

# Why is it possible to enhance moral status and why doing so is wrong?

Nicholas Agar

## Correspondence to

Dr Nicholas Agar, Victoria University of Wellington, Philosophy Program, PO Box 600, Wellington 6140, New Zealand; nicholas.agar@vuw.ac.nz

Received 23 February 2012  
Accepted 20 July 2012  
Published Online First  
24 August 2012

## ABSTRACT

This paper presents arguments for two claims. First, post-persons, beings with a moral status superior to that of mere persons, are possible. Second, it would be bad to create such beings. Actions that risk bringing them into existence should be avoided. According to Allen Buchanan, it is possible to enhance moral status up to the level of personhood. But attempts to improve status beyond that fail for want of a target – there is no category of moral status superior to that of personhood. Buchanan presents personhood as a threshold. He allows that persons may succeed in enhancing their cognitive and physical powers but insists that they cannot enhance their moral status. I argue that it is an implication of accounts that make a cognitive capacity, or collection of such capacities, constitutive of moral status, that those who do not satisfy the criteria for a given status find these criteria impossible to adequately describe. This obstacle notwithstanding, I offer an inductive argument for the existence of moral statuses superior to personhood, moral statuses that are necessarily beyond human expressive powers. The second part of this paper presents an argument that it is wrong to risk producing beings with moral status higher than persons. We should look upon moral status enhancement as creating especially morally needy beings. We are subject to no obligation to create them in the first place. We avoid creating their needs by avoiding creating them.

This paper advances two claims about human enhancement. First, it is possible to use biotechnological or cybernetic means to enhance the moral status of humans. Second, doing so is morally wrong.

The paper's first task is to defend the possibility of enhancing moral status. According to what has become the received view, it is possible to enhance moral status up to the level of personhood. But attempts to improve status beyond personhood fail for want of a target—there is no category of moral status superior to that of personhood. The first part of this paper challenges a canonical presentation of this view by Allen Buchanan.<sup>1</sup> According to Buchanan, genetic and cybernetic technologies may succeed in enhancing our cognitive and physical powers but they cannot enhance our moral status. He disputes the possibility of post-persons, where post-personhood denotes a status superior to personhood. Buchanan asks why, if there are moral statuses higher than personhood, we find it so difficult to describe them? I call this the *inexpressibility problem*. It is an implication of accounts that make a cognitive capacity, or collection of such capacities, constitutive of moral status, that those who do not satisfy the criteria for a given status find these criteria impossible to adequately describe. I offer an

inductive argument for the existence of moral statuses higher than personhood.

The second part of this paper presents an argument that it is wrong to risk producing beings with moral status higher than persons. We should look upon moral status enhancement as creating especially morally needy beings, beings whose needs should take precedence over our own. We are subject to no obligation to create them in the first place. We avoid creating their needs by avoiding creating them.

## PART 1: AN INDUCTIVE ARGUMENT IN FAVOUR OF THE POSSIBILITY OF ENHANCING MORAL STATUS

There are two ways in which human beings might be morally enhanced. Humans could undergo either *moral disposition* enhancements or *moral status* enhancements. For brevity's sake I shall sometimes refer to the first as disposition enhancements and the second as status enhancements.

The aim of moral disposition enhancements is to increase the moral value of an agent's actions or character.<sup>2–4</sup> Julian Savulescu and Ingmar Persson<sup>5</sup> describe how we might enhance moral dispositions by boosting empathy and cooperativeness. Thomas Douglas<sup>4</sup> argues that we increase this moral value by attenuating certain counter-moral emotions. This paper advances no claims about the feasibility or justice of disposition enhancements.

The aim of moral status enhancement is not to increase the moral value of our actions or characters. Rather it increases a being's entitlement to certain forms of beneficial treatment and reduces its eligibility for certain forms of harmful treatment. The following elaboration of this basic idea is due to Buchanan. He proposes that a being has moral standing 'if it counts morally, in its own right.'<sup>1</sup> (p. 346) Moral status differs from moral standing in being a comparative notion. Suppose two beings both have moral standing. One may have higher moral status than the other. That is, one being may count for *more* morally in its own right than the other.

Some accounts find connections between status enhancements and disposition enhancements. For example, those who take a Kantian outlook think that 'personhood' indicates both a moral status and a moral disposition. It is the capacity of persons to make moral judgments that entitles them to respect.

Much of the discussion in this paper concerns a particular moral status—that indicated by the term 'person.' The concept of personhood that occupies the central location in Buchanan's discussion is a Kantian one according to which a person



- ▶ <http://dx.doi.org/10.1136/medethics-2012-100834>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-100835>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-100836>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-100837>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-100843>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101095>



Scan to access more free content

is a being with a capacity for practical rationality. Persons can both be held accountable and hold others accountable. While much of this paper's discussion assumes this account of personhood, the points I make should apply to a Lockean account that identifies persons as rational, self-conscious beings, who are aware that they have interests that persist over time. They apply to any account that makes a cognitive capacity essential to personhood.

One might enhance the moral status of a sentient nonperson by introducing into it the cognitive traits sufficient for personhood. For example, suppose one were to make cybernetic modifications to a sheep, giving it mental powers identical to those of human persons. It is possible that the precise modifications of cognitive and affective powers required to turn sheep into Kantian persons will differ from those required to turn them into Lockean persons. Each should allow, depending on the precise nature of the psychological enhancements, that the sheep has undergone status enhancement. It should now qualify for the moral protections due to persons.

### WHY IS IT SO DIFFICULT TO ENHANCE THE MORAL STATUS OF PERSONS?

It is possible to enhance status up to the level of personhood. But is it possible to enhance status beyond this point? Buchanan's serious doubts about the possibility of enhancing the status of human persons derive from a claim he calls the moral equality assumption. This is the idea that 'all who have the characteristics that are sufficient for being a person have the same moral status.' (p. 347)

I follow Buchanan in using the term 'mere person' to indicate a being who satisfies the criteria for personhood but fails to satisfy any criteria for a higher moral status. Post-persons would, on this understanding, be persons, but not *mere* persons. So why does Buchanan find it unlikely that such beings could exist?

Buchanan presents personhood as a *threshold concept*. Giving a being who does not meet the criteria for personhood greater powers of practical rationality may improve its moral status by enabling it to satisfy those criteria. Enhancements beyond this point make no difference to whether or not a being satisfies these criteria. They should therefore not place it in a moral category superior to that of persons. Buchanan says 'If a person's capacity for practical rationality or for engaging in practices of mutual accountability or for conceiving of herself as an agent with interests persisting over time were increased, the result presumably would be an enhanced person, not a new kind of being with a higher moral status than that of person.' (p. 359)

### THREE OBSTACLES TO MORAL ENHANCEMENT

There are three obstacles to recognising a moral status higher than personhood.

1. *The problem of the logic of thresholds*: When used to indicate moral status, personhood is, according to most analyses, a threshold concept. Once one satisfies it, additional increments of the properties relevant to satisfying it make no moral difference. How can a higher moral category exist if personhood corresponds with a threshold?
2. *The problem of how to improve on inviolability*: According to an analysis favoured by Buchanan persons differ from nonpersons in being morally inviolable. Suppose that we accept this analysis. How could enhancement improve the most status of a being whose most fundamental rights already cannot be violated?

3. *The problem of describing a higher moral status*: What are the criteria for moral statuses higher than personhood? It is difficult to imagine what the criteria for post-personhood might be. Attempts to specify them seem to succeed only in adding to the powers of persons in ways that enhance them but make no difference to their moral status. It is relatively easy to imagine enhancements that make persons more intelligent. What is difficult is seeing how these changes could enhance moral status.

### The problem of the logic of thresholds

Buchanan presents the concept of moral status indicated by personhood as a threshold concept, not a scalar concept. Wealth is a scalar concept. The degree of one's wealth increases with the acquisition of additional quantities of money and items of monetary value. There is no point at which additional quantities of money or valuable items cease to make a difference to one's wealth—one can always become wealthier by acquiring more money or items of value. Having the moral status of a person is, in contrast, a threshold concept. Buchanan says 'according to theories that accord moral status (or the highest moral status) to persons, understood as beings who have the capacity for practical rationality or for engaging in practices of mutual accountability, what matters is whether one has the capacity in question. Once the threshold is reached, *how well* one reasons practically or *how well* one engages in practices of mutual accountability does not affect one's moral status.' (p. 357)

The threshold view has much appeal. It explains an observed moral equality among persons who satisfy the criteria for personhood. Among those who satisfy the criteria there is wide variation in the relevant capacities. It is apparent that some human persons are better than others at reasoning practically. Yet we strongly resist acknowledging many moral statuses to correspond with different levels of attainment in practical reasoning.

In what follows, I explore two ways in which we might explain the observed moral equality of all human persons. One posits what I will call a *strong threshold* in moral status. The other posits *weak thresholds*. Weak thresholds are compatible with the enhancement of moral status beyond personhood. Thomas Douglas<sup>5</sup> offers a very useful taxonomy of possible patterns that status enhancements may follow.

*A strong moral status threshold*: a point or region beyond which *no* improvement to the capacities relevant to moral status makes any difference to status.

*A weak moral status threshold*: a point or region beyond which *moderate* improvements to capacities relevant to moral status make no difference to status. Improvements of greater magnitude *could* make a difference to status.

Buchanan's moral equality assumption posits a strong threshold in moral status. A weak threshold differs in allowing that improvements to the capacities relevant to moral status could make a difference to status so long as these improvements are more than moderate. If we understand moderate improvements as encompassing the entire observed human range above the minimum criteria for personhood, then the existence of a weak threshold in moral status is compatible with the observed moral equality of all human persons. Positing a weak threshold in moral status permits but does not settle the question of whether enhancement of practical reason well beyond the human range might have this effect.

The notion of a weak threshold may seem less intuitive than that of a strong threshold. It is not as theoretically tidy. But it is a notion that we make frequent use of. We often place more than one threshold along a single axis of human ability. Since the placement of one strong threshold precludes the placement of a second strong threshold these must be weak thresholds.

Consider an example from outside of the moral domain. Suppose you decide to improve your Spanish language abilities. You sign up for classes at a school which administers a test. The test uses weak thresholds to place you in an appropriate class. If you know no, or close to no Spanish you go into an introductory class. More knowledgeable students are placed in an intermediate class. Students with the best language skills enter an advanced class. There is some variation in the language abilities of students who find themselves in the intermediate class. Some barely avoid placement in the introductory class; others fall just short of the standard required for enrolment in the advanced class. The existence of a weak threshold means that variation in language abilities over this range makes no difference to the class in which a student is placed. But variation of greater magnitude does.

If we suppose that personhood indicates a weak threshold in moral status then the logic of thresholds does not preclude the existence of moral statuses higher than personhood.

### The problem of how to improve on inviolability

If persons are already morally inviolable then what higher form of respect is due to post-persons? How could the respect we owe to morally inviolable persons differ from that we owe to post-persons?

As part of his argument that post-persons (McMahan's preferred term is 'supra-persons') are possible, Jeff McMahan rejects an absolutist reading of inviolability according to which there are no circumstances in which it could be right to sacrifice an inviolable being.<sup>6</sup> Instead we should allow that there are degrees of inviolability. It is conceptually possible to create beings who are more inviolable (or less violable) than persons.

McMahan argues that this approach explains commonsense verdicts about what can and cannot be done to morally inviolable being. There is no absolute prohibition on sacrificing persons. It is permissible to intentionally kill persons in some supreme emergencies—circumstances in which the sacrifice is required to prevent the deaths of a very great number of persons. For example, it would be morally acceptable to sacrifice a small number of nuclear power plant workers to prevent its reactor from going critical and destroying a local town. According to McMahan, this judgment about inviolability is best explained by ascribing to humans a very a very high degree of inviolability rather than attributing absolute inviolability. This analysis of inviolability would permit post-persons to differ morally from persons in having an even lower degree of violability.

McMahan gives an example that illustrates this higher degree of inviolability. He supposes that there is some number of innocent lives for which it would be right to sacrifice an innocent person. McMahan reasons that the higher inviolability of post-persons could make it impermissible to sacrifice them for the purpose of saving this number of innocent humans.

### The problem of describing a higher moral status

We come now to the most serious obstacle to higher categories of moral status. There seems a significant barrier in grasping the criteria that one must satisfy to be correctly pronounced a post-person. It is easy to imagine beings who are more

intelligent than we are. But it is difficult to see how this greater intelligence could place them in a higher moral category.

Buchanan allows that this point is not decisive. He urges that we not 'confuse a failure of imagination with conceptual incoherence.' (1, p. 359) But he nevertheless says 'In the absence of an account of what the higher threshold would be like, the claim that there could be beings at a higher threshold who would have a higher moral status is not convincing.' (p. 363). Those who assert the possibility of higher moral statuses owe either a description of the criteria for a higher moral status, or an explanation of why the great difficulty in producing such an account might nevertheless be compatible with the existence of such categories. In what follows I offer an explanation for the apparent inexpressibility of moral statuses higher than personhood. I argue that, if these criteria are constituted by a cognitive capacity, or collection of cognitive capacities, then it is reasonable to expect that they will be difficult for mere persons to formulate. This fact notwithstanding, it is possible to infer their probable existence. It should be viewed as improbable that there would be no categories of moral status higher than persons.

### CRITERIA FOR HIGHER MORAL STATUSES AND THE INEXPRESSIBILITY PROBLEM

We seemingly have little difficulty in describing many of the things that cognitively enhanced beings might do. They might perform fiendishly complicated mathematical calculations in their heads; they might take minutes to read and internalise the entire contents of the 20 volume edition of the Oxford English Dictionary, and so on. We cannot do these things but we have little difficulty in imagining much of what doing them would involve. Of course, some mental performances of enhanced humans might not occur to us. But if we were to be alerted to the possibility of these performances then there should be no obstacle to imagining them. These imaginings may not be particularly detailed—but they should be sufficiently detailed for us to be reasonably confident that what we are attempting to imagine is, in fact, possible. There is, however, a species of enhancements which causes a systematic failure of our imaginative powers. We cannot imagine enhanced humans doing logically impossible things—such as proving that the square root of 2 is a rational number. And this is something that no degree of cognitive enhancement could permit. If we cannot imagine the improvement of cognitive powers enhancing moral status then might it, like proving the rationality of the square root of 2, be impossible?

In what follows I argue that we cannot express these criteria because they are constituted by capacities that are *cognitive*. The fact that criteria for post-personhood are cognitive is a barrier to mere persons' powers of expression and imagination. We, nevertheless, have reason to believe in them.

The Kantian analysis preferred by Buchanan places beings in the category of persons by virtue of possession a cognitive capacity—one is a person by virtue of the capacity for practical reason, the capacity to engage in practices of mutual accountability. It is a feature of a criterion determined by a cognitive capacity that those who do not satisfy it are typically unable to properly understand it. If we really understood the forms of practical reasoning that were constitutive of post-personhood then we would be able to perform them and as such would satisfy the criteria for post-personhood. Even error-riddled, halting performances of these cognitive tasks should qualify us for inclusion in this moral category. Our fellow post-persons may view us as particularly



dull members of their moral category, but they should, nonetheless, accept us as their moral equals.

We should not underestimate the manner of understanding that the Kantian approach to moral status supposes. It is possible that a being who just fails to satisfy the requirements of higher moral status would notice some of the consequences of having a lower status. This being might notice that beings that resemble it are frequently caused to suffer by beings that differ from it. This increased probability of suffering could be a consequence of the being's possessing a lower status. Higher status beings might, correctly, be sparing themselves suffering by redirecting it to lower status beings. The lower status being might make such observations without without truly understanding why it has a lower status. We might make such observations of the actions of post-persons while understanding very little about how they justify their choices. We would fail to possess a moral status implied by that understanding. Consider an analogous case in respect of mathematical understanding. Someone asked whether she *really* understands Gödel's Incompleteness Theorem is being asked to do more than just confidently assert the proof shows that arithmetic is incomplete. She is being asked whether, if handed pencil and paper, she could perform the key steps in the proof. Among those who really understand Gödel's proof some produce those steps with greater ease than others. But all who claim a genuine understanding ought to be able to produce them under some circumstances.

So it is at least compatible with our apparent cluelessness in respect of higher moral statuses constituted by a cognitive capacity that they do exist but we are unable to properly describe them. Of course, it is also compatible with our apparent cluelessness that there are no such criteria. In what follows I present a *moderately strong inductive argument* for the existence of such criteria.

### WHY SHOULD WE DEFER TO THE VIEWS ABOUT MORAL STATUS OF (SINCERE) COGNITIVELY ENHANCED BEINGS?

How can we overcome the inexpressibility problem? I have argued that the fact that higher moral statuses are constituted by a cognitive capacity or a collection of such capacities means that we who do not satisfy these criteria cannot adequately describe them. We can, in principle, overcome the limitations on our powers of expression by deferring to beings who are properly able to grasp these criteria. For example, in disputes about moral status we should recognise the assessments of beings who lack our cognitive and imaginative limits as, in principle, superior to our own. The question of whether we should recognise the existence of higher moral statuses becomes the question of whether beings who are sufficiently cognitively superior to us and lack our imaginative limits would recognise the existence of such statuses.

Our deference to beings who lack our imaginative and intellectual limits resembles that which moderately talented students of mathematics grant to those whose mathematical skills are manifestly superior to their own. If you believe in the sincerity and superior mathematical skills of an interlocutor you should sometimes believe her even when she presents conclusions about mathematics that appear unlikely. For example, it continues to seem absurd to me that 0.99 recurring could be identical to 1. But my deference to the superior mathematical judgment of others leads me to believe it.

This approach makes sense according to the prominent view of moral truth defended by Michael Smith.<sup>7</sup> According to Smith, true moral claims are those that would be assented to

under conditions of ideal rationality and ideal information. We will never achieve these conditions. Nor will cognitively enhanced beings. Smart though they may be, they cannot know every morally relevant fact and be logically incapable of error. But we should acknowledge their perspective as superior to our own. Cognitively enhanced beings should know all the morally relevant facts known by persons. In addition, they should know facts about moral status that unenhanced humans do not. There is reason to think that their conclusions are being more likely to be vindicated, at the limit of rational inquiry, than are our own.

### WHY SUFFICIENTLY COGNITIVELY ENHANCED BEINGS WILL (PROBABLY) FIND THAT COGNITIVE DIFFERENCES BETWEEN THEM AND US MARK A DIFFERENCE IN MORAL STATUS?

So far I have argued that we should defer to the assessments of moral status offered by sincere cognitively superior beings. We should acknowledge them as better informed, and better able to reason about relevant moral facts than we are. They lack some of our cognitive and imaginative limits and so are better able to assess claims to a higher moral status than personhood. So, how likely is it that cognitively enhanced beings will recognise the existence of higher moral statuses?

In the remainder of this first section I advance two claims. First, there is no upper limit to the capacities that determine moral status. Second, it is likely that cognitively enhanced beings would recognise a higher moral status somewhere in the spectrum of capacities higher than human persons.

Consider the following examples of bounded and unbounded capacities. There is an upper limit on the capacity to speak the English language. It is possible to specify perfect knowledge of English. This might involve knowing all of the vocabulary items and rules of grammar that are properly part of the language. It is unlikely that any current or past speaker of English has perfect knowledge of his or her language. But it is, at least in principle possible, to have knowledge of the English language that could not be bettered. There is a finitely large community of speakers of the language, and one could not do better than knowing everything about the language that these speakers collectively know.

There is, in contrast, no limit on possible improvements of the capacity for mathematics. It is likely that there is a very large but finite collection of mathematical truths knowable by humans. But there are almost certainly truths beyond these. There is no reason to believe that the limits of mathematics must be tied to the limits of the understanding of humans or of any other being.

Knowledge of practical reasoning seems more like knowledge about mathematics than it is like knowledge of the English language. It differs from knowledge of English in making no indispensable reference to principles or ideas known by a community. It is something that can be improved. There is no reason to believe in preset limits on logical reasoning, the power of abstraction, memory, or any other of the abilities that jointly constitute practical reasoning.

### TWO HYPOTHESES ABOUT HIGHER MORAL STATUSES

Here are two hypotheses about moral enhancement that are logically compatible with the fact that powers constitutive of practical reasoning can always be improved.

Hypothesis 1 There is *some degree* of improvement of capacities constitutive of status that cognitively superior beings

would recognise as indicating a moral status higher than personhood.

Hypothesis 2 There is *no degree* of improvement of these capacities that cognitively superior beings would recognise as indicating a moral status higher than personhood.

The available evidence should lead us to prefer hypothesis 1 to hypothesis 2. There is a limitless space of possible improvements of practical reason. There is inductive support for the notion that some degree of improvement of traits relevant to status produces a moral status superior to personhood.

Many philosophers acknowledge at least three different moral statuses up to and including persons. There are inanimate objects such as rocks that possess zero moral status. They are properly counted as possessing a moral status rather than lacking one. An example of something that fails to have a moral status is the concept of roundness. If presented with a dilemma in which you were required to sacrifice either a person or the concept of roundness you would have difficulty in establishing what was being requested of you. Those who seek moral trade-offs between objects that fall into the category of objects with zero status and objects that belong to higher categories receive clear advice on which objects should be favoured. Rocks should be sacrificed ahead of persons. Then there are sentient nonpersons. These include sheep, cats, and dogs. They count morally, in their own right, but to a lesser extent than persons, members of the third moral status. You would be making a moral mistake if you rescued a rock from a burning building, deliberately leaving a cat to be consumed by the flames. You would also be mistaken if you chose to rescue the cat rather than a person.

Consider now the vast expanse of possible improvements to the human capacity for practical reason. Given the existence of three distinct moral statuses in the range of mental powers of which we currently have direct experience, it seems unlikely that no moral statuses higher than personhood could occupy this expanse. Our modest cognitive powers mean that we won't understand exactly how enhanced cognitive powers would grant a higher moral status. Remember, however, that in these matters, we are deferring to beings with understanding about morality superior to ours.

This is an inductive argument. It has some of the limitations common to inductive arguments with small evidential bases. Compare it with another inductive reasoning that seeks to reach beyond the limits of human experience. Scientists engaged in the Search for Extraterrestrial Intelligence (SETI) are searching for signs of intelligence originating from distant stars. They are also interested in determining how likely, given what we have observed so far, this search is likely to be successful. We have direct evidence for intelligent life only on one planet—our own. On the other hand, the universe is vast. There are many billions of unobserved planets. Intelligent life evolved on Earth so it is unlikely that there could be a law of nature preventing its evolution elsewhere.

The inductive argument for intelligent life beyond Earth is comparatively weak because it moves from a single observation. There is only one planet in the Universe that we know to have intelligent life. The inductive argument for higher moral statuses resembles it in moving beyond evidence directly available to humans. It is stronger than SETI's inductive argument in virtue of the fact that it extrapolates from three observed moral statuses.

Moral status category 1: The zero moral status possessed by rocks.

Moral status category 2: The moral status possessed by sentient nonpersons such as sheep and toads.

Moral status category 3: The moral status possessed by persons.

These observations of moral status make it reasonable to believe in the existence of moral status category 4 that includes post-persons.

I propose that our observations of moral statuses make it likely that beings lacking our cognitive limits will recognise moral statuses superior to personhood.

## PART 2: WHY MORAL STATUS ENHANCEMENT IS MORALLY WRONG?

In what follows I present an argument for avoiding the creation of post-persons that appeals to consequences. Degrees of cognitive enhancement that risk moral status enhancement should, by implication, also be avoided. The bad consequences that I discuss are not certain. But they are sufficiently probable and bad to justify limiting cognitive enhancement. The argument presents post-persons as especially morally needy beings. Once brought into existence their needs should take precedence over our own. But there is no obligation to create post-persons. We appropriately prevent the existence of the needs of post-persons by refraining from creating post-persons.

Consider the widely discussed argument against anthropogenic climate change. This argument draws support from models of the climate that attribute some significant part of climate change to human causes. According to this model, current or increased human production of greenhouse gasses may have quite disastrous consequences for the planet's human and nonhuman inhabitants. While not certain, these bad consequences are sufficiently bad to justify reducing greenhouse gas production. I claim to be certain neither of the possibility of moral status enhancement nor of its bad consequences. My claim is that the bad consequences are, in moral terms, so bad that a moderate probability of their occurrence makes it wrong not to seek to prevent them.

The key objection against further anthropogenic climate change is not that no one will benefit. Technologies that produce greenhouse gases benefit many—they provide employment and returns on investments. Rather it's that there are possible or probable consequences of climate change so bad that these jobs, profits, and other good effects do not compensate for them. The consequences of moral status enhancement conform with this pattern. Any benefits received by recipients of moral enhancement do not, in moral terms, make up for the costs imposed on others.

## THE MORAL COSTS OF MORAL ENHANCEMENT

How might moral status enhancement lead to bad consequences? According to a scenario frequently represented in science fiction, cognitively and technologically superior beings victimise humans. This is a significant concern of Francis Fukuyama who worries that posthumans might enslave humans.<sup>8</sup> This discussion supposes that degrees of moral status are recognised and respected. I assume that if costs are to be imposed on mere persons they must be morally justified. There will be an orderly transition from societies composed exclusively of mere persons to what I will refer to as *mixed societies*—societies constituted by both mere persons and post-persons. This orderly transition occurs without significant violations of the moral entitlements of mere persons.

Consider an analogous issue in respect of the climate change. Some commentators think that a loss of habitable and cultivable land will lead to widespread unrest as groups fight for control over diminishing supplies of life-sustaining resources.

But this is no necessary consequence. It is possible that there will be an entirely orderly transition from circumstances of relative plenty to circumstances of relative scarcity; decisions about how to distribute the diminished supplies of vital resources will be made in accordance with all relevant moral criteria. Any complaints from those deprived of food and shelter should be directed not at the decisions which were the proximal causes of their deaths but at earlier choices that permitted circumstances such as these to arise.

For the purposes of objecting to moral status enhancement I make an analogous assumption. Post-persons will not abuse their power over mere persons. They will treat mere persons exactly as they deserve. This will entail sacrifices by mere persons. Mere persons may not complain about the decision by post-persons to impose those sacrifices. The difference in moral status between post-persons and mere persons makes them legitimate. But mere persons can complain about choices that anticipate the morally correct choices by post-persons—namely the choice to create beings with enhanced moral status.

### SUPREME EMERGENCIES AND SUPREME OPPORTUNITIES

For some idea of the circumstances under which mere persons might suffer losses we should begin with McMahan's discussions of supreme emergencies. These rare circumstances permit the sacrifice of individuals with the highest moral status to save a (much) greater number of individuals with that same status. McMahan surmises that mere persons are likely to be better candidates for sacrifice in supreme emergencies than are beings with a moral status superior to persons. If supreme emergencies remain rare then this may not be such a big loss. Perhaps mere persons will feel adequately compensated by increases in productivity and the expansion of scientific knowledge brought by the cognitive enhancements that enhance the moral status of other citizens.

The chief danger for mere persons lies in a foreseeable expansion of the range of cases for which they may be justifiably sacrificed. Consider another category of circumstances in which mere persons might be sacrificed. In supreme emergencies morally significant beings are sacrificed to prevent significant harms. *Supreme opportunities* arise in respect of significant potential benefits best secured by sacrificing morally considerable beings.

Situations in which we might sacrifice morally considerable beings to secure benefits arise more frequently than situations in which we might sacrifice morally considerable beings to prevent emergencies. Suppose that emergencies are rare. Humans routinely seek benefits by sacrificing morally valuable individuals. We eat the flesh of sentient nonpersons because we find it both tasty and nutritious. We conduct painful and lethal medical experiments on sentient nonpersons for clues about human diseases. We insist that suffering be minimised and not be inflicted for frivolous reasons. But we nevertheless permit it.

### HOW SUPREME OPPORTUNITIES COULD JUSTIFY THE SACRIFICE OF MERE PERSONS?

Could the promise of benefits for post-persons justify the sacrifice of mere persons? One difference between supreme emergencies and supreme opportunities suggests that the answer to this question may be no.

Suppose that we limit our attention to supreme emergencies and opportunities that involve only persons. Some supreme emergencies permit the sacrifice of persons. But supreme opportunities involving only persons seem not to permit the sacrifice of persons. Consider the example of possible supreme benefits

arising in respect of medical research. There are strict rules restricting what can be done to human subjects in medical experiments. Such restrictions seem appropriate even when we believe that great benefits could result from a more permissive attitude.

What should we take out of the asymmetry between supreme emergencies and supreme opportunities? Supreme emergencies seem to permit the sacrifice of persons to prevent large numbers of persons from suffering significant harms while supreme opportunities seem not to permit persons to be sacrificed to bring great benefits.

There are two ways in which we can explain why persons cannot be sacrificed in pursuit of benefits for other mere persons. The first explanation points to a non-relational property of persons. If we assume Buchanan's approach to persons, we would say that persons cannot be sacrificed to produce great benefits because they are capable of practical reason. This interpretation should transfer to situations involving post-persons. The act of creating post-persons would not, it itself, render mere persons incapable of practical reason.

According to a second interpretation of the prohibition, a relational property of persons prevents their sacrifice. They have the highest moral status of all involved beings. On this second interpretation, the act of creating post-persons would, in itself, dislodge mere persons from the role of having the highest moral status. Their moral demotion would render them eligible for sacrifice to provide significant benefits for post-persons.

There is inductive support for the second interpretation. Other gaps in moral status of which we are aware seem to permit the sacrifice of lower status beings to benefit their moral betters. It is morally permissible to sacrifice objects with moral status zero to produce benefits for sentient persons—we find it acceptable to feed carrots to rabbits. It is morally permitted to sacrifice sentient non-persons to produce benefits for persons—we find it acceptable to conduct painful and lethal experiments on rhesus monkeys to find better treatments for serious diseases suffered by human persons. These permissions provide inductive support for a permission to sacrifice mere persons to benefit post-persons.

There is some support for the first interpretation in Kantian talk of the incomparable moral value of persons. If post-persons are possible then there could be beings whose value is not only comparable with that of mere persons, but superior to it. We could, however, grant all of Kant's normative conclusions about the situations that most interested him. In situations involving only mere persons and beings of equivalent or lower status, one would be making a serious mistake if one proposed that benefits to mere persons could justify the deliberate sacrifice of another mere person. We are required to emend Kantian views about the incomparable value of (mere) persons only in circumstances involving post-persons.

McMahan's discussion focuses on supreme emergencies. He proposes that different treatment in supreme emergencies would suffice to give one category of beings a moral status higher than members of the other category. In limiting his discussion to supreme emergencies, McMahan does not do more than is required to establish the possible existence of beings with a status higher than persons. His argument should not be read as implying that supreme emergencies are the *only* circumstances in which differences between mere persons and post-persons become apparent. It is perfectly compatible with McMahan's view that the difference in status licences other forms of discrimination.



Thomas Douglas has suggested to me that if the differences in status are real but small then morality might require that mere persons be distinguished from post-persons *only* in supreme emergencies. There are good inductive grounds for thinking that the differences in moral status between mere persons and post-persons will be quite significant. The gaps between the status of non-sentient things and sentient nonpersons and between sentient nonpersons and persons make big differences to permissible or required treatment. The differences are certainly not limited to supreme emergencies.

Remember that mere-personhood and post-personhood are different weak thresholds and not just different points on a graph registering continuous improvements of a morally relevant scalar property. If different points on a line registering continuous moral status improvement are close to one another it is reasonable to think that moral requirements and permissions may be similar. Weak thresholds combine a comparatively wide range of different degrees of morally relevant attributes. As a consequence there's likely to be a big difference between requirements and permissions appropriate for each weak moral threshold.

Consider the Spanish language school example I used to introduce the idea of weak thresholds. Weak thresholds separate students enrolled in introductory, intermediate, and advanced language classes. There would be a small difference between these classes if only a small number of vocabulary items or grammatical principles were taught in the higher class and not in the lower class. But this is unlikely to be the case. The instruction properly offered in introductory classes differs quite markedly from the instruction offered in intermediate language classes. This is so even if there is only a small difference in the language skills of the most knowledgeable student in the introductory class and the least talented student in the intermediate class.

### **WHY POST-PERSONS WILL PROBABLY IDENTIFY MANY SUPREME OPPORTUNITIES REQUIRING THE SACRIFICE OF MERE PERSONS?**

How does the addition of supreme opportunities add to the likely burdens of mere persons in a society some of whose members are post-persons? Suppose that there is an orderly transition to a mixed society. Supreme opportunities will permit mere persons to be sacrificed to provide significant benefits for post-persons. Just as human persons are morally entitled to sacrifice sentient nonpersons in pursuit of better treatments for serious human diseases, so too post-persons may be entitled to sacrifice mere persons to gain a better understanding of ailments afflicting them. Human medical researchers use monkeys because their relatedness to us makes them a useful model of human disease. Suppose that diseases afflict post-persons. The emergence of post-persons from human mere persons may make mere persons ideal subjects for medical experiments.

This is but one example of a possible use that post-persons may make of mere persons. The cognitive enhancement that may turn mere persons into post-persons is likely to generate uses for their human person ancestors that we cannot identify. The enhancement of cognitive powers that occurred with the evolution of humans from ape-like ancestors has created beneficial uses for many parts of the environment for which apes have no use. Super-intelligent post-persons are likely find beneficial uses of parts of their environment that we cannot think of. Some of those parts of their environment could include human brains and bodies.

We shouldn't presume too much insight into the designs of beings with radically enhanced intellects. The futurist Ray Kurzweil has a suggestion that should scare mere persons.<sup>9</sup> Kurzweil predicts that advances in information technologies will soon set off a progression of increasingly powerful cognitive enhancements. One way to enhance the processing power of our minds is to physically expand them. Cognitively enhanced beings would do the same thing to their minds that computer engineers do when they add more transistors to a computer. Kurzweil predicts a future in which the minds of enhanced humans colonise the universe. Every bit of matter and energy will become the substrate of, and fuel for, thought. This could include the matter and energy that constitute the brains and bodies of human mere persons.

I conclude that it is reasonable to think that the creation of post-persons will leave mere persons more likely to suffer significant harms.

### **WHY REDUCED IMMUNITY IS UNLIKELY TO BE ADEQUATELY COMPENSATED?**

Suppose that the creation of post-persons increases the frequency and severity of harms suffered by mere persons. These forecast harms might be justified if we believed that sufficient benefits accrued to those who will suffer them.

There are good inductive grounds for thinking that compensating benefits will not be forthcoming.

The status of the lower status beings seems more effective at protecting against harms inflicted by higher status beings than it is at procuring benefits from them. Take a case in which higher status beings produce benefits through inflicting harms on lower status beings. There are many medical experiments on nonhuman animals. Some of the diseases addressed by these experiments afflict the species that provide the experimental subjects. The fact that the members of a species have suffered so that a benefit can be produced seems to do little to establish a claim on that benefit. Rhesus monkeys may perform valuable services in research on Parkinson's disease. But we would consider treating them with the expensive therapies that such research produces only after all human patients have been adequately treated. We would appeal to facts about relative moral status to vigorously challenge a doctor who sought to place a rhesus monkey patient on a waiting list for treatment for Parkinson's disease ahead of a human patient.

The same points apply to other benefits resulting from technological advances. It would be viewed as a serious misallocation of resources to use new construction techniques to build very stimulating environments for sheep before using them to improve the residential circumstances of needy human persons. This would be so even if we thought that it was wrong to benefit the humans by causing the sheep to suffer.

There is, therefore, some inductive support for the notion that post-persons will allocate benefits to mere persons only when all of the needs of post-persons are met. The hopes of mere persons will depend on the predictions of some futurists that technological progress will create a super-abundance that enables the all of the interests of post-persons and mere persons to be concurrently satisfied.

### **CONCLUDING COMMENTS**

I have argued for two claims. First, it is likely that some degree of cognitive enhancement will enhance moral status. There are good inductive grounds for thinking that it will bring into existence post-persons—beings with a status superior to mere persons. Second, the creation of post-persons would be a

morally bad thing. It is likely to impose significant penalties on mere persons. The fact that there is no moral obligation to create post-persons means that we should not.

**Acknowledgements** Thomas Douglas, Allen Buchanan, Felice Marshall, Mark Walker.

**Contributors** NA is the sole contributor to this article.

**Competing interests** None.

**Provenance and peer review** Commissioned; externally peer reviewed.

## REFERENCES

1. **Buchanan A.** Moral status and human enhancement. *Philos Publ Aff* 2009;**37**:346–81.
2. **Walker M.** Enhancing genetic virtue: a project for twenty-first century humanity? *Pol Life Sci* 2009;**28**:27–47.
3. **Persson I, Savulescu J.** The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity. *J Appl Philos* 2008;**25**:162–77.
4. **Douglas T.** Moral enhancement. *J Appl Philos* 2008;**25**:228–45.
5. **Douglas T.** Cognitive enhancement and the supra-personal moral status. *Philos Stud* [in press].
6. **McMahan J.** Cognitive disability and cognitive enhancement. *Metaphilos* 2009;**40**:582–605.
7. **Smith M.** *The moral problem*. Oxford: Blackwell, 1994.
8. **Fukuyama F.** *Our posthuman future: consequences of the biotechnology revolution*. New York: Farrar, Straus & Giroux, 2002.
9. **Kurzweil R.** *The singularity is near: when humans transcend biology*. London: Penguin, 2005.