

A Sensible Speciesism?

Christopher Grau

Abstract: In his essay “The Human Prejudice” Bernard Williams presented a sophisticated defense of the moral relevance of the concept “human being”. Here I offer both an analysis of his essay and a defense of his conclusions against criticisms made by Julian Savulescu and Peter Singer. After a discussion of the structure of Williams’s argument, I focus on several complaints from Savulescu: that Williams underestimates the similarities between speciesism and racism or sexism, that Williams relies on a disputable internalism about reasons to make his case, that Williams ignores the arbitrariness of species membership, and (a complaint also made by Singer) that Williams attacks a straw man in considering only negative utilitarianism when criticizing the idea of an Impartial Observer. I defend Williams against these charges and end with a brief discussion of how his alien invasion example illustrates the fundamentally different philosophical sensibility of Bernard Williams in comparison to critics like Savulescu and Singer.

Keywords: humanism; speciesism; criterialism; Bernard Williams.

Too often, philosophers’ contributions to these questions [about the nature of values] seem designed only to reduce the number of thoughts that people can have, by suggesting that they have no right to some conceptions that they have or think they have. But equally philosophy should be able to liberate, by suggesting to people that they really have a right to some conception, which has been condemned by a simple or restrictive notion of how we may reasonably think.

Bernard Williams,

‘Must a concern for the environment be centered on human beings?’ (1995: 233)

And it is perhaps worth saying before I begin that what I am trying to do here, what Gaita was trying to do, is hard: hard to articulate, hard to state clearly, hard to see in the round, hard to apply. Repeatedly, criterialism wins out in philoso-

phers' discussions of personhood because, conversely, it is dead easy; it is such a simple and straightforward view of personhood. There are plenty of spurious analogies between science and ethics available, but at any rate we can trust this one: no more in the case of the nature of the person than in the case of particle physics does the fact that a view is simple in any way improve its chances of being true.

Timothy Chappell,
'On the Very Idea of Criteria for Personhood' (2011: 21)

1. *Introduction*

There are people who embrace the sort of position that has come to be known as 'speciesism' because they believe it provides support for their own rather callous and dismissive attitudes towards non-human animals. As you might expect, the philosophical defenses marshaled for this cause tend to be shallow and easily knocked down. There are others, however, who have defended a 'speciesist' view that come from a rather different place and are motivated by more humane considerations. Philosophers such as Cora Diamond and Elizabeth Anderson have put forward views that get grouped under the 'speciesist' label yet they have made clear their concerns regarding our general tendency to *underappreciate* the moral status of animals (Diamond 1978; 1991; Anderson 2004). In a similar spirit, Bernard Williams has offered a sophisticated defense of the moral relevance of the concept of humanity, and it is one that he clearly believes is compatible with sincere efforts to *extend* our range of concerns to non-human animals (Williams 2008). I think Williams's original contribution to these debates is worthy of serious consideration, and here I want to both analyze and defend his account while addressing criticisms of it that have recently been offered by Julian Savulescu (2010) and Peter Singer (2009).

Before I move on to discuss Williams's argument, however, I want to say a bit more about my own motivations for defending a position that some like to call 'speciesist'. I find the moral arguments for vegetarianism compelling, and the moral arguments against factory farming utterly convincing. I am not interested in theoretically justifying a framework that allows for excusing animal suffering. Rather, I'm interested in defending a moral outlook that recognizes that being a human can matter morally. In other words, I believe that, along with other categories, the category 'human being' has moral relevance, and I think attempts to deny this result in an impoverished moral landscape. Recognizing the moral importance of such a category does not imply that a

non-human animal could not (in certain circumstances) rightly be regarded as having a moral status equivalent to or even higher than a human. It is one thing to claim membership in the class ‘human being’ can count among legitimate moral considerations, it is another to decide *when* it should count and *how much* it counts for. Those are difficult tasks I do not attempt here.

2. *Williams’s argument*

In order to appreciate Williams’s more specific claims about morality and ethics it is important to understand his background assumptions, for they are rather different than those of many (perhaps most) ethicists. In ‘The Human Prejudice’ Williams (2008) does make these background assumptions fairly clear, but nonetheless the essay has been misread and misunderstood by philosophers working from a different place.¹ ‘The Human Prejudice’ is a sustained attempt to consider what ethical work the concept of ‘the human’ can do once one has recognized that the ideas of ‘absolute importance’ and a (moral) ‘cosmic point of view’ are either false or incoherent. The issue of coming to accept a ‘disenchanted’ world is one that Williams has pursued in many places – he was perhaps our most eloquent 20th century philosopher to repeatedly warn against the dangers of harboring false hopes regarding tempting but misguided naive conceptions of moral thought. He wisely cautioned us to be wary of universal ethical claims that covertly require or presuppose the impossible.

He begins this essay by reminding us that our ethics comes *from us*.² It is not surprising, then, that it is in large part *about us*.³ (This isn’t a *necessary* connection – it is just an understandable feature of the content of our ethical thought given its source.) So, it also isn’t surprising that our species membership – our humanity – figures importantly in our ethical thought. We don’t need to posit a mistaken malingered trace of the notion of ‘Cosmic Importance’ to explain why we care about human beings *as human beings*. We need not think humanity matters *period*. It is enough that, “human beings matter *to us*” (139).⁴

Many philosophers (some of whom fall under such labels as ‘personist’ or ‘criterialist’ or ‘moral individualist’) claim that this commitment stands in need

¹ Future references to this work will cite page number only.

² “There is no other point of view except ours in which our activities can have or lack a significance” (137).

³ “Whether a creature is a human being or not makes a large difference, a lot of the time, to the ways we treat that creature or at least think that we should treat it” (138).

⁴ This echoes his earlier discussion in *Ethics and the Limits of Philosophy*: “The word ‘speciesism’ has been used for an attitude some regard as our ultimate prejudice, that in favor of humanity. It is more revealingly called “humanism,” and it is not a prejudice...” (Williams 1990: 118).

of justification. After all, at first glance an apparent preference towards human beings in virtue of a biological fact like species membership can look a lot like other biology-based prejudices we now reject as unjustified, such as racism and sexism. (Thus the unhappy term ‘speciesism’).

In the case of racism and sexism, the justifications that have been put forward don’t hold water. In the case of humanism, Williams points out that most humanists have not even *attempted* to offer the sort of justifications provided by racists and sexists. (There is an interesting structural difference here.) Perhaps this is just because there are no non-humans around in a position to demand such justifications. But Williams thinks there is another factor that is relevant: an attempt at justification in this context is likely to appeal to those features that supposedly distinguish human beings from other creatures (e.g., our ability to use language, our capacity for culture, etc.). However, what is it that makes such features valuable? It is hard to see how they could be defended from some truly impartial perspective as ‘simply better’ than, say, the ‘amazing’ capacities of insects. (141) We’re kidding ourselves to think that – we would be falling prey again to the idea that there’s a cosmic scale of importance that we happen to score highly on. There is no reason, however, to believe that such features are, as he puts it, being “cheered on by the universe” (144).

So, aspects of human life like culture, language, etc. are valuable because they are valuable *to us*. Why are they valuable to us? Presumably in large part because we happen to possess them. This, however, can seem to be “another expression of the human prejudice” (141). Not only are these features not *more* obviously justifiable than a commitment to humanity, it appears as though they may derive from our *prior* commitment to humanity.

Perhaps this just shows us how *deep* and perverse the prejudice is? Williams suggests our commitment here does go quite deep, but he doesn’t think this makes it illicit. To help make his case he considers the alternative: what does moral thought tend to look like *without* the human prejudice? Peter Singer provides a valuable case study, both because he’s one of the most prominent living ethicists and because his approach is widely shared by others working in moral philosophy. Williams points out that a rejection of humanism tends to bring with it the acceptance instead of “some more substantial set of properties, supposedly better fitted to give a reason” (142). (This is a view he labels ‘personism’. Others have called such an approach ‘criterialism’.) Popular properties for the personist’s list of criteria have included sentience, rationality, and the capacity for moral agency.

This approach raises an obvious question: why those properties and not others? If the reason is that they matter *to us*, well we are back where we started: being a human being *also* matters to us. If the reason is that they are

‘simply better’, the defender needs to explain what this could mean in a way that doesn’t end up resorting to an enchanted picture of the universe:

A different answer would be that it is *simply better* that the world should instantiate the fancy properties of personhood, and not *simply better* that human beings as such should flourish. But that is once more our now familiar friend, absolute importance, that survivor from the enchanted world, bringing with it the equally familiar and encouraging thought that the properties we possess – well, most of us, not counting the infants, the Alzheimer’s patients, and some others – are being cheered on by the universe (144).

Philosophers like Singer certainly don’t *think* they are presupposing an enchanted view of the universe. So what does ‘simply better’ mean for them? Many invoke something like the idea of an Ideal or Impartial Observer to justify the objectivity of their value judgments. Williams suggests that this sort of model is far from free of remnants of an enchanted worldview. (It even appears to be modeled on Christ!).

More importantly, however, Williams points out that those who endorse the IO model don’t appear to have taken seriously enough the idea they are proposing. Let’s try to take seriously the idea of a hypothetical impartial observer taking on the suffering of all individuals. What would result? Williams claims that the impartial observer would suffer “an ultimate horror, an unendurable nightmare” and anyone looking to base behavior on such insight would want to “annihilate the planet” (147).

The IO approach has things inside out. We can’t get purchase from such an external perspective, but we don’t need such a perspective: we have the ethical resources already, on the inside. We already have reasons “to listen to our sympathies and extend them, not only to wider groups of human beings, but into a concern for other animals, so far as they are in our power. This is already a human disposition” (147). (That we often refer to such extensions as ‘humane’ is telling).

We can act intelligibly from these concerns only if we see them as aspects of human life. It is not an accident or a limitation or a prejudice that we cannot care equally about all the suffering in the world: it is a condition of our existence and our sanity. [...] it is a total illusion to think that this enterprise can be licensed in some respects and condemned by others by credentials that come from another source, a source that is not already involved in the peculiarities of the human enterprise (147).⁵

⁵ He points out that there is a funny irony here: the defender of the IO ends up embracing something like the notion of Cosmic Importance, though in a Lutheran style (i.e., we humans don’t score high); but that is combined with an Enlightenment Optimism in the power of rationality and “these theories” to save us.

In essence, Williams is suggesting that the proponent of personism doesn't have the philosophical resources she thinks she has. She has only the same grounding for her position as we do, and is misled by the pretensions of theory to think otherwise. Once the personist sees that she is on the same metaphysical footing as the humanist, however, she ought to reconsider her revisionary tendencies. What's so bad about an ethical attachment to the human species after all?

Williams considers the possibility that there could be another point of view beyond our own: not a God's eye perspective but an *alien* perspective. What does this possibility show us about the status of the human prejudice? What if the aliens were smarter and better than us and concluded that the universe would be better off if our prejudices were removed (or *we* were removed)? Should we go along with them?⁶

Williams suspects personists might end up as collaborators. Or at least he's not sure how they could resist that call without falling back on something like the human prejudice: "I do not see how they could be sure that they were not the victims of what in their terms would be just another self-serving prejudice" (152).

How could the resisters justify their resistance? Williams draws an analogy with ethnic or cultural loyalty:

The situation that this fantasy presents is in some ways familiar. It is like that of a human group defending its cultural, possibly ethnic, identity against some other human group which claims to dominate or assimilate them. [...] The relevant concept is something like: loyalty to, or identity with, one's ethnic or cultural grouping; and in the fantasy case the ethical concept is: loyalty to, or identity with, one's species. Moreover – and this is the main lesson of this fantasy – this is an ethical concept we already have. This is the ethical concept that is at work when, to the puzzlement of the critics, we afford special consideration to human beings because they are human beings. [...] So the idea of there being an ethical concept that appeals to our species membership is entirely coherent (150).

Of course this ethical idea is not uncontestable, just as the idea of loyalty to an ethnic or cultural group is not uncontroversial. Williams is acutely aware of "coercive rhetoric, the lies about difference, and the sheer violence that are often associated with such ideas" (151). However, he appears to be on the side of those who "may be respectful of the energizing power of such conceptions, and of the sense they can give of a life that has a rich and particular character" (150). He ends his essay with the memorable line: "Personally I think there

⁶ His discussion here is similar to his earlier discussion in Williams 1990 (which also includes the idea of choosing between a confederation and a nonaggression pact). See pp. 102-104.

are many things to loathe about human beings, but their sense of their ethical identity as a species is not one of them” (152).

3. *Savulescu's response*

In his essay ‘The Human Prejudice and the Moral Status of Enhanced Beings: What Do We Owe the Gods?’ Oxford bioethicist Julian Savulescu argues that we have duties to posthuman entities and radically enhanced human beings and, in the process, offers an extended and detailed critique of Williams’s defense of humanism. Savulescu acknowledges that humanism/speciesism is widely held as a ‘folk view’ and he describes Williams’s essay as “the most sophisticated defense of that view” (2010: 216).⁷

4. *Structural Differences*

Savulescu begins his critique by calling into question Williams’s assertion that ‘It’s a human being’ can *really* operate as a reason, suggesting it is indeed just another prejudice like racism or sexism (219). Savulescu thinks he can show this by attacking Williams’s claim that there is a ‘structural difference’ between humanism on the one hand and racism and sexism on the other. However, on this topic Savulescu seems to fundamentally misunderstand Williams. Williams is talking about how these prejudices have *actually* functioned most of the time. There are interesting differences here (which he highlights): racists and sexists tend to either be operating at a level prior to assuming race and sex provide reasons (i.e. at the level of what he calls a ‘barely articulated practice of discrimination’), or else they are beyond that and are attempting to offer justifications (which turn out to be mere rationalizations) for their ‘ism’ (e.g., justifications that appeal to supposed differences in intelligence). The humanist, however, tends to be doing something different:

On the one hand, it is not simply a matter of inarticulate or unexpressed discrimination: it is no secret that we are in favor of human rights, for instance. On the other hand, ‘it’s a human being’ does seem to operate as a reason, but it does not seem to be helped out by some further reach of supposedly more relevant reasons, of the kind which in the other case of prejudice turned out to be rationalizations (Williams 2008: 140).

⁷ Future citations to Savulescu will be to page number only. Peter Singer’s brief response to Williams also begins with kind words towards Williams: “Hence a defense of speciesism by a philosopher as skilled and eloquent as Bernard Williams is significant and will prove instructive” (Singer 2009: 97).

Unconvinced, Savulescu points out the *mere contingency* of these structural differences, and imagines a possible world where it could all be different:

Moreover, it is a contingent feature of other prejudices that they appeal to rationalizations. They could be simply like the human prejudice. [...] I have argued that racists and sexists can use race and sex in the same ways as humanists use human. The reverse is true. Humanist could endorse the human prejudice in the early basic way. They could go on treating other human beings badly, just as racists treated blacks badly, without thinking to justify it. [...] Or they could offer bad rationalizations – like humans have souls (219-220).

Savulescu seems to think that Williams is asserting these structural differences as *necessary*, but Williams here (and elsewhere in his philosophical writings) is not particularly interested in talking about what is necessarily the case. As he says at one point in a response essay from *World, Mind, Ethics*: “philosophy needs to give up its long obsession with necessity” (Williams 1995b: 224).

Savulescu also implies that Williams is using the ‘structural difference’ point to convince us that humanism isn’t as disreputable as racism or sexism, but Williams explicitly *acknowledges* that the differences here don’t “necessarily show that it isn’t a prejudice” (141). Williams points out, however, that these differences do suggest that this attitude is *deeper* (and so different) than those others, in the sense that attempts at justification here are going to usually fail in a way that is different from how the racists’ and sexists’ justifications typically turn out to be rationalizations. For the humanist, an attempt to further justify humanism is likely to fail because a concern for humanity turns out to be (for many of us) a more basic value (or at least just as basic) as the values likely to be appealed to in the purported justification. That there could be *other* types of people who think differently about these things does little to weaken Williams’s point here.⁸ (A contingent fact about us can still be a *deep* fact about us).

Having said that, I acknowledge that despite Williams’s general emphasis on the actual (as opposed to merely possible) texture of human life, in this essay he somewhat uncharacteristically *does* go on to consider a hypothetical scenario concerning possible alien life, and I’ll consider Savulescu’s remarks on that scenario later on.

⁸ This consideration of the ‘deepness’ of the attitude leads Williams to point out a further important dissimilarity between humanism and the bad ‘isms’: “Oppressed human groups come of age in the search for emancipation when they speak for themselves [...] but other animals will never come of age” (141). (As he puts in Williams 1990, the “only question here is how we should treat them”.) Interestingly, Savulescu ignores this point (perhaps because he views it as merely a ‘contingent’ fact about how animals happen to currently be), even though Williams comes back to it later in his essay (148).

5. *Desire-Based Reasons vs. Value-Based Reasons*

Savulescu goes on in his essay to suggest that Williams is guilty of assuming that all of his opponents are necessarily clinging to some questionable notion of a cosmic order. Savulescu rightly points out that this need not be the case, saying: “personism, however, need not have any commitment to any supernatural entity or cosmic order [...] Personists appeal to properties human beings value and to normative properties, like the badness of suffering” (226).

The former claim (that personists appeal to properties human beings typically value) is uncontentious, and I think Williams has no real argument with someone who simply claims to value sentience and rationality but doesn't value humanity *per se*. (His 'argument' will just consist of pointing out that such a person's values have no firmer foundation than our own humanism/speciesism.) However, the latter suggestion in that passage (the appeal to 'normative properties') implies what has come to be called *externalism* about reasons, and it seems likely, given his other work on the topic, that Williams has in mind that sort of meta-ethical view (among many others) when he criticizes moral philosophers for clinging to remnants of an enchanted worldview.

Whether Williams had externalism in mind with such remarks or not, Savulescu certainly thinks that he did, and he spends a fair amount of time talking about the internalism/externalism debate. Unfortunately, he somewhat question-beggingly puts the debate in terms of 'value-based reasons' (VB) vs. 'desire-based reasons' (DB) with VB reasons corresponding to what Williams means by 'external' and DB reasons 'internal'. Savulescu couches many of his criticisms of Williams in these terms (VB vs. DB), suggesting that the feasibility of Williams's argument regarding the concept of humanity somehow hinges on this impossibility of the existence of external reasons.

Savulescu's discussion of this debate further illuminates how removed his perspective on moral philosophy is from Williams's own. The basic difference between the internalist and externalist is, as described by the SEP entry on this topic: “[for the externalist] the possibility of being motivated to do *A* can be explained by the existence of a reason to do *A*, while Williams' view is that the existence of a reason to do *A* must be explained by the possibility of being motivated to do *A*” (Finlay & Schroeder 2012). Williams himself lays out his position slightly differently in different places, but here's one sketch he offered in an exchange with John McDowell:

The central idea is that if B can say truly of A that A has a reason to phi, then (leaving aside the qualifications needed because it may not be his strongest reason) there must be a sound deliberative route to phi-ing which starts from A's existing motiva-

tions. It follows that what an agent has a reason to do will be a function of what I called his 'S' – that is to say, the existing set of his motivational states (Williams 1995b: 186-187).

Savulescu's treatment of Williams's position is surprisingly uncharitable. So, for example, on p. 221 Savulescu says:

Williams is correct that we favour and esteem these properties [The properties persons are fond of]. But what he neglects is that we have a *reason* to favour and esteem them. The concept of a normative reason is a primitive one. We have a reason to relieve someone's pain. This reason is provided by or related to the badness of pain. It is not provided by the fact that we disapprove of that person being in pain (221).

These are misleading remarks given that Savulescu knows full well that Williams does not *neglect* this possibility – he *attacks* it as incoherent in his defense of internal reasons. Savulescu simply *asserts* the truth of the external reasons thesis here with his talk of the supposedly 'primitive' concept of a 'normative' reason.

When Savulescu goes on to consider Williams's own position on these matters, Williams gets unfairly equated with Hume, and the sophistication of Williams's own view is ignored. As many commentators on this debate (both pro and con) have pointed out, Williams's position differs from Hume's in several important ways.

First, the persons' 'S' ('subjective motivational set') includes not just currently held desires but "dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent" (Williams 1979: 105). Also, this set isn't static – ends can come to be modified as means are evaluated. In addition, Williams explicitly allows for non-instrumental reasoning and a significant role for the imagination. As Christopher Cowley summarizes: "Williams is keen to avoid the narrow Humean skepticism about normative reason which would come about if [A has reason to phi] were true only when the agent were *already* motivated to phi" (Cowley 2005: 350).

This subtlety does not survive Savulescu's re-description of Williams's position, however, so we get a crude description of internalism that makes Williams's view sound more than a little bizarre:

Thus, if we happened not to care about human beings, or persons, we would have no reasons, on the Williams-style account, to care about them. If parents did not care about their children, they would have no reason to care about them. And if we did care about persons, we would have a good reason to care about them. If we accept this kind of defense of the human prejudice, anything goes, or at least, anything could

go depending on what we happened to care about. There is no reason to care about anything! God, your mother, your love, your children, or yourself (225).

As we've seen, however, what is left out is the way in which Williams's account goes well beyond Hume's in allowing a rich role for deliberation to adjust our desires. Thus, if 'we happened not to care' about a particular thing (persons, God, mom) it is perfectly compatible with Williams's account that there are all sorts of ways in which we could reasonably *come to care* about those things.⁹ And when that is possible, then a reason to care can indeed be present. What Williams denies is that we can make sense out of talk of practical reasons that are not in this sense connected to motivation.

In addition to distorting Williams's position, Savulescu is also silent regarding the many difficulties attached to his preferred rival (externalist) view. Perhaps more importantly, however, Savulescu has in a way changed the subject and distracted us from the issues with which he began. This is because *even if* Savulescu offered both a convincing account of the nature of external reasons and an explanation of their motivational power, this still won't settle the issue at hand, since it seems entirely coherent that *if* external reasons actually exist, *there could be external reasons for humans to care about humanity*. It is hard to see how a demonstration of the existence of an external reason for a parent to care about their child (say) could be offered which isn't similar in structure to something that could be said on behalf of an external reason for humanism.

Tellingly, Savulescu does not attempt to offer specific arguments for the existence of external reasons for any of the capacities he takes to be of importance to personists. Instead the claim he makes, at various points, is that what most of us *really* care about (despite appearances) is not whether a creature is a human being but "the characteristics that make [individuals] persons" (228). (In other words, rather than argue for an external reason, he spends his time arguing that we are already motivated by personist considerations.) Interestingly, he thinks he can show this by an appeal to *how we act*. So, at one point he claims: "This is revealed by our practices of letting brain damaged human beings die and even killing them: those in a persistent vegetative state, sufferers from severe brain injury or advanced dementia. These are human beings but we do not value their lives when we disconnect their feeding tubes" (228). Putting aside the fact that it is far from obvious that removing feeding tubes is

⁹ A more detailed but similar defense of Williams's notion of internal reasons can be found in Patrone 2013. Patrone does not consider Williams's discussion of (or Savulescu's response to) the alien-invasion thought experiment, his analogy with ethnic or cultural grouping, or Williams's claim that there is something not quite *sane* about the I/O model, but she offers an insightful discussion which I think is generally compatible with the diagnoses and analysis offered here.

an clear demonstration that (quote) ‘we do not value their lives’, the larger issue here is that Savulescu apparently thinks it legitimate to *appeal to our practices* to show what we ‘really’ value, yet he chooses to ignore the large variety of ways in which our practices demonstrate, every day, that we value humanity in a way that cannot simply be reduced to the value we also happen to place on sentience or rationality or moral agency. To take just one dramatic example effectively discussed by Cora Diamond (1978): we don’t eat our dead.¹⁰

6. *The Impartial Observer*

I’ve argued that Savulescu has been unfair in his characterization of Williams’s position on internal reasons, but I’ve also claimed that his extended discussion of (what he calls) VB and DB reasons is somewhat beside the point. However, as we’ve seen, in ‘The Human Prejudice’ Williams does criticize an influential model (that of the Impartial Observer) which many have taken to undergird and support the claim that there are external reasons to prioritize a concern with sentience in our moral thought. Savulescu attempts to defend the Impartial Observer model criticized by Williams, and I think Savulescu’s response amounts to another illuminating misunderstanding. This time the misunderstanding is shared by Peter Singer. Both Singer and Savulescu sharply criticize Williams’s remarks about the feasibility of the Impartial Observer model, and both offer the same sort of complaint: They take Williams to have offered a criticism that only applies to *negative* utilitarianism, and they reply that this objection doesn’t affect more sophisticated versions of utilitarianism (like the sort actually embraced by Singer):

This seems to be an attack on negative utilitarianism – the view that we should minimize suffering. However, it does not seem to apply to positive utilitarianism, that we should maximize the balance of pleasure over suffering. After all, if we consider happiness as well as suffering, there is much to be encouraged by. We would only annihilate the planet if the suffering outweighed the happiness and life for all or the majority was not worth living (Savulescu 2010: 226).

¹⁰ See also Elizabeth Anderson’s illuminating discussion (in Anderson 2004) of the ways we value humanity (e.g. we view a child as having a right to learn language, while even the smartest chimp (capable of language) has no such right). Jeff McMahan is more upfront than Savulescu on these matters concerning the revisionary nature of moral individualism: “It is true that it is difficult to justify these differences in our eating and funerary practices in terms that a moral individualist would find acceptable. The question is whether the practices challenge moral individualism or whether moral individualism challenges the practices” (McMahan 2005: 373).

But whoever said that the ideal observer would have to have the one fixed idea of being against suffering? Williams's ideal observer would seem designed for use by a negative utilitarian – that is, one who holds that we should minimize suffering, but not that we should maximize happiness. Since Williams mentions 'Utilitarians such as Singer' in connection with this model, he appears to assume that it is the kind of model I would use. But I have never been a negative utilitarian. [...] In *Practical Ethics* and at all other times when I set out the foundations of my ethical position, I write of doing 'what has the best consequences, on balance, for all affected' or 'furthers the interests of those affected'. That obviously includes the pleasures and satisfactions that those affected can experience, as well as the pain and unsatisfied desires. One would have to be an extreme pessimist to assume that considering this would lead an ideal observer to desire to annihilate the planet (Singer 2009: 98-99).

Even the harshest critics of Williams usually acknowledge that he was far from a sloppy or naïve thinker, and he was surely aware that Peter Singer is not a crude negative utilitarian. Why, then, did Williams speak of a simplified Impartial Observer? I take it that Williams presumably chose the simplest version of utilitarianism possible not because he was *dodging* more nuanced versions, but because he thought his point applied to *all* possible versions. As he says:

The model comes in various versions, in many of which the figure is not exactly dispassionate: rather, he is benevolent. This can mean several things, in terms of there being a positive value to preference-satisfaction, and so on, but let us concentrate on the simplest application of the idea – that the Ideal Observer (IO) is against suffering and wants there to be as little of it as possible. With his omniscience and impartiality he, so to speak, *takes on* all suffering, however exactly we are to conceive of that, and takes it all on equally. [...] So I want to take the model seriously: perhaps more seriously, from a certain point of view, than those who use it. [...] I wonder whether they ever consider what it would really be like to take on what the IO supposedly takes on (146).

How, then, could this objection apply to the preference-based utilitarianism of Singer or the positive utilitarianism Savulescu mentions? As he suggests above, Williams believes he can show that none of the philosophers proposing various permutations of the IO are actually taking the idea *seriously*.¹¹ To be frank, I think what Williams is suggesting is that there's been some bullshitting going on here. Harry Frankfurt's discussion (in his classic essay

¹¹ Compare his similar discussion in Williams 1990: "It [the World Agent interpretation of the Ideal Observer theory] is appropriate, though, only if the model is taken literally; and if it is taken literally, even to a slight degree, it becomes clear how bizarre it is. Any one agent who had projects as conflicting, competitive, and diversely based as the World Agent's would be (to put it mildly) in bad shape" (87-88). [Note, as he points out on 84, if you pick a different dispassionate model it isn't clear utilitarianism wins the competition...].

'On Bullshit') of a retort from Wittgenstein to Fania Pascal is relevant. Frankfurt characterizes Wittgenstein as the ultimate anti-bullshitter, and talks of an incident where Pascal said to Wittgenstein that she felt 'just like a dog that has been run over'. Wittgenstein apparently replied, disgusted: '*You don't know what a dog that has been run over feels like*' (Frankfurt 2005). I take Williams to, in essence, be saying something like: '*You don't know what being an Impartial Observer feels like!*'

What would it be like to *actually take on* the sufferings of the world? If you think it through, it is rather difficult to conclude that the Impartial Observer would be anything like the figure utilitarians tend to sketch: some sort of Super Agent calmly weighing the scales and calculating how to increase the ratio of 'hedons' or 'utils' in the universe. Instead, this figure, if it in any way resembles *us*, would be tormented beyond words. Tormented, I imagine, to a point where being offered the additional ability to soak up the pleasures on the universe would be of little consolation. (This is why I think this is not merely an attack on negative utilitarianism).

Compare Williams's scenario with Dostoevsky's presentation in *The Brothers Karamazov* of the scene in which Ivan offers Alyosha a now quite famous "thought experiment" involving the creation of a world:

Imagine that you are creating a fabric of human destiny with the object of making men happy in the end, giving them peace and rest at last, but that it was essential and inevitable to torture to death only one tiny creature -- that baby beating its breast with its fist, for instance -- and to found that edifice on its unavenged tears, would you consent to be the architect on those conditions? Tell me, and tell the truth.

"No, I wouldn't consent", said Alyosha softly (Dostoevsky 2002).

The refusal to create such a world seems to Alyosha (and, I'm assuming, to most of us) to be preferable to bringing about the suffering of an innocent child, preferable *despite the joy* that would also be present in that world. Similarly, Williams's thought might have been: any IO that in any relevant sense resembles an actual agent is one who would conclude that the *destruction* of the world is preferable to the continuation of a world with so much suffering *regardless* of how much joy might *also* be present in said world.

Now I don't want to deny important differences between a case where one is choosing whether or not to *create* a world versus the choice to *destroy* a pre-existing world. I merely want to suggest that in both scenarios seriously taking on the suffering involved may just be too much for an actual, sane, decent person to bear – too much *even if* that person is also aware of the pleasure and happiness that is possible in such a world. Both scenarios (Dostoevsky's and Williams's) gain their power from forcing us to take the issues at hand *personally*.

I don't know if Williams would approve of the comparison to *Karamozov* here, but I think this interpretation connects up nicely with some of the other things he goes on to say, like the line I quoted earlier: "It is not an accident or a limitation or a prejudice that we cannot care equally about all the suffering in the world: it is a condition of our existence and our sanity" (Williams 2008: 147). If you think you are successfully imagining an impartial observer that can adequately take on all that suffering, you should ask yourself whether the figure you are imagining could possibly be sane. If you come to agree with Williams that sanity would be hard to maintain in such a situation, then you might want to reconsider the degree of authority that such a figure should possess over judgments concerning how we should best live our lives.

Savulescu also considers Williams's remark about the impossibility of caring equally about all suffering and says the following:

Perhaps the claim here is, that without giving greater concern to *human* suffering, we could not exist. This again is not true – some animal rights activists seem to be able to exist caring equally about all suffering. Peter Singer exists, even though he attempts in perhaps an imperfect human way to care about all suffering (226).

Savulescu does not seem to recognize that Williams is here talking of someone *actually being able to care* about all suffering equally, not just someone who in some distant intellectual sense recognizes a supposed duty to care equally. In other words, Savulescu is not taking Williams very seriously here if he thinks Peter Singer (the actual person) comes in any way close to embodying the sort of hypothetical example Williams has been considering.

Why the disconnect? I think Savulescu (here, and in his earlier dismissal of the 'mere' contingency of the structural difference) is failing to appreciate the way in which Williams insists that our philosophical reflection have a foothold in *actual human life*. On reflection, this IO model is revealed to be an other-worldly fantasy, and as such it distorts our thinking about the (complex and often messy) moral reality of this world.

7. *Species membership as 'arbitrary'*

A repeated mantra in Savulescu's essay [offered at least six times] is that caring about species membership is obviously irrational and confused since species membership is arbitrary, while features like consciousness and rationality are not.

Being human is merely having the property of being able to interbreed or having a certain chromosomal structure. These facts are not in themselves of normative

significance. It is these normative facts that differentiate personism from humanism, racism, and sexism. The latter do not appeal to relevant moral facts. Personism is not an arbitrary club (221).

Species distinctions are arbitrary and turn on capacity to interbreed or some genetic structure. These are not what people have in mind when they say human beings have value (227).¹²

My response in defense of Williams is one that is familiar from the writings of Raimond Gaita and Cora Diamond: The relevant notion of species here is the *folk* concept of a kind of which we are members (a reasonably ‘thick’ concept that contains normative and evaluative components), not the scientific classification of interest to biological taxonomists. The concept that is relevant is one that existed prior to a scientific understanding of the nature of species and does not hinge on the details of any specific genetic account.

This point (about the relevance of our folk concept vs. a scientific or scientific notion of human being) has been made repeatedly by philosophers inspired by Wittgenstein (like Chappell, Diamond, Mulhall, and Gaita) but it also gets repeatedly ignored or misunderstood by criterialists/personists. Two nice examples of explicit appeals to folk concepts are below.

At least for creatures like us, whose temporal experience runs only over decades, hardly ever over centuries, let alone the millions of years that speciation requires, it should not be a controversial thought that a population of creatures sharing a common genetic heritage, physiology, and ethology can sensibly be treated as a unitary grouping, as in fact common sense treats it. And that thought is all I am committed to meaning by ‘species’ (Chappell 2011: 21).

The essence of the kind *homo sapiens* is a genetic code. Anything that has that code is a human being, no matter what it looks like. Anything which doesn’t have that code is not a human being no matter what it looks like. But the conception of humankind that is built out of our responses – and which is interdependent with concepts as basic as sensation – takes little notice of the scientific criterion for *homo sapiens* (Gaita 2002: 269).

Once it becomes clear that the relevant concept here is not simply the idea of a creature with a specific genetic code, Savulescu’s attempt to discredit speciesism as arbitrary simply misses the point. For example, he proposes a thought experiment involving ‘covert alien coexistence’ in which we are to suppose that we discover half of the human population is silicon based and not carbon based. He thinks that our response (i.e., we probably wouldn’t be very bothered) shows that species membership can’t possibly matter to us, but of

¹² See also similar remarks on pp. 223, 236, 239 and 242.

course all this shows is that the *scientific* notion of species doesn't carry moral import.¹³ It doesn't show the concept of humanity is irrelevant, because (pretty clearly) both 'species' here would be taken to fall under that concept.¹⁴

8. *The Alien Example*

We come now, finally, to Williams's discussion of hypothetical cases involving an encounter with aliens. The first thing to note is that although Savulescu quotes the relevant passage, he does not really focus on the case that Williams suggests is difficult: where we imagine that "painlessly, they [the aliens] will rid us, certainly of our prejudices, and, to the required extent, of some of our cultural and other peculiarities" (Savulescu 2010: 237). (Savulescu seems to think such brainwashing is not offensive or morally problematic).¹⁵ Instead Savulescu focuses on the case where the aliens want to *destroy* us and a variety of other cases (which he parallels with supposedly similar cases involve the deaf) that concern saving one life over another or choosing one embryo over another.

Considering the case in which a post-human / alien tries to kill a human for the sake of a better world (and the 'parallel' case in which someone tries to kill a deaf person for the sake of a better world) Savulescu concludes that resistance is justified. But note the actual justification he offers:

Even on a VB view of reasons, self-interest and morality can be independent sources of reasons for action. Self-interest, like Autonomy, can give weight to what we do care about and value and the fact that we want our lives to go a certain way. [...] People can be morally required to give up their life – for example for their country. But they might have the most reason to desert their moral duty, depending on the values. Self-interest might rationally justify their desertion (241).

¹³ Note that Savulescu presents this example as if it is original, but it is actually very similar to a thought experiment presented by Peter Carruthers back in 1992 and discussed in Chappell 2011. Chappell's answer differs somewhat from mine. I suggest we'd still consider the silicon humans as members of the same kind (humanity). He thinks it is better to claim that the lesson is that more than one species would be accepted as moral persons. I don't think much of substance hinges on our different descriptions here. Note also that a similar scenario is discussed (and a similar diagnosis offered) in MacLean 2015.

¹⁴ This sort of response to Savulescu also appears in note 12 of the SEP entry for 'Cognitive Disability and Moral Status' (Wasserman, 2012). Tatiana Patrone also suggests that Williams is not working with a simple biological concept in Patrone 2013. I have previously recommended understanding the relevant concept here as a folk concept in Grau 2010.

¹⁵ Oddly, he interprets Williams's example as involving only the removal of a speciesist prejudice (Savulescu 2010: 237), when it is clear Williams has in mind a scenario where any aspects of our form of life which lead to excessive "cultural autonomy" would be removed (Williams 2008: 149).

This is a very odd response to Williams's discussion. Williams is considering whether the universal moralist can possibly justify (from *within* their theory) resistance to the aliens. Savulescu responds that any of us can always justify resistance in terms of self-interest. In other words, his response is to say we could give the intuitively appealing answer (that resistance is justified) by *abandoning* moral considerations. Savulescu has here changed the subject. He has not actually addressed the challenge: can the universal moralist justify resistance without resorting to humanism?

To the extent that Savulescu does actually consider the challenge from Williams, he seems, in the end, to admit that collaboration may be morally required. He waffles on what his own actual position is but at one point *seems* to embrace scalar personism: "my own position based on personism and moral status being a scalar and not a threshold concept" (241). Given that approach, he admits that on "value-based personism, smarter people DO have more of a right to live" (238). And while at various points he seems to follow Jeff McMahan in granting that a *small* amount of partiality towards humans *may* be justifiable in terms of 'special relations', when discussing the alien scenario Savulescu concludes: "We might have reason to save or create such vastly superior lives, rather than continue the human line" (244).¹⁶

It is fair to say I have significant philosophical disagreements with Peter Singer, but on this topic I admire the fact that he is at least refreshingly candid and straightforward. He cuts to the chase and gives just the answer to Williams's challenge we expect from him:

In these circumstances, the right thing to do, and the courageous thing to do, is not to listen to the tribal instincts that prompt us to say 'My tribe (country, race, ethnic group, religion, species, etc) right or wrong' but to say: 'I'm on the side that does what is right'. Although it is fantastic to imagine that a fair-minded, well-informed, far-sighted judge could ever decide that there was no alternative to the 'removal' of our species in order to avoid much greater injustice and misery, if this really were the case, we should reject the tribal – or species – instinct, and answer Williams's question in the same way (Singer 2009: 101).

This brings us to a very fundamental issue (on which I'll end): what do we do when one person's *reductio* is another person's courageous biting of the bullet? At one point in his essay Williams speaks of Singer and makes the quip: "he and the pro-lifers both argue 'if abortion, then infanticide' but they take it as an objection, and he takes it as an encouragement" (144). Or consider Williams's remarks, in this same essay, about our duties to nature:

¹⁶ I discuss McMahan's views on this issue in Grau (2015).

The question arises, whether we should not be in the business of reducing the harm that other animals cause one another, and generally the suffering that goes on in nature.... There is something altogether crazy about the idea, that it misrepresents our relations to nature (146).

And compare that with Jeff McMahan's discussion in 'The Stone' column in the *New York Times*:

If we could bring about the end of predation by one or the other of these means at little cost to ourselves, ought we to do it? [...] I am therefore inclined to embrace the heretical conclusion that we have reason to desire the extinction of all carnivorous species, and I await the usual fate of heretics when this article is opened to comment (McMahan 2010).

What we have here is a clash between two fundamentally different sensibilities: what you might call Moral Rationalists/Revisionists (Singer, McMahan, Savulescu) vs. others (like Williams, Diamond, Mulhall, Gaita, etc.) who are less inclined to take the deliverances of moral theory as authoritative and are thereby less revisionary in their tendencies. I am somewhat skeptical that the tensions between these camps will ever be finally resolved, but of course in *giving that diagnosis* I align myself with the side in the debate that acknowledges significant limits on how far reason can take us towards convergence in ethical (or indeed philosophical) thought. (That is, the group I've dubbed 'the rationalists' are going to be unlikely to accept the claim that fundamentally different sensibilities lay at the bottom the dispute. They are going to be more likely to conclude that their opponents have made a mistake or are simply dumb.) Skeptical as I may be of rationalistic pretensions in moral theory, I still think it is worth talking about these issues and *trying* to make some progress, and that's what I have attempted to do here in trying to defend Bernard Williams's essay from what I take to be misguided but revealing criticisms.

Acknowledgments: An earlier version of this paper was presented at the CUNY Graduate Center Philosophy Colloquium and the Third Colloquium on the Modalities of the Good, Charles University, Prague. Thanks also to Elisa Aaltola, Sophie-Grace Chappell, John Hadley, Tatiana Patrone, Nicholas Smyth, David Wasserman, Federico Zuolo, and two anonymous reviewers for comments.

References

- Anderson, Elizabeth, 2004, "Animal Rights and the Values of Nonhuman Life", in Sunstein, Cass R. and Nussbaum, Martha C. (eds.) *Animal Rights: Current Debates and New Directions*, Oxford University Press, Oxford: 277-298.
- Chappell, Timothy, 2011, "On the Very Idea of Criteria for Personhood", in *The Southern Journal of Philosophy*, 49, 1: 1-27.
- Cowley, Christopher, 2005, "A New Defence of Williams's Reasons-Internalism", in *Philosophical Investigations*, 28, 4: 346-368.
- Diamond, Cora, 1991, "The Importance of Being Human", in *Royal Institute of Philosophy Supplement*, 29: 35-62.
- Diamond, Cora, 1978, "Eating Meat and Eating People", in *Philosophy*, 53, 206: 465-479.
- Dostoevsky, Fyodor, 2002, Eng. tr. by Richard Pevear *The Brothers Karamazov*, 12th edition, Farrar, Straus and Giroux, New York.
- Finlay, Stephen and Schroeder, Mark, translator; Pevear, Richard, 2012, "Reasons for Action: Internal vs. External" in the *Stanford Encyclopedia of Philosophy*, url: <http://plato.stanford.edu/entries/reasons-internal-external/>
- Frankfurt, Harry, 2005, *On Bullshit*, Princeton University Press, Princeton.
- Gaita, Raimond, 2002, *A Common Humanity: Thinking about Love and Truth and Justice*, Routledge, New York.
- Grau, Christopher, 2015, "McMahan on Speciesism and Deprivation", in *The Southern Journal of Philosophy*, 53, 2: 216-226.
- Grau, Christopher, 2010, "Moral Status, Speciesism, and Liao's Genetic Account", in *Journal of Moral Philosophy*, 7: 387-396.
- MacLean, Douglas, 2015, "A Plea for Humanism," 6th Russell Philosophy Conference, Sonoma, Ca, March 16, 2015.
- McMahan, Jeff, 2010, "The Meat Eaters", in *The New York Times, Opinionator: The Stone*, accessed: Sept. 19, 2010, url: <http://opinionator.blogs.nytimes.com/2010/09/19/the-meat-eaters/>
- McMahan, Jeff, 2005, "Our Fellow Creatures", in *The Journal of Ethics*, 9, 3-4: 353-380.
- Patrone, Tatiana, 2013, "In Defense of the 'Human Prejudice'", in *International Journal of Technoethics*, 4, 1: 26-38.
- Savulescu, Julian, 2010, "The Human Prejudice and the Moral Status of Enhanced Beings: What Do We Owe the Gods?" in Julian Savulescu and Nick Bostrom, eds., *Human Enhancement*, Oxford University Press, Oxford.
- Singer, Peter, 2009, "Reply by Peter Singer" in Schaler, Jeffrey, ed., *Peter Singer Under Fire: The Moral Iconoclast Faces His Critics*, Open Court, New York: 97-101.
- Wasserman, David, Jeffrey Blustein, Daniel Putnam, accessed: July 6, 2012, "Cognitive Disability and Moral Status", in *Stanford Encyclopedia of Philosophy*. url: <http://plato.stanford.edu/entries/cognitive-disability/>
- Williams, Bernard, 2008, "The Human Prejudice" in Williams, Bernard, *Philosophy as*

- a Humanistic Discipline*, Princeton University Press, Princeton: 135-152.
- Williams, Bernard, 1995a, *Making Sense of Humanity*, Cambridge University Press, Cambridge.
- Williams, Bernard, 1995b, "Replies" in Altham, J.E.J. and Harrison, Ross, eds., *World, Mind, and Ethics: Essays on the ethical philosophy of Bernard Williams*, Cambridge University Press, Cambridge.
- Williams, Bernard, 1990, *Ethics and The Limits of Philosophy*, Routledge, London.
- Williams, Bernard, 1979, "Internal and External Reasons", in Harrison, Ross, ed., *Rational Action*, Cambridge University Press, Cambridge: 101-113.

