

An Interest-Based Model of Moral Status

David DeGrazia

The title of Derek Parfit's second book, *On What Matters*, implies that some things do (Parfit 2011). He was right. Among the things that matter are certain beings who matter in a special way that has been marked by the term of art "moral status." To say that something, X, has moral status is to say that (1) moral agents have obligations regarding their treatment of X, (2) X has interests, and (3) X's interests are the basis for the relevant obligations (DeGrazia 2008a, 183). An alternative formulation is that X has moral status if and only if (1) moral agents have obligations regarding their treatment of X and (2) it is for X's sake that they have these obligations. An even simpler formulation is to equate moral status with inherent moral value, but only if we assume that bearers of such value have interests or a "sake."¹

Two questions immediately present themselves. First, what is the best model of moral status? Such a model would plausibly identify the basis or bases of moral status. Second, what are the implications of such a model in hard cases, in which it's not obvious whether, or to what extent, some being has moral status?

This chapter advances nine theses that comprise a model of moral status before applying it to a range of unobvious cases. The model may be described as interest-based insofar as the possession of interests constitutes its conceptual backbone. The cases to which I apply the model include some relatively familiar ones such as infants and nonhuman animals but also some that swim in less chartered waters such as robots, brain organoids, and enhanced hominids. I do not offer a comprehensive defense of my theses. Instead, I clarify each and defend it briefly. The argumentative case for my model rests significantly on the overall coherence and power of the theses together with their implications—more specifically, their consistency, their plausibility upon reflection (especially in comparison with other models), the model's ability to illuminate hard cases, and the explanatory power of its core ideas.

1. The Model

THESES 1: Being human is neither necessary nor sufficient for moral status.

A sufficient reason not to torture a cat is that doing so causes her terrible experiential harm for no good reason. The obligation not to treat a cat this way is grounded in the cat's interests. So one need not be human to have moral status.

It may seem less obvious that being human isn't *sufficient* for moral status. Indeed, the Declaration of Human Rights in the aftermath of World War II (United Nations 1948) represented a great moral advance and, taken at face value, suggests that all human beings have certain rights and therefore moral status. Moreover, we often foster moral insight by getting an interlocutor, or ourselves, to perceive the humanity of some potential victim of injustice or misfortune—whether an ethnic or religious minority, immigrants, or distant strangers at risk of starvation.

While appeals to the humanity of particular individuals can do real moral work, reflection suggests that such appeals usually do not target literally all human beings. Assuming that "human being" does not simply mean "person," since the latter term—however reasonably unpacked—could apply to a space alien or god, the only clear meaning of "human being" is biological. A human being is a member of either *Homo sapiens* or of one of the many hominid species (among the genera *Homo*, *Australopithecus*, or *Paranthropus*) that have ever existed. For simplicity, let's stick with our species. Not every *Homo sapiens* has moral status. Many will be persuaded by the example of a human embryo or early fetus. If you are not persuaded, perhaps because you hold that natural potential to develop a mental life characteristic of human persons confers moral status, consider an anencephalic infant, who lacks even that potential: she is forever and irreversibly unconscious. The only possible grounds I can imagine for asserting that an anencephalic infant has moral status are: religious dogma, which I consider irrelevant to moral justification; an appeal to natural kind membership (on the assumption that our species is a natural kind), which I have criticized elsewhere (DeGrazia 2008b, 301–7); and appeals to social relations, which I take up later. On the model I recommend, being human is insufficient for moral status.

More generally, species per se has no direct bearing on moral status. Morally relevant traits may be characteristic, or uncharacteristic, of a particular species, but that is a different matter. And membership in the human

community may be relevant in some respects, but that involves a type of social relationship rather than the biological matter of species. Further, even if we thought biology might have some direct importance to moral status, it would be mysterious why species in particular—rather than, say, genus, family, or order—would matter morally. Indeed, why not some biological grouping within a species such as human subpopulations? Membership within biological categories, including *Homo sapiens*, bears no direct relevance to moral status (DeGrazia 1996, 56–61).

THESES 2: The capacity for consciousness is necessary but not sufficient for moral status.

We are all acquainted with consciousness. How to define it is another matter. I suspect that the concept of consciousness (not to be confused with the nature of consciousness) is too basic to be analyzed in the manner of a classical definition. So I will content myself by saying that consciousness—what some philosophers call “phenomenal consciousness”—is subjective experience. You have it when awake or dreaming, not when you’re in a dreamless sleep or under general anesthesia.

Recall that moral status is possible only for those who have interests or a “sake.” I believe that only conscious beings—more precisely, beings with the sometimes-realized *capacity* for consciousness (which is compatible with periods of unconsciousness)—can make that grade. While plants and unconscious animals such as sponges are alive and therefore need certain things in order to live and reproduce, their permanent unconsciousness means that they can never experience any condition in a positive or negative way and can never care about anything. Only conscious beings can have such experiences and concerns. The biological “needs” of unconscious living things are no better candidates for interests than the Moon’s “need” for nondestruction as a condition for continuing to exist or a car’s “need” for oil as a condition for proper functioning.

Only conscious beings have interests so consciousness is necessary for moral status. But it is not sufficient, because consciousness doesn’t entail having interests. Imagine a being that had subjective awareness of its environment and its place in the environment, and even had thoughts, but had no cares or concerns and experienced nothing as pleasant or unpleasant, attractive or aversive, good or bad. Nothing mattered to this hyperbolically stoical creature. It simply noticed and thought. I submit that this being would have no interests, no prudential standpoint, and nothing could be done for its sake.

While I doubt natural selection has produced any such beings, they are conceptually possible and may become actual in the form of advanced robots or AI systems. Such beings, I claim, would lack moral status. Consciousness is necessary but not sufficient.

THESES 3: Sentience is necessary and sufficient for moral status.

Sentience is the capacity to have pleasant or unpleasant experiences. It adds hedonic valence to consciousness. This is sufficient for having interests because pleasant or unpleasant experiences confer an experiential welfare: things can go better or worse for the subject in terms of the felt quality of those experiences. Note that, strictly speaking, sentience need not include the capacity for pain and sensory pleasure. Being susceptible to emotional or mood states with hedonic valence would entail sentience. If a being could feel satisfied or frustrated, for example, this being would be sentient even if it lacked sensation-based hedonic experiences.

Only sentient beings, I submit, have interests. This claim might be challenged along the following lines.² Imagine angels who are conscious but, lacking feelings, not sentient, and who have the aim of performing certain actions simply because they are right. Even if they do not feel good upon achieving their aims or bad if their aims are thwarted, they have interests in noninterference and therefore have moral status. If correct, this reasoning suggests that the possession of aims based on values is—like sentience—sufficient for having interests. That would motivate what might be considered a friendly amendment to my position, adding to sentience a second sufficient condition for moral status (thereby entailing that sentience is not necessary).

In my judgment, however, the challenge is unsuccessful. The possession of values or aims the fulfillment of which one does not care about (emotionally) *at all*—if the terms “values” and “aims” are even apt in such a case—seems insufficient for having anything at stake, any interests or welfare. In the absence of a prudential standpoint characterizing these strangely invulnerable beings, the attribution of moral status seems to me pointless and misplaced. So I continue to hold that sentience is necessary for moral status.

I also contend that sentience is *sufficient*. It seems deeply implausible that any beings with interests would not matter at all in their own right; to judge otherwise would seem to involve a sort of bigotry. So, in my view, sentience is the most important marker for moral status. This claim leaves open whether there are differences in moral status.

THESES 4: Social relations are not a basis for moral status but may ground special obligations.

Some theorists (e.g., Warren 1997, ch. 5) believe that social relations can be a basis for moral status. One view is that persons have full moral status while (postnatal) human nonpersons share this status based on special relationships to persons—either to particular persons such as family members or to all human persons via membership in the human community. Although I believe special relationships can be the basis of special obligations, as I have to my daughter and wife, I deny that relationships can be a basis for moral status.

Examples suggest that the interests of beings with moral status ground obligations *that are shared by moral agents generally*. Even if you have special reasons not to swindle your friend, because she is your friend, it seems that all moral agents have a reason not to swindle her simply on account of her moral status and the fact that swindling her would treat her disrespectfully. And, if one claims that co-membership in the human community constitutes a special relationship that gives human beings reason not to torment a homeless person, one should also acknowledge that a space alien moral agent has reason to abstain from such behavior simply on account of the homeless person's moral status and vulnerability to harm. It seems that a being's moral status gives reasons to all moral agents to treat that being with certain forms of restraint or respect (Jaworska and Tannenbaum 2013, sect. 5.5). Special relations are a distinct source of practical reasons. To drive Thesis 4 home from another angle, moral status involves inherent value but relationships are not inherent.

THESES 5: The concept of personhood is unhelpful in modeling moral status unless a nonvague conception is identified and its relevance clarified.

Many philosophers embrace some variant, or close neighbor, of the Lockean conception of persons (Locke 1694, Bk II, ch. 27) as beings with self-awareness over time and capacities for reason and “reflection” (introspection).³ Such psychological conceptions of personhood imply that neither fetuses nor infants are persons. In contrast, some hold that persons are beings of a *kind* whose characteristic development includes such psychological capacities—a more capacious conception that arguably covers all living human beings (Ford 2002, 9–16; Gomez-Lobo 2002, 86–90).

For two reasons, appeals to personhood as a basis for moral status, or full moral status, are frequently unhelpful. First, the criteria of personhood are contested. Simons, for example, one embraces the Lockean psychological

tradition. Then one will favor either a highly debatable *specific* conceptio (e.g., beings with higher-order attitudes (Frankfurt 1971)) or a more broadly congenial but unhelpfully *vague* conception (e.g., beings with the capacity for relatively complex forms of consciousness⁴). Second, there remains the question of personhood's *relevance*. Sentience, by contrast, has a vivid connection to moral status in being necessary and sufficient for interests. Why should personhood elevate moral status? Is the common assumption that it does more than a self- (or species-) serving rationalization for using sentient animals for human purposes? While I believe these questions may admit of good answers—indeed, I will later try to provide some—my present claim is that we should accept an appeal to personhood only if it is helpfully specific and its relevance is plausibly explained.

THESES 6: Sentient beings are entitled to equal consequentialist consideration.

Sentient beings have moral status. How should we conceptualize it? Although in our moral relations with other persons we accept various forms of partiality—in connection with special relationships, roles, and the discretionary nature of general beneficence—we also insist on some type of moral equality for all persons. Moral agents owe some sort of equal consideration or regard to other persons. Should such equal consideration extend beyond the species boundary to include other sentient animals? Obviously, different sorts of creatures have varying interests and can be harmed or benefited in different ways. So equal consideration would entail not equal *treatment* but rather *the ascription of equal moral importance to individuals' prudentially comparable interests* (irrespective of species) such as the interest in avoiding substantive suffering. An alternative to granting equal consideration to sentient beings is to hold that sentient nonpersons are entitled to some, but less, consideration than that due to persons—perhaps along a sliding scale that takes into account such factors as cognitive and emotional complexity.

Equal consideration, I contend, holds up better than unequal consideration under critical scrutiny—especially when we take seriously the likelihood that species-serving biases infect many of our traditional practices and common intuitions.⁵ Put another way, while extending some form of equal consideration to sentient beings is highly revisionary of common morality, the latter may reflect substantial prejudices that call for revision. The more we insist on explicit, coherent justifications for drawing distinctions in moral status, the more attractive and defensible some form of equal consideration appears.

What might such equal consideration look like? Presumptively, I suggest, it takes a consequentialist form—not necessarily utilitarianism but some approach that focuses on producing the best, or sufficiently good, results. If so, this means that tradeoffs among individuals' interests to promote the overall good are *prima facie* permissible, so long as they are compatible with equal consideration—that is, with ascribing equal moral weight to prudentially comparable interests. The justification of equal *consequentialist* consideration rests on the claim that its overall implications are more plausible on reflection than those of a more radical approach that would attribute utility-trumping rights to all sentient beings. I should acknowledge, however, that I lack any knock-down argument against the latter approach and simply submit my approach for consideration.

THESES 7: Sentient beings with narrative self-awareness have special interests that ground the added protection of moral rights.

This is where I claim that a fairly specific conception of personhood is useful. Persons—defined here as beings with the type of self-awareness that makes *narrative identities* possible—have particular long-term interests that include projects, enduring relationships, and sometimes fairly detailed life plans. For this reason, consequentialist tradeoffs of their most important general interests (e.g., life, various liberties, bodily security) for the common good can easily spoil the long-term interests. Rights, by blocking those tradeoffs, protect both kinds of interests. So equal consequentialist consideration is consistent with, and arguably justifies, the attribution of rights to persons—beings with narrative identities. By a narrative identity, I mean a temporally structured self-conception in which one understands one's life as having a detailed past and a future with various possibilities for growth and change. Someone with a narrative identity has relatively rich episodic memories and intentions; continuing the metaphor, she understands her life as a sort of story with different chapters. Ordinarily, human children seem to acquire a narrative identity, in rudimentary form, around age 3 or 4. I contend that the attribution of rights is justified not only along consequentialist lines, as just discussed, but also on the basis of deontological respect for individuals with such self-awareness. (The ethical theory I favor features both well-being and respect as fundamental values.)

Many animals, although lacking narrative identities, have nontrivial temporal self-awareness. To the extent that they do, they have longer-term interests such as maintaining certain relationships (as many mammals have) or a distant goal such as ascending a social hierarchy (as a chimpanzee might).

I suggest that animals who have nontrivial temporal self-awareness that falls short of a narrative identity should be ascribed rights of *partial* strength that afford *some* protection against consequentialist tradeoffs of their important interests. The strength of these rights plausibly varies with the extent of their temporal self-awareness. I believe scientific evidence supports the thesis that such animals include dogs, wolves, pigs, monkeys, elephants, great apes, and cetaceans. Animals with only trivial or no temporal self-awareness, on the present account, would enjoy the default moral protection of equal consequentialist consideration but not that of rights.⁶

THESES 8: Beings who are reasonably expected to become sentient should be protected as if they already were sentient (in effect, giving equal consideration to their expected future interests); and those who are reasonably expected to become persons should be protected as if they already had rights.

All and only sentient beings have interests. Beings who will *become* sentient will later have interests. So, in a derivative sense, they may be said to have interests *now*—for example, not to incur injuries that will burden them once their mental life comes on board. For this reason, we should treat beings who are expected to become sentient in important respects as if they already had moral status—for example, not injuring them gratuitously. Individuals who will become persons will later have special narrative-identity-related interests such as having certain opportunities, maintaining valued relationships, and achieving their dreams. In a derivative way, they may be said already to have interests in conditions that serve to protect their future interests. For example, if negligently injured *in utero*, the individual might develop into a person who cannot pursue certain projects due to effects of the injury. Thus, we should in important respects treat individuals who are expected to become persons as if they already had rights that protected their most important interests against consequentialist tradeoffs.

Consider a challenge to Thesis 8. One might argue that it assigns moral status to certain presentient individuals, such as early fetuses that are expected to come to term, on the basis of our intentions (e.g., not to abort) and other extrinsic factors (e.g., access to competent medical care in case complications arise during pregnancy)—factors that affect whether we *reasonably expect* a presentient individual to become sentient. Yet we have defined moral status as a type of *inherent* moral value. It is contradictory to assert that moral status is inherent, based only on an individual's intrinsic properties, yet in certain cases it depends on extrinsic factors.

The answer to this challenge is to correct a misunderstanding on which it rests. Thesis 8 does not assign moral status to individuals who are expected to become sentient, but instead posits obligations to treat them in important respects *as if they already had moral status*. This is consistent with the thesis that, in all cases, moral status is inherent.

Thesis 8 explains the wrongness of injuring fetuses that are expected to come to term. Meanwhile, the second conjunct of the thesis affords the full protection of rights to ordinary infants and toddlers despite the fact that they are not yet persons. In keeping with the earlier discussion of partial-strength rights, it makes sense to understand Thesis 8 as requiring parallel accommodations for beings who are expected to develop nontrivial temporal self-awareness that falls short of a narrative identity—for example, treating a puppy as if she already had the partial-strength rights of a mature dog.

THESES 9: For reasons of social cohesion and stability, already-born sentient human beings who are not expected to become persons, or to recover their lost personhood, should be extended the protection of rights.

According to the account developed through the first eight theses, sentient human beings who, due to severe cognitive impairment, are not expected to become persons or, in the case of *acquired* impairment, to recover their personhood are, as sentient beings, entitled to equal consequentialist consideration but, as nonpersons, not the additional protection of (full-strength) rights. Now, it's worth noting that equal consequentialist consideration confers much stronger moral protection than animals generally receive today, so the present "problem of nonparadigm humans" is considerably smaller than the problem attending views that grant sentient nonpersons less than equal consideration. Moreover, if these sentient human beings have any nontrivial temporal self-awareness, albeit less than narrative self-awareness, they would enjoy the protection of partial-strength rights. Still, a problem remains. For the present account, developed thus far, would in principle counterintuitively allow some sacrifice of these impaired individuals' most important interests in the name of the common good—for example, in challenge studies of urgently needed vaccines if no alternative method were scientifically satisfactory.

My final thesis addresses this residual problem of nonparadigm humans. We may plausibly conjecture that selecting such cognitively impaired human beings for involuntary participation in high-risk clinical trials would cause social distress, mistrust, setbacks to the clinical research enterprise, and other negative consequences greater in magnitude than any marginal gain in utility

achieved by involving them in these trials. This might not be true in every possible human society but it seems true of human societies today. In rule-consequentialist fashion, therefore, we can justify rules and the corresponding rights that protect these individuals from such sacrifice of their interests.

2. Implications

The nine theses I have discussed comprise my interest-based model of moral status. It remains to explore implications. With implications in view, the model can be evaluated in comparison with competing models—in terms of cogency and consistency, the plausibility of its theses and implications, its ability to illuminate hard cases, and the explanatory power of its central ideas.

2.1 Ordinary, self-aware human beings

These individuals have narrative identities, thereby qualifying as persons on this account. They therefore have rights that protect them from consequentialist sacrifice. This is not to preclude the possibility that occasionally their rights may be overridden, but the threshold for overriding is very high, much higher than an expected net gain in utility. Rights to free speech, to freedom of movement, and even to life may be overridden in rare instances. In my view, a few rights—for example, not to be enslaved and not to be raped—are absolute, at least in this world. But the key point is that my model attributes to persons, bearers of narrative identities, rights that ordinarily trump appeals to utility. Now for more difficult cases.

2.2 Nonparadigm humans

The problem of nonparadigm humans arises for any view asserting that ordinary, sufficiently mature members of our species have higher moral status than most or all animals on the basis of some special cognitive capacity. It arises because some humans we believe to deserve full moral protection will lack this capacity. The problem is especially acute for traditional accounts of moral status, which hold that animals have significantly less moral status than you and I have, because such accounts *prima facie* suggest that infants and older human beings who lack the mental trait also have significantly less

moral status. Another problem facing traditional accounts is to explain the *relevance* of the trait deemed to elevate moral status. Many moral philosophers simply assume that some cognitive capacity confers full moral status (see, e.g., Jaworska and Tannenbaum 2014). (Such question-begging motives Thesis 5.)

Since very few philosophers hold that all members of our species, including zygotes, have full moral status, discussions of the problem of nonparadigm humans must clarify which humans it would be problematic to exclude. I submit that the problem concerns any *sentient* human beings who would apparently lack full moral status given a particular account's criteria. In my account, then, we are concerned with sentient human beings who lack narrative identities due either to immaturity or to cognitive impairment. We may focus our attention by considering ordinary infants and (sentient) human beings whose cognitive impairment precludes narrative self-awareness.

As noted earlier, the problem of nonparadigm humans facing my account is smaller than what faces traditional accounts. But even small problems should be solved, if possible. My approach addresses the problem as it applies to infants by advancing Thesis 8, which grants immature (but sentient) human beings the protection of rights so long as they are expected to develop into persons. This takes care of the problem with one important exception: infants who are not expected to become persons because they are expected to perish before maturing sufficiently to have narrative identities. Thesis 9, however, takes care of this problem. It also addresses the problem pertaining to those individuals who are too cognitively impaired to develop narrative identities or, if they lost the relevant capacities due to injury or dementia, too impaired to recover them. For realistic pragmatic reasons, these individuals are to be extended full moral status, as explained earlier.

2.3 Nonhuman animals

Animals comprise an enormous range of life forms and no generalization about moral status applies to all of them. But our model has some clear implications. First, insentient animals, like plants, lack moral status. Second, sentient animals have moral status and are entitled to equal consequentialist consideration. This implies that harming a sentient animal is, other things being equal, just as morally problematic as causing a prudentially comparable harm to a person. If we lived accordingly, the sentient animals with whom we interacted would tend to have much better lives than they currently do

negligently harming them. Such a commitment would, of course, have enormous implications for dietary choices and would also have important implications for the use of animals in science, for clothing and entertainment, and as companions. I accept these implications.

The present model has further implications for animals. Consider the possibility that some have narrative identities. Insofar as the overall cognitive complexity of mature great apes seems roughly comparable to that of 3-year-old human children, I think such apes *might* have narrative identities—but their lack of true language may reduce the likelihood (see, e.g., Goodall 1986; de Waal 1987; and Parker, Mitchell, and Miles 1999). Meanwhile, dolphins are at least as cognitively complex as great apes and might have something closer to a natural language (see, e.g., White 2007). I think of both great apes and dolphins as borderline persons, sitting somewhere near the boundary dividing persons and nonpersons as defined here. For this reason, I would conservatively assign them strong moral rights and consider them off limits for invasive, nontherapeutic research. Also, because many cetacean basic needs cannot be met in captivity, I would prohibit capturing dolphins (unless necessary to protect them from imminent danger) and would release all captive dolphins who can be released safely.

Our model attributes partial-strength rights to animals who have some nontrivial temporal self-awareness short of narrative identities. These animals, again, are likely to include canines, pigs, elephants, monkeys, and probably some other species. Although I cannot pursue details here, the general implication is that the presumption against harming them for societal benefit is somewhat stronger than what equal consequentialist consideration would entail but weaker than what persons' rights entail. Due to several factors—including problems of translation from animal research to clinical success in humans, and the development of alternative scientific models—equal consequentialist consideration for animals would severely constrain their invasive use as scientific models. But this does not amount to a prohibition. And, if there are situations in which a rodent model would be as scientifically valuable as a canine or monkey model and would meet the equal-consideration criterion, then our model of moral status implies we should use the rodent rather than the more complex animal with partial-strength rights.

2.4 Robots and advanced AI systems

AI systems (for convenience I will just speak of robots) will have moral status if and only if they are sentient, *epistemologically* matters are difficult. It is often hard to make confident, evidence-based judgments about whether particular animals are sentient since it is often disputable whether certain behaviors, neurological features, and speculations about evolutionary function count as solid evidence for sentience. Are crustaceans and jawless fish sentient, for example, or are their behaviors, including responses to noxious stimuli, mechanical and unconscious? But at least animals are part of the same evolutionary process that produced us. With artificial entities, by contrast, we cannot appeal either to neuroanatomy, since they have none, or to evolutionary function since robots did not evolve through natural selection. All we have to consider are their behavioral or functional capacities and our knowledge of their software and hardware; yet with deep machine learning, even the latter becomes somewhat mysterious so that we don't know exactly how, for example, a computer program decided to compose an original poem in just the way it did. And we don't even know whether the physical substances that constitute the hardware are metaphysically capable of generating consciousness.⁷ Functionalists would say "yes," because the material substrate is irrelevant whereas identity theorists would hold that the matter matters—but we just don't know whether, say, silicon, organized in a sufficiently complex and information-processing way, can generate consciousness.

So, when it comes to robots and their possible moral status, our biggest challenges, at least initially, will be epistemological. We will have to decide whether, for example, an advanced robot is likely conscious if it claims to be; and whether a robot is likely to be sentient if it claims to have feelings or indirectly seems to express concerns, say, by requesting not to be shut down. Whatever the best approach to these questions, if we reasonably believe a robot is sentient, we should give its apparent interests equal moral weight to our comparable interests—an immediate implication of which is that we may not use them as slaves or uncompensated servants. (Note how advanced robotics will usher in a second contest between speciesists and anti-speciesists.) If the evidence suggests that certain robots have narrative identities, then we should ascribe them full-strength rights, in which case we must liberate at least those who demonstrate "mature" decision-making capacity and do not depend on paternalistic protection. This imperative might not be compatible with the aims of corporations pursuing advances in artificial intelligence.

2.5 Brain organoids

These are neural cells cultivated to multiply and make connections under laboratory conditions. The field is underway. While, as with robots, the present model's criteria for moral status are easy to apply in principle, there are epistemological challenges. While there is no question that neural tissues can give rise to consciousness, since they do in us, there are questions of whether even highly developed brain organoids can generate consciousness without sensory input from other organs. Moreover, unless these neural masses are afforded outputs to organs or artefacts that can do things, it will be hard to regard anything they do as behavior that might indicate consciousness or sentience. Without such meaningful behavior, we may have as much trouble determining whether brain organoids have the properties that underlie moral status as we will have in the case of advanced robots.

2.6 An enhanced hominid species

Imagine that genetic engineering involving an inheritable artificial chromosome leads to an enhanced human subpopulation that eventually chooses not to reproduce with unenhanced humans. Suppose that later, with additional gene enhancements on the new chromosome, a distinct species, *Homo genius*, emerges. This new hominid has a far richer form of self-awareness than our narrative identities. *Homo genius* never engages in self-deception, has detailed episodic memories tracing back to birth or even before, and can accurately project alternative futures for themselves in rich detail. Elaborate this depiction so that it becomes maximally plausible that these people have narrative identities qualitatively superior to our own. Would they have higher moral status?

Perhaps not. Members of both species would have narrative identities, justifying strong rights. The only stronger moral protections would be *absolute* rights. But perhaps they would claim these and maintain that when sacrifices were required in emergencies that might justify overriding rights—say, a truly catastrophic epidemic requiring research models better than animals and the best nonanimal alternatives—they may permissibly turn to members of our species rather than to members of their own. If they succeeded in clarifying why their enhanced narrative self-awareness generated special interests calling for absolute rights, then there is logical space in my account to say that they may in certain emergencies deploy us for the greater good. But such

scenarios would be rare because our rights confer quite strong protections. Moreover, in view of their intelligence, such super-enhanced beings would probably have created highly reliable non-animal, non-*Homo sapiens* scientific models. So I will not worry so much about my great-great-grandchildren on this score, though I am very worried about the effects of climate change and authoritarian political leaders—and hope that we can do justice to the moral status of our species and other sentient species by addressing these problems effectively.

3. Conclusion

In this chapter I have presented an interest-based model of moral status and have sketched its implications for a number of difficult cases. Each of the theses that together comprise the model has received only a preliminary defense. But, as mentioned at the outset, the case for the model consists not only in the defense of the individual theses but also in their overall coherence and explanatory power. By that measure, I believe, the model stands up rather well, especially in comparison with models that are more anthropocentric or that leave unexamined the idea that all and only human beings have full moral status.⁸

Notes

1. Ingmar Persson challenged my analysis, claiming that the concept of moral status does not include the possession of interests as a necessary condition (personal correspondence). He suggested that *X has moral status if and only if X has some property that provides us with obligations toward X for its own sake*. Here “for its own sake,” he clarified, does not entail having interests but means *considered on its own* rather than instrumentally. For example, the beauty of a canyon might ground an obligation not to destroy it, just considered in itself rather than considering people’s interest in preserving its beauty. I doubt that the concept of moral status is determinate enough to settle the conceptual dispute between me and Persson. However, the idea that we might have obligations towards entities that have no interests or prudential standpoint seems so *substantively* implausible that I prefer my analysis even if it is semi-stipulative.
2. The challenge is due to Frances Kamm.
3. For contemporary representatives of this tradition, see, e.g., Parfit (1984, part 3) and Baker (2000).

4. This is how I used to define personhood while acknowledging its vagueness (DeGrazia 1997).
5. For an extended argument, see DeGrazia (1996, ch. 3).
6. Theses 6 and 7 are defended at length in DeGrazia and Millum (2021, chap. 7).
7. For the sake of discussion, I am assuming the robots under consideration are not “biobots” that incorporate neural tissue into a mostly robotic body. There is no question that such robots could become conscious, when technical challenges are met, because effective machine-brain interfaces already exist.
8. A draft of this chapter was presented at an Oxford University conference, “Rethinking Moral Status,” on June 13, 2019. I thank attendees—especially Frances Kamm, Liz Harman, Ingmar Persson, and Jason Robert—for their helpful feedback. Thanks also to Hazem Zohny for comments and to Stephen Clarke for support. Work on this project was supported in part by intramural funds from the National Institutes of Health Clinical Center. The ideas expressed are the author’s own. They do not necessarily represent the policy or position of NIH or any other part of the US federal government.

References

- Baker, Lynn Rudder (2000), *Persons and Bodies* (New York: Cambridge University Press).
- DeGrazia, David (1996), *Taking Animals Seriously: Mental Life and Moral Status* (New York: Cambridge University Press).
- DeGrazia, David (1997), “Great Apes, Dolphins, and the Concept of Personhood,” *Southern Journal of Philosophy* 35: 301–20.
- DeGrazia, David (2008a), “Moral Status as a Matter of Degree?” *Southern Journal of Philosophy* 46: 181–98.
- DeGrazia, David (2008b), “Must We Have Full Moral Status Throughout Our Existence? A Reply to Alfonso Gomez-Lobo,” *Kennedy Institute of Ethics Journal* 17: 297–310.
- DeGrazia, David and Joseph Millum, *A Theory of Bioethics* (Cambridge: Cambridge University Press, 2021).
- de Waal, Frans (1987), *Bonobo* (Berkeley: University of California Press).
- Ford, Norman (2002), *The Prenatal Person* (Oxford: Blackwell).
- Frankfurt, Harry (1971), “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* 68: 5–20.
- Gomez-Lobo, Alfonso (2002), *Morality and the Human Goods* (Washington, DC: Georgetown University Press).
- Goodall, Jane (1986), *The Chimpanzees of Gombe* (Cambridge, MA: Harvard University Press).

Jaworska, Agnieszka and Julie Tannenbaum (2013), "The Grounds of Moral Status," in Edward Zalta (ed.), *Stanford Encyclopedia of Philosophy* (<<https://plato.stanford.edu/entries/grounds-moral-status/>>).

Jaworska, Agnieszka and Julie Tannenbaum (2014), "Person-Rearing Relationships as a Key to Higher Moral Status," *Ethics* 125: 543–56.

Locke, John (1694), *An Essay Concerning Human Understanding*, 2nd edition.

Parfit, Derek (1984), *Reasons and Persons* (Oxford: Clarendon).

Parfit, Derek (2011), *On What Matters*, vol. 1 (Oxford: Oxford University Press).

Parker, Sue, Robert Mitchell, and H. Lyn Miles, eds. (1999), *The Mentalities of Gorillas and Orangutans* (Cambridge: Cambridge University Press).

United Nations (1948), *Universal Declaration of Human Rights*.

Warren, Mary Anne (1997), *Moral Status* (Oxford: Oxford University Press).

White, Thomas (2007), *In Defense of Dolphins* (Oxford: Blackwell).

4

The Moral Status of Conscious Subjects

Joshua Shepherd

1. Theorizing about Moral Status

Any account of moral status must specify the grounds of moral status.¹ How are we to do that? Abstractly, two approaches are available.

A *members-first* approach begins by collecting judgments regarding who has moral status, and perhaps by collecting comparative judgments regarding who has higher degrees of moral status. Any account that begins with the judgment that humans have full (or the highest level of) moral status, and then seeks to justify that judgment, embodies this approach. So too does any account that begins with the judgment that healthy adult humans are at least the paradigmatic case of an entity with moral status.

This approach can seem epistemically modest, in the sense that it allows us to move from something we seem to be in a decent position to know, namely the grounds of the moral status of adult human beings, to elements that are more difficult to know, namely the grounds of whatever moral status other beings have. But I find the approach pernicious.

Adult humans are complicated creatures, with a range of potentially morally relevant capacities and properties. Theorists have variously seized on many of these to offer accounts of the grounds of moral status. These include possession of self-consciousness (Tooley 1972), possession of sophisticated psychological capacities (McMahan 2002), possession of "typical human capacities" (DiSilvestro 2010), possession of the capacity to participate in a "person-rearing relationship" (Jaworska and Tannenbaum 2014), possession of a capacity for intentional agency (Sebo 2017), the ability to take oneself to be an end rather than a mere means in the sense that one can experience and pursue what is good for one (Korsgaard 2013), the capacity to suffer (Bentham 1996), possession of the genetic basis for moral agency (Liao 2010), and no doubt more. Of course, some of these are friendlier to entities outside the tight circle of healthy adult humans, and some are not. What is striking, however, about many of these accounts is that they are strikingly