to satisfy any Vulcan rights activist, since it only protects the rights of the privileged while neglecting those of the most vulnerable.

Ultimately, I think the only viable strategy for defending inclusive sentientism is to abandon the Resonance Constraint. On this view, all conscious beings are welfare subjects regardless of whether or not they have any pro-attitudes towards their own welfare. The problem with this strategy is that it includes humanoid Vulcans within the moral circle only by including Blobs. Personally, I cannot bring myself to accept this inclusive version of sentientism, since I find it so implausible that Blobs are welfare subjects: their conscious life may have some intrinsic value, but it has no welfare value for them. Ultimately, though, my goal is to show that is not the only option for proponents of sentientism. There is a more restrictive version of hedonic sentientism that not only avoids this implausible result, but also provides a better explanation of why sentientism is true.²⁰

²⁰ Earlier versions of this paper were presented at Patrick Butlin's Digital Minds Workshop at the Future of Humanity Institute in November 2021, Uriah Kriegel's Value of Consciousness series in September 2023, and Andreas Mogensen's Mind and Moral State series at the Global Priorities Institute in May 2024. I am grateful to audiences on those occasions and also for conversation and email correspondence with David Chalmers, Tyler Cook, Cameron Green, Preston Lennon, Eden Lin, Luke Roelofs, Charles Siewert, and Inchul Yum.

References

- Arpaly, Nomy. 2002. 'Moral Worth.' *Journal of Philosophy* 99: 223-245.
- Bayne, Tim, and Montague, Michelle. 2011. *Cognitive Phenomenology*. New York: Oxford University Press.
- Bradford, Gwen. 2023. 'Consciousness and Welfare Subjectivity.' *Noûs* 57: 905-921.
- Chalmers, David. 2022. *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: Norton.
- Chang, Ruth. 2004. 'Can Desires Provide Reasons for Action?' In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by Jay Wallace, Philip Petit, Samuel Scheffler, and Michael Smith, 56-90. New York: Oxford University Press.
- Davis, Wayne. 1986. 'Two Senses of Desire.' In *The Ways of Desire*, edited by Joel Marks, 63-82. Chicago: Precedent.
- Feldman, Fred. 1988. 'Two Questions about Pleasure.' In *Philosophical Analysis: A Defense by Example*, edited by David Austin, 59-81. Dordrecht: Kluwer.
- Fletcher, Guy. 2013. 'A Fresh Start for the Objective-List Theory of Well-Being.' *Utilitas* 25: 206-220.
- Heathwood, Chris. 2006. 'Desire Satisfactionism and Hedonism.' *Philosophical Studies* 128: 539-563.
- Heathwood, Chris. 2007. 'The Reduction of Sensory Pleasure to Desire.' *Philosophical Studies* 133: 23-44.
- Heathwood, Chris. 2019. 'Which Desires are Relevant to Well-Being?' *Noûs* 53: 664-688.
- Jacobson, Hilla. 2019. 'Not Only a Messenger: Towards an Attitudinal-Representational Theory of Pain.' *Philosophy and Phenomenological Research* 99: 382-408.
- Kriegel, Uriah. Forthcoming. 'The Value of Consciousness to the One Who Has It.' In *The Importance of Being Conscious*, edited by Geoffrey Lee and Adam Pautz. New York: Oxford University Press.
- Lee, Andrew. 2018. 'Is Consciousness Intrinsically Valuable?' *Philosophical Studies* 176: 655-671.
- Lewis, David. 1980. Mad Pain and Martian Pain. In *Readings in the Philosophy of Psychology, Volume 1*, edited by Ned Block, 216-222. Harvard University Press.
- Lin, Eden. 2020. 'Attitudinal and Phenomenological Theories of Pleasure.' *Philosophy and Phenomenological Research* 100: 510-524.
- Lin, Eden. 2021. 'The Experience Requirement on Well-Being.' *Philosophical Studies* 178: 867-886.
- Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

- Oddie, Graham. 2005. *Value, Reality, and Desire*. New York: Oxford University Press.
- Parfit, Derek. 1984. *Reasons and Persons*. New York: Oxford University Press.
- Quinn, Warren. 1993. 'Putting Rationality in its Place.' In *Morality and Action*, by Warren Quinn, 228-255. Cambridge: Cambridge University Press.
- Roelofs, Luke. 2023. 'Sentientism, Motivation, and Philosophical Vulcans.' *Pacific Philosophical Quarterly*. 104: 301-323.
- Railton, Peter. 1986. 'Facts and Values.' *Philosophical Topics* 14: 5-31.
- Rosati, Connie. 1996. 'Internalism and the Good for a Person.' *Ethics* 106: 297-326.
- Rosati, Connie. 2009. 'Relational Good and the Multiplicity Problem.' *Philosophical Issues* 19: 205-234.
- Schafer, Karl. 2013. 'Perception and the Rational Force of Desire.' *Journal of Philosophy* 110: 258-281.
- Schapiro, Tamar. 2009. 'The Nature of Inclination.' *Ethics* 119: 229-256.
- Schroeder, Mark. 2008. 'Having Reasons.' *Philosophical Studies* 139: 57-71.
- Shepherd, Joshua. 2018. *Consciousness and Moral Status*. New York: Routledge.
- Siewert, Charles. 2021. 'Consciousness: Value, Concern, Respect.' *Oxford Studies in Philosophy of Mind* 1: 3-40.
- Singer, Peter. 1975. *Animal Liberation*. New York: Harper Collins.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.
- Smithies, Declan. MS. 'Hedonic Theories of Desire.'
- Smithies, Declan and Weiss, Jeremy. 2019. 'Affective Experience, Desire, and Reasons for Action.' *Analytic Philosophy* 60: 27-54.
- Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT Press.
- Stampe, Dennis. 1987. 'The Authority of Desire.' *Philosophical Review* 96: 335-381.
- Strawson, Galen. 1994. *Mental Reality*. New York: Oxford University Press.
- Tenenbaum, Sergio. 2007. *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press.
- Vadas, Melissa. 1984. 'Affective and Non-Affective Desire.' *Philosophy and Phenomenological Research* 45: 273-279.
- Van der Deijl, Willem. 2019. 'Is Pleasure All That is Good About Experience?' *Philosophical Studies* 176: 1769-1787.
- Van der Deijl, Willem. 2020. 'The Sentience Argument for Experientialism about Welfare.' *Philosophical Studies* 178: 187-208.
- Velleman, David. 1992. 'The Guise of the Good.' *Noûs* 26: 3-26.

Velleman, David. 1998. 'Is Motivation Internal to Value?' In *Preferences*, edited by Christoph Fehige and Ulla Wessels, 88-102. Berlin: De Gruyter.