

In Defense of Artificial Replacement

published in *Bioethics* 31: 393-399.

Derek Shiller

Abstract

If it is within our power to provide a significantly better world for future generations at a comparatively small cost to ourselves, we have a strong moral reason to do so. One way of providing a significantly better world may involve replacing our species with something better. It is plausible that in the not-too-distant future, we will be able to create artificially intelligent creatures with whatever physical and psychological traits we choose. Granted this assumption, it is argued that we should engineer our extinction so that our planet's resources can be devoted to making artificial creatures with better lives.

The pace of technological change is very difficult to predict far in advance, but our current trajectory makes it reasonable to guess that we will have the power to create genuine artificial intelligence – artificially created individuals that equal or surpass human beings in all dimensions of cognition, including creativity, power, insight, and wisdom – by the close of this century. Some futurists¹ have worried about our species' continued existence after this development. Such concerns are motivated by the recognition that it may be difficult to predict and control artificial creatures that are smarter than we are. There is something selfish about this fear and the ethically responsible thing for us to do may be to engineer our own extinction.

In this paper, I will present a simple speculative argument for what I will call the *Artificial Replacement Thesis*. The Artificial Replacement Thesis suggests that we should replace our species with artificial creatures who are

1 V. Müller, ed. 2015. *Risks of Artificial Intelligence*. Boca Raton, FL: CRC Press.

capable of living better lives. I will start by introducing several assumptions that will be integral to my argument. In the second section, I will defend a supplemental principle that I call the *Future Beneficence Principle*, that says that we should go out of our way to improve the well-being of future generations, even if our actions will change who comes to exist. With these foundations laid, I will present my argument for the Artificial Replacement Thesis in the third section. I will spend the remainder of this paper formulating and replying to salient objections.

ASSUMPTIONS

My argument for the Artificial Replacement Thesis relies on two assumptions, which I will take for granted in the remainder of this paper.

First, I will assume that we will have the power to create intelligent artificial minds that resemble natural minds in every morally relevant way we wish. Morally relevant ways might include: consciousness, cognitive flexibility, emotional capacity, capacity for happiness and unhappiness, ability to engage in interpersonal relationships, creativity, freedom of the will (in whatever sense we have it), and philosophical, religious, or artistic insight. Whatever nature can do with clumps of neurons, we will be able to do artificially. If this assumption is correct, it means that with the right design, artificial creatures will be able to fall in love, experience exquisite joy, write novels that probe existential self-doubt, ponder the basic metaphysical structure of reality, and appreciate the beauty of mathematical theorems.

Perhaps the most controversial part of this assumption is the claim that it is possible for us to create conscious minds artificially. While this assumption is integral to my argument, it involves complex issues in the philosophy of mind, and so is not something that I can properly defend here.²

² A very basic argument goes as follows: The primary functions performed by neurons are being replicated by artificial analogues with increasingly impressive results. If pure neural interactions ultimately initiate and control our actions (all of the evidence from neuroscience suggests they do) and if the function of neurons can be artificially replicated, then there is little reason to doubt that we will be able to build artificial creatures that act just as we do by combining the artificial neurons into the right structures. Whether artificial creatures would not only act like us, but also experience things in the way that we do is open to debate, but there would be

Second, I will assume that it will be comparatively easy to avoid instilling our artificially intelligent constructs with the vicious traits that presently afflict humanity. Human bodies are the result of a variety of evolutionary pressures. These pressures promoted traits that enhanced our ancestors' survival, not their well-being. Consequently, human life was often nasty, brutish, and short.³ Though modern technology helps us to fare much better than our ancestors, we share many of their shortcomings.

Some of these shortcomings are physiological. Perhaps the foremost among these relate to the process of aging. As we age, our bodies deteriorate in a variety of ways: we feel joint pain, we lose our mobility and flexibility, our minds become less sharp, we get sick, and we die. Many of us are numbed to this inevitable tragedy because we are so accustomed to it, but it is a substantive limit to the amount of well-being that we can enjoy in our lives.

Other shortcomings are psychological. We suffer from unreasonable and unfulfillable desires. We want to be kinds of people that we cannot be. In pursuit of fleeting temptations, we are disposed to make decisions that go against our own interest. We are aggressive, callous, and cruel to each other. We harbor arbitrary biases against our fellow creatures based on irrelevant characteristics or group membership.

These are not the inevitable vices of any intelligent being. They are part of our species. We need not pass them on to our artificial creations. While there might be some constraints to the kinds of artificial minds that it is possible for us to make, I will assume that we will be able to build artificial creatures that are not afflicted by the same kinds of vices that we are. Our creations could be rational, intelligent, deeply caring, loyal, respectful, wise, good-humored, emotionally stable, and eternally healthy.

considerable pressure to think that they would.

The case for intelligence in artificial minds is similar to the case for intelligences in other animals. If they behave in a way that is similar to the way that we behave, as a result of similar cognitive processes, then we should think that they have similar experiences. Artificial minds might differ from ours in ways that the minds of other creatures do not, but they may also be more like us in other ways, such as in their higher-order cognitive structure.

³ For a more detailed argument, see D. Benatar. 2006. *Better Never to Have Been: The Harm of Coming Into Existence*. New York, NY: Oxford University Press.

THE FUTURE BENEFICENCE PRINCIPLE

In addition to the assumptions discussed in the last section, my argument will also make use of the Future Beneficence Principle.

Future Beneficence Principle: Where it is possible to greatly improve the well-being of future generations at a comparatively low cost to ourselves, we should do so, even if doing so will affect the identity of those future beings.

This principle borrows plausibility from its relation to the Future Nonmaleficence Principle.

Future Nonmaleficence Principle: Where it is possible to improve our well-being, at a comparatively far greater cost to future generations, we should not do so, even if doing so will affect the identity of those future beings.

Our obligations to the future are notoriously complicated by the fact that our present actions may influence who will come to exist in the future.⁴ The choice to conserve resources now may make future generations happier, but it will most likely also change the makeup of those generations. Since existence is a precondition for a valuable life, everyone who possesses a life worth living and whose existence depends on our choice will have benefited from that choice, no matter which choice we make.

Still, it is widely thought that we have moral reasons to act in ways that will lead to a better future. This may be because we have reasons to produce as many people as possible at maximal levels of well-being⁵. Alternatively, it may be

4 D. Parfit, 1984. *Reasons and Persons*. Oxford. Oxford University Press; T. Schwartz, 1978. Obligations to Posterity. In R. Sikora and B. Barry, eds., *Obligations to Future Generations*. Philadelphia, PA: Temple University Press: 3-13.

5 R. Sikora. 1978. Is It Wrong to Prevent the Existence of Future generations. In *Obligations to Future Generations*. R. Sikora and B. Barry, eds. Philadelphia, PA: Temple University Press: 112 –166; J. Leslie. The Need to Generate Happy People. *Philosophia* 1989; 19 (1):29-33; S. Rachels. Is It Good to Make Happy People? *Bioethics* 1998; 12 (2):93-110; M. Huemer. In Defence of Repugnance. *Mind* 2008; 117 (468):899-933; M. Gardner. Beneficence and Procreation. *Philos Stud* 2016; 173

because we have reasons to make sure that the people who will exist are as well-off as possible. Or it may be for some other reason. I will not offer any particular account of what explains our obligations to future generations. Instead, I will suggest that there is no clear dividing line between the Future Beneficence Principle and the Future Nonmaleficence Principle. If we cannot bear to deny the latter, we should accept the former as well.

While there is sure to be disagreement about the reasons, many will agree that it is wrong to do things that will greatly harm future generations in exchange for a comparatively small benefit to ourselves. It would be wrong, for example, to continue activities that contribute significantly to global warming for our own economic gain on the grounds that the people who will primarily suffer will also owe their existence to our decision. The Future Nonmaleficence Principle formalizes this idea.

The Future Beneficence Principle suggests that we have similar reasons to provide benefits to future generations when we have the ability to do so. The difference between acting to raise future well-being and abstaining from lowering future well-being is hard to make precise.

On a first pass, choosing not to make sacrifices to better others amounts to merely allowing them to live a worse life, whereas choosing to benefit ourselves at the expense of others amounts to acting to produce that worse life. This suggests that the difference between the two principles may depend on the doing/allowing distinction, which holds that in at least some cases, actively doing something is morally worse than merely allowing the same thing to occur.

The doing/allowing distinction is highly suspect⁶, and gains plausibility from its correlation with the intended/foreseen distinction, which suggests that it can make a moral difference whether or not one intended one's act to have certain consequences. While there might be a significant moral difference between intending to lower well-being and merely foreseeing that such will be a result of one's actions, the two principles do not make any assumptions about our intentions. The viability of the intended/foreseen distinction will not help justify differing verdicts on these principles.

(2):321-336.

⁶ For some criticisms of distinctions between acting and allowing, see J. Bennett. Negation and Abstention: Two Theories of Allowing. *Ethics* 1993; 104 (1):75-96.

Alternatively, we might think that the difference between future beneficence and future nonmaleficence turns on the difference between raising and lowering well-being. It is permissible to intentionally act in ways that fail to improve people's lives, but not permissible to intentionally act in ways that worsen people's lives.

There is a basic problem in applying this distinction to actions that affect the far future. If the identities of the individuals who come to exist are very sensitive to our actions, then there is little we can do to change the levels of well-being of specific future people: we can act in ways that will bring one group into existence with some level of well-being, or some other group with some other level of well-being. A requirement not to lower well-being must then be a requirement not to opt for lower levels of overall well-being, rather than a requirement not to lower any individual's well-being. If there is an imperative to not act in ways that will produce lesser levels of well-being over ways that will produce greater, we should expect to be symmetrically obliged to choose greater levels over lesser.⁷

I think we should accept the Future Beneficence Principle on the grounds that we should accept the Future Nonmaleficence Principle, and that it is hard to justify a distinction between them. However, even if this argument is found to be insufficient, there is an additional reason to think that the two principles stand or fall together: there would be unwelcome consequences to attributing a significant asymmetry to our moral reasons.⁸

⁷ It might be suggested that the difference comes down to the direction of divergence that the action produces from the current trajectory of well-being: we have no obligation to divert the future in a direction of more well-being, but we do have an obligation not to divert it in a direction of less. This may make sense when we can evaluate the trajectory in which history is proceeding independent of our own actions. However, when the shape of the future integrally depends on which actions we choose, as is the case with the Artificial Replacement Thesis, it is not obvious whether there is any good way to pin down a trajectory. In other words, the notion of moral inertia (see C. Sartorio. *Moral Inertia*. *Philos Stud* 2008; 140 (1):117-133.), which may drive many of our intuitions between the differences between beneficence and nonmaleficence, makes less sense in situations where the outcome is extremely sensitive to the precise action we choose to take.

⁸ This bears similarity to an argument made by Sikora, *op. cit.* note 5. See also

Surely, many of our actions will have both positive and negative effects in the far future. If we had special reasons not to harm that were not balanced by reasons to benefit, then we might be barred from doing things that on the whole lead to a positive situation but which included some significant negative consequences. It is plausible that the vast majority of choices about policies that have significant future effects are like that. In order to allow good consequences to balance out the bad, we need to give some moral weight to improving the lives of future generations.

In light of these two arguments, the Future Beneficence Principle should appeal to those who are taken by the Future Nonmaleficence Principle.

THE ARGUMENT

My particular application of the Future Beneficence Principle involves its implication that, given the choice, we should create beings with greater well-being over beings with substantially less well-being, when it is not too costly to us.

This application parallels the Principle of Procreative Beneficence,⁹ which enjoins us to 'select the child, of the possible children [we] could have, who is expected to have the best life, or at least as good a life as the others, based on the relevant, available information.' (p. 415)

Savulescu originally presented his principle in the context of choices of medical intervention into normal human reproduction. It implies that we should take steps to avoid having children with traits that are harmful to their well-being, and that we should opt for children with traits that are conducive to their

J. McMahan, 2009. Asymmetries in the Morality of Causing People to Exist. In D. Wasserman and M. Roberts eds. *Harming Future Persons*. Springer: 49 – 68; B. Bradley. Asymmetries in Benefiting, Harming and Creating. *J Ethics* 2013; 17 (1-2):37-49.

9 J. Savulescu. Procreative Beneficence: Why We Should Select the Best Children. *Bioethics* 2001; 15 (5-6):413-426; J. Savulescu and G. Kahane. The Moral Obligation to Create Children with the Best Chance of the Best Life. *Bioethics* 2009; 23 (5):274-290; J. Harris. 2007. *Enhancing Evolution: The Ethical Case for Making Better People*. Princeton, NJ: Princeton University Press.

well-being.¹⁰ This idea can be extended to apply to the question of whether or not we should opt to have normal human progeny or to create artificially intelligent beings. If we can provide artificial progeny a better life than biological progeny, Savulescu's principle seems to suggest that we should forgo natural reproduction on the grounds of beneficence. We should choose to create creatures with the best life. The fact that such creatures are made of silicon and do not emerge directly from our genitals is morally irrelevant.

Here is my argument: human beings live lives that are quite suboptimal. With a good design, we will be able to produce artificial creatures whose lives are much closer to being optimal. We will then be faced with the choice of continuing to populate the world with humans, or devoting resources over to creating creatures who are capable of much higher levels of well-being.

Our resources are finite, and the same resources that might allow human beings to live – effort, land, energy, raw materials – could be more effectively spent on creating and sustaining artificial creatures. When that becomes the case, the beneficent thing to do is to choose that our children be artificial, rather than natural. It will not harm us too much, and it will greatly benefit future generations.

Artificial Replacement Thesis: Once it is possible to design artificial creatures whose lives are significantly better than human lives (and at a more efficient use of our resources), we should engineer the extinction of the human race in order to route available resources to creating and sustaining them.¹¹

10 Savulescu's principle would do most of the work that I want the Principle of Future Beneficence to do, so why involve it? One main motivation is that I think that the source of our reasons to pursue replacement lies not just with the immediate succeeding generation, but all generations that follow. If we could create a perpetual utopia, as may be possible, our reasons for bringing it about will be stronger because the advantages will continue to accrue, and it will significantly raise the strength of our moral reason for making the switch.

11 A stronger thesis, which I take to be about as plausible, says that we should eradicate all animal life on Earth to devote resources to creating and sustaining artificial creatures. This would solve the problem of the immense amount of animal suffering integral to the sustenance of natural environments. While it is controversial whether we have any duty to intervene in nature (see for instance J. Hadley. *The Duty to Aid Nonhuman Animals*

This proposal should not be read as a justification for forcibly bringing about such a change against the wishes of currently existing people. Nor should it be read as involving the purposeful suicide of anyone. The extinction called for could be achieved by generational replacement, or perhaps a gradual petering out of humanity (where each generation is significantly smaller than the previous).

The Future Beneficence Principle only instructs us to act in the interests of future generations when it is not comparatively costly to ourselves. So in order for the thesis to be supported by the principle, it would need to be possible for human extinction to be carried out in relative comfort. Though one might imagine that the last generation of humans would feel anguish, despair, and loneliness, there is no reason why this must be the case. The last humans would have the company of not just each other, but also of their artificial progeny.

I purposefully have left the strength of the normative claim ambiguous. I am neither affirming nor denying that we have a moral *obligation*, as opposed to something less morally stringent. It may be that by not replacing our species, we will have done nothing wrong. I am not claiming that anyone should be barred from having natural children, if they so choose. A right to reproductive autonomy may grant us final moral license to choose for ourselves what kind of children to have. The conclusion of my argument should be read as saying only that the good and decent thing for us to do as a species is to replace ourselves.

This proposal is likely to be met with a great deal of skepticism. We value humanity, and treat the extinction of our species with fear. Though it has been convincingly argued¹² that the end of humanity sooner rather than later is not inherently bad in itself (and this should be especially true if we are replaced with something objectively preferable), we have an understandable attachment to our own species.¹³ This attachment manifests itself both in our comparative lack of

in Dire Need. *J Appl Philos* 2006; 23 (4):445–451 and J. Milburn. Rabbits, Stoats and the Predator Problem: Why a Strong Animal Rights Position Need Not Call for Human Intervention to Protect Prey from Predators. *Res Publica* 2015; 21 (3):273-289.) it seems like it would be good to replace the desperate and uncertain lives of wild animals with lives of optimal well-being.

12 J. Lenman. On Becoming Extinct. *Pac Phil Q* 2002; 83 (3):253–269.

13 Though see also J. Bennett. 1978. On Maximizing Happiness. In R. Sikora and B. Barry eds. *Obligations to Future Generations*. Philadelphia, PA: Temple University Press: 61 – 73.

concern for other animals, and in a repulsion to the Artificial Replacement Thesis. We would like human beings to always be around. Nevertheless, I do not believe that any weighty justification can be found for this preference. Before concluding, I will consider three objections and argue that none succeeds.

OBJECTIONS AND REPLIES

Objection 1: Imperfection is good.

According to the first objection, human imperfections actually make our lives better. If we were all extremely intelligent, coolly rational, disease- and disability-free, our lives would lose something of the depth and complexity that makes them valuable.¹⁴

While good in small doses, perhaps too much happiness and self-determination may be detrimental to us. A good life involves making the most with what we have. Perhaps it is valuable to accomplish difficult things,¹⁵ and this is not possible without being substantially imperfect.

There are surely some maladies that we are better off without, but perfection would not be worth the loss of depth of experience that our imperfections provide us.

Reply: Artificial creatures can also be imperfect.

I grant that lives are better when they are constrained. Overcoming obstacles contributes to the quality of our lives, and if we could get whatever we wanted whenever we wanted it, our lives might be happy but meaningless. It might be a necessary feature of any obstacle worth overcoming that we face some chance of failure to overcome it. And it may be that triumphing over self-imposed limitations does not add as much to the quality of our lives as meeting and exceeding externally-imposed limitations. If so, then we might think that our imperfections actually work in our favor.

14 E. Barnes. Disability, Minority, and Difference. *J Appl Philos* 2009; 26 (4):337-355; R. Garland-Thomson. The Case for Conserving Disability. *J Bioeth Inq* 2012; 9 (3):339-355.

15 G. Bradford. The Value of Achievements. *Pac Philos Q* 2013; 94 (2):204-224.

This only succeeds as an objection to the Artificial Replacement Thesis, however, if either our lives are already fairly close to being optimal, or if we are less able to imbue our artificial creations with the kinds of imperfections that make our lives better. Both ideas are dubious.

First, it seems extremely unlikely that we should have lucked into just the right amounts of misery and impediments for an optimal life. It is implausible that the majority of detriments we face are beneficial. It is certainly not something we generally think about our own lives. We strive to improve them by bettering ourselves. We don't consider ourselves lucky to be rash, cruel, or susceptible to cancer. We don't attempt to teach our children poor money management skills to make their eventual financial stability all the more meaningful. So we shouldn't think that all of our imperfections really add much value to our lives.

Second, we have a great deal of control over the lives of our artificial creations. If we want to make artificial creatures that die, we can design them to do so. If we want to make artificial creatures with gambling problems, we can. If we want to make artificial creatures that strive to find romantic love and fail most of the time, due to their own neurotic insecurities, that is, by assumption, our choice.

If we can actually choose the imperfections of our creations, then we can do so in a way that is most beneficial to them. It is unlikely that our particular array of limitations is conducive to maximal well-being. It is unlikely that the optimal life-span for intelligent creatures happens to be how long human beings live. So we could design our progeny to live as long as it is good for a creature to live, and no longer. Plausibly, it would be better for us to have fewer kinds of pain and suffering, or have them more evenly distributed throughout the population and throughout our lives. Insofar as imperfection is good, artificial lives could be made to have just the perfect amount of imperfection. This should be far preferable to our present state.

Objection 2: Human beings can themselves be perfected.

Medical technology is rapidly improving. We are much better able to treat conditions now than we were a century ago. We may be much better still in another century. Instead of replacing ourselves with artificial creatures, we should

simply use technology to get rid of the detrimental traits we currently have. We can cure disease and forestall death inevitably through medical advances, and the more persistent psychological problems of humanity might also be curable through genetic manipulation.

Reply: It will most likely be easier to build new lives that avoid our imperfections than it would be to rid ourselves of those imperfections.

Making up for our species' problems requires working within a very specific set of constraints. Our bodies were not created to be easy to fix, so there is no guarantee that we could reach perfection through feasible alterations on this existing design nearly as easily or effectively as could be achieved starting from scratch.

It also remains to be seen how effective genetic tampering, social engineering, or psychiatric treatment can be. Genes don't nearly influence phenotypical traits. Who we are is largely determined by a set of extremely complex interactions between our genes and our environment. All this means that many of our imperfections may be extremely difficult or impossible to fully remove without creating larger problems elsewhere. Psychology is presently far from solving our emotional maladies, and we are making much more rapid progress toward artificial intelligence than we are to completely resolving our negative traits.

Even if we could create human beings with optimal lives in the same time frame in which we could create optimal artificial lives, we cannot get around the fact that our lives take a lot of resources to sustain. Humans have bodies of particular sizes that come along with particular caloric needs. Those needs have so far led us to convert a large percentage of the Earth's surface to agriculture. Our ideal lifestyle requires a lot of space and property for each individual. Sustaining a large human population has had catastrophic effects on our environment that an artificial population might avoid. If we could produce vastly more efficient artificial creatures that did not require organic material to live, then by allowing humanity to go extinct, we could make our present resources go a lot farther.

Objection 3: Human beings have value insofar as they contribute to diversity.

It is good that human beings exist, even if our lives are not as good as they can be, because we add to the diversity of the world. Imperfection can add to the value of the world even apart from contributing to the value of the lives bearing the imperfection.¹⁶ We should preserve humanity for the same reasons why we should preserve gorillas, giant pandas, and blue whales. These animals may not have lives that are as valuable as the average human being, but that doesn't mean that we should eradicate them and settle their former environments.

Reply: Diversity could be achieved in other ways.

Something of value would be lost if we eradicated the natural environment, and used all of Earth's resources to produce a maximal number of identical artificial creatures, no matter how good their lives. Whether or not it would be worthwhile to make the trade may depend on the specifics of the trade-off and fine details of a theory of value. However, it is possible for us to agree that homogeneity should be avoided without concluding that we must strive to keep human beings around as long as possible. Seeing value in the existence of a species need not mean that we should preserve that species at any cost. The cost of preserving humanity would plausibly involve a cost to diversity greater than the benefit it would achieve.

First, it may be possible to produce a greater amount of diversity artificially than naturally with a given allotment of resources. Human beings differ from each other, but mostly only in fairly constrained ways. We share a common physiological and psychological profile, including similar goals and interests, with each other. Artificial creatures, designed to our specifications, could have a far greater variety of forms. The same resources that would be devoted to keeping a large population of human beings alive could be used to produce more heterogeneous communities of artificial creatures.

It may be objected that not all kinds of diversity are valuable. Natural diversity is perhaps more valuable than artificial diversity. While I have trouble sympathizing with this response, it must be noted that humanity has had a hugely negative impact on natural diversity – in species and environments – in

¹⁶ Garland-Thomson, *op. cit.* note 14.

the world. We would be better off sustaining a large and well-off artificial population on some of the Earth's resources and returning the rest of the Earth to its natural state, than we would be on our present trajectory.

Third, even if we accept that diversity is a good thing, it may not be fair to create beings who are worse off in order to increase diversity.¹⁷ The fact that the existence of people with debilitating conditions is valuable in a population does not entail that we ought to preserve it, if it makes the lives of those with those conditions worse off, just as valuing disability does not entail that it is morally permissible to cause disabilities in nondisabled individuals.¹⁸

Finally, it is consistent with the spirit of the proposal that we could keep a token number of human beings around in small communities for diversity's sake. If there is anything inherently valuable in preserving a species, it is unlikely to matter whether we sustain more than a fairly token quantity.¹⁹ As a result, even if we have reason to keep the human species alive, we would only have reason to do so in trivial numbers. The vast majority of our resources could still be devoted to producing diverse artificial creatures with maximal well-being.

CONCLUSION

Many of the assumptions made in the course of the argument are likely to be all the more contentious because of what they collectively imply about the Artificial Replacement Thesis. The central argument in this paper may be read as a reductio of the assumptions upon which it is built. If the argument is successful given those assumptions, then skeptics should feel pressure to reject at least one of those assumptions.

What should they reject? I think that there are three natural places for skeptics to disagree with the argument.

First, it may be held that artificial minds cannot be created, or that we can never know whether they've been created. This is a very reasonable stance,

17 R. Sparrow. Imposing Genetic Diversity. *Am J Bioeth* 2015; 15 (6):2-10.

18 E. Barnes. Valuing Disability, Causing Disability. *Ethics* 2014; 125 (1):88-113.

19 T. Hurka. Value and Population Size. *Ethics* 1983; 93 (3):496-507.

especially in its second form,²⁰ but the skeptic who takes this way out leaves the moral question untouched. This is a precarious position for anyone who strongly disagrees with the conclusion, as an improved understanding of what makes our minds special, and how similar artificial intelligence may function, may force us to reevaluate.

Second, it could be maintained that our value cannot easily be replaced. I considered various reasons why we might think this is the case, and argued against them. But there might be something else that I did not sufficiently consider. If the assumptions about our future capabilities are accepted, then it is not easy to see what might be of value in us that cannot be replaced. One obvious possibility is naturalness. We cannot replicate our naturalness – interpreted as the product of undirected natural forces – in our artificial creations by definition, but elevating naturalness to a chief virtue is ad hoc.

Third, it might be held that we owe little to future generations in terms of moral considerations. This is an interesting view, and versions of it have been defended in response to the nonidentity problem. But it is an extreme view that has found few adherents.

Hesitance to accept the conclusion is understandable from creatures attached to their own existence. Such attachments, however, lead to well-known biases in favor of our sex, subculture, race, or species against others. Evaluating the argument fairly requires doing one's best to set such biases aside.

I think that we cannot ignore the possibility that we will be able to create artificial creatures with lives of optimal well-being in the not-too-distant future. If we can do that, a genuine utopia on Earth may be within our grasp. We must merely have the grace to step out of the way to let it happen.

20 J. Prinz. Level-headed Mysterianism and Artificial Experience. *J Conscious Stud* 2003; 10 (4-5):111-132.