

DR PREPBN
This paper summarizes a couple of the main arguments from my new book, *Humanity's End*. In the book I argue against radical enhancement – the adjustment of human attributes and abilities to levels that *greatly exceed* what is currently possible for human beings. I'm curious to see what reaction this elicits in a journal whose readership includes some of radical enhancement's most imaginative and committed advocates.

Humanity's End is motivated by the conviction that the debate about human enhancement must move beyond the binary “yes, I love it; no, it's evil” dialectic that has tended to dominate philosophical discussion up till now. When we focus on the multiple means – cybernetic, genetic, nanotechnological, and so on – by which humans are likely to be enhanced, we find significant moral differences. Some human enhancements should be endorsed; many should be rejected.

An additional motivation for *Humanity's End* is a sense that the technologies of human enhancement are on the verge of something really big. If Ray Kurzweil is right, then improvements of the information technologies that may be used to enhance human attributes track an exponential path (Kurzweil 2005). A feature of these ever-increasing patterns of improvement is that they deliver dramatic improvements quickly. Kurzweil's law of accelerating returns is controversial. Perhaps new means of enhancement won't arrive according to its schedule. But I know from my own nostalgia for rotary dial telephones and incredulity about computers that wirelessly access the Internet that human adaptation to technology lags behind the pace of technological change. We mustn't just assume a

gradualist scenario in which new human enhancements arrive in small increments with plenty of time for us to adjust between each instalment.

Rapid advances in the technologies of enhancement raise the possibility of radical enhancement which I define as the improvement of human attributes and abilities to levels that greatly exceed what is currently possible for human beings. *Humanity's End* presents an argument for rejecting this degree of enhancement.

Rejecting radical enhancement does not entail rejecting all the ways in which humans might enhance themselves. I defend a moderate stance on human enhancement. Some advocates of enhancement hear calls for moderation in much the same way that frat party attendees respond to calls for moderate consumption of beer. Indeed, there are many moral debates in which moderation is hard to achieve. For example, you either think that the state should be empowered to put condemned criminals to death or you think it shouldn't. It's difficult to work out what might count as a moderate position on this issue (arguing for the semi-execution of the nastiest criminals?). By contrast, a proper understanding of enhancement technologies reveals plenty of ground for moderation. I argue that our shared humanity imposes moral and prudential limits on enhancement. Many of the enhancements people most want are fully compatible with their humanity. They're valued by human beings precisely because of their promotion of enduring and significant human values. In matters of human enhancement, however, more is not always better. Proper scrutiny reveals radical enhancement to be incompatible with our humanity, and worth avoiding because of that fact. Slippery slope arguments to the effect that endorsing moderate enhancement entails endorsing radical enhancement warrant the same respect as proposals by party hosts that accepting offers of spritzers commits you to draining entire liquor cabinets.

So what would make our humanity incompatible with the machinations (literal and metaphorical) of advocates of radical enhancement? In *Humanity's End*, I identify humanity with the biological species *Homo sapiens*. According to Ernst Mayr's influential definition, biological species are "groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1963, 30). Racism may occasionally erect barriers between Malawians, Finns, Koreans, or Samoans. But these are only ever temporary and disappear together with the bigotry that generates them. The reproductive barrier between *Homo sapiens* and *Pan troglodytes* is, in contrast, no mere cultural artefact. Humans could mate with chimpanzees only under the most thoroughly artificial circumstances.

I propose that radical enhancement creates reproductive barriers in much the same way as would altering members of *Homo sapiens* to be genetically, physiologically, and psychologically indistinguishable from typical members of *Pan troglodytes*. The term "posthuman" is not just for show: it indicates a significant difference between the radically enhanced and the unenhanced. Human reproduction is about more than the sexual act. It's about having offspring that can be acknowledged as children and successfully raised to adulthood. While it certainly doesn't emerge as a matter of definition, radical enhancement is likely to isolate its recipients from humans. The act of giving yourself an intelligence that greatly exceeds that of Einstein is likely to make you less interested in the pillow talk of beings whose intelligence is inferior to Einstein's, and therefore in the activities that normally precede pillow talk. There's a good chance that any resulting offspring will be viewed as scientific curiosities and not as beloved children. This is no necessary truth. But it is likely. It's no temporary artefact of racist or speciesist ideology. I'm currently halfway through Peter F. Hamilton's hugely enjoyable space opera *The Dreaming Void*. I find that one of the less credible aspects of Hamilton's novels is the mutual interest and involvement he postulates between unimproved humans and beings who have undergone radical cognitive enhancement. It seems to me that the radically enhanced will get as much out of discussing their principal concerns with us as human literary critics get out of discussing Umberto Eco with chimpanzees.

In *Humanity's End*, I argue for a category of moral claims that track and respond to facts about species membership. According to species-relativism, certain experiences and ways of existing

properly valued by members of one species may lack value for the members of another species. Species-relativism is a version of moral relativism whose most widely discussed instance – cultural relativism – has come in for a great deal of hostility from philosophers. Cultural relativism presents morality as a cultural product. The prospects for species relativism are superior to those for cultural relativism simply because species differences matter more than cultural differences.

Species-relativism should not be confused with speciesism – the philosophically dubious doctrine that membership of a given species makes a difference to one’s moral worth (see Singer 1993).² Rather, it’s the view that certain kinds of valuable experience are more readily available to the members of some biological species than they are to non-members. It does not licence distinctions in moral worth on the basis of the capacity for these experiences.

I present species-relativist analyses of a variety of valuable human experiences. The successful pursuit of monumental intellects and millennial life spans makes these human experiences less accessible to us. To the extent that we value such experiences we’re right to reject radical enhancement.

So what, more precisely, are these valuable human experiences? I argue that some values are tied to our human limitations. Sometimes you have an instrumental interest in covering 42.195 kilometres as quickly as possible. Perhaps you’re couriering important documents. You’d be glad of the radical enhancement of your lungs and legs – or to have a friend with a Porsche. But sometimes you have an intrinsic interest in running the distance. You’re competing in a marathon. You thereby seek a connection with all other human marathoners, past and present. You’re less likely to achieve this connection with a posthuman whose cybernetic enhancements propel her over the distance in five minutes flat. I urge readers to consult the pages of *Humanity’s End* for less allusive descriptions of human values too fragile to survive radical enhancement.

Four riders of the Singularity

Humanity’s End uses the work of four prominent advocates of radical enhancement to explore the degree of enhancement that is compatible with our humanity. Investigation of these four reveals the perils of sending our brains and bodies in whichever direction technology can take them. The four focus on a variety of facets of radical enhancement. I certainly don’t have the completist ambitions of Morris Zapp, the character in David Lodge’s book *Trading Places* who aspires to say absolutely everything there could ever be said about Jane Austen. Many worthy advocates of radical human enhancement don’t feature in *Humanity’s End*. Furthermore, the book’s four focuses don’t agree about everything. They’re not like Marxist-Leninist conspirators who must march in lockstep or the revolution will fail.

Humanity’s End is not really about transhumanism. It’s true that some of radical enhancement’s most vigorous and persuasive advocates do identify as transhumanists – Nick Bostrom and James Hughes are two. But others don’t – Aubrey de Grey and Ray Kurzweil are examples. In any event, I’m a philosopher and there’s more philosophical value in ideas than in social movements.

The futurist and inventor Ray Kurzweil and the gerontologist Aubrey de Grey are principally interested in the technologies and therapies of radical enhancement (Kurzweil 1990, 2000, 2005; de Grey and Rae 2007). Kurzweil’s main focus is the human mind. He derives billion-fold increases in human intelligence from his law of accelerating returns. De Grey’s chief concern is the human body. He’s overseeing the search for therapies that he hopes will fix the damage aging inflicts on our bodies, thereby granting us millennial life spans.

While Kurzweil’s law of accelerating returns applies to all technologies, its implications for Artificial Intelligence have special significance for humans. According to Kurzweil, AI isn’t principally about making artificial things intelligent. It’s about making us artificially super-intelligent. The message from AI is that anything done by neurons can potentially be done better by electronic chips, without the attendant risk of Alzheimer’s. He predicts the transfer of our minds from fallible, disease prone

neurobiology to machines. The conversion of our minds into technology subjects it to the law of accelerating returns. It will swiftly transform the human mind into an intelligence that is “about one billion times more powerful than all human intelligence today” (Kurzweil 2005, 136).

Aubrey de Grey is a big-bearded dissident gerontologist who dares to deny the inevitability of aging. He is developing rejuvenation technologies that will, he believes, soon add instalments not of one or two years to human life spans but instead of centuries and millennia. De Grey’s aim is negligible senescence – an end to aging. According to him the multiple endogenous causes of human morbidity and mortality correspond to just seven things that can go wrong with cells and the relationships between them. He has testable theories about how to reverse these causes of aging. All he needs is (a lot of) cash. With sufficient financial support de Grey thinks that there’s a good chance that we can achieve millennial life spans within twenty five to thirty years.

Kurzweil and de Grey take charge of the “how to” of radical enhancement. Nick Bostrom and James Hughes address radical enhancement’s philosophical and social implications. Bostrom (2009) attempts a philosophical outflanking of radical enhancement’s opponents. According to his account, the very human values that opponents of enhancement claim to be defending actually call for dramatically enlarged intellects and extended life spans. A proper understanding of our human values makes radical enhancement urgent. Meanwhile, Hughes attends to the social realities of a society containing both radically enhanced and unenhanced citizens (Hughes 2004). He argues for a democratic transhumanism according to which the vast differentials in power between the unenhanced and the radically enhanced have no bearing on their moral worth. Humans, posthumans, and mountain gorillas differ in many ways. But they are all persons. Hughes proposes that democratic transhumanism can ensure a harmonious future for societies that comprise individuals who are making the transition from humanity to posthumanity at varying speeds, or not at all, by grounding moral and political status in the personhood that they all possess in equal measure.

These presentations of radical enhancement serve as different illustrations of the perils of immoderation. *Humanity’s End* tailors responses specific to their technological and philosophical details.

A precautionary approach

I recommend a precautionary approach to radical enhancement. This precautionary approach is not to be conflated with the (justly) infamous precautionary principle, which burdens advocates of technological change with an essentially unsatisfiable obligation to prove that their proposed changes could lead to no harm to us or to the environment. Extreme versions of the principle would have denied us incandescent light bulbs and antibiotics (see Starr 2003). Opponents of the principle are surely right to reject its near exclusive focus on the potential downsides of change. No action – visiting your local doctor, scratching your nose, calmly adopting the lotus position – is entirely free of risk. Risk is to be managed not avoided.

Advocacy of human enhancement manifests a contrary defect. There’s a tendency to be mesmerized by the upsides of radical enhancement. Millennial life spans and monumental intellects do seem to sell themselves. In some moods, advocates of radical enhancement refuse to take their opponents seriously. One riposte made by advocates of radical life extension is to accuse their opponents of suicidal urges and to complain that they shouldn’t foist their loathing of life on others. They act as if declaring “I love life” settles the dispute about the desirability of radical life extension. If opponents of cognitive enhancement are serious then why aren’t they taking to their heads with mallets to trim any surplus IQ points imposed on them by nature? The precautionary approach taken in *Humanity’s End* is motivated by the recognition that if a deal sounds too good to be true it might be.

A necessary first step in a decision-theoretic evaluation is to represent all of a choice’s possible outcomes, both positive and negative. We should attempt the same in respect of radical enhancement even if we cannot undertake the next step in decision-theoretic analysis of assigning specific

probabilities and utilities to each of these outcomes. We're left with an informal analogue of this exercise. When adopting the precautionary approach we should be aware of work done by psychologists on the glitches in human rationality that lead us to overlook certain possible outcomes and overemphasize others.

Consider one recent mis- or non-application of the approach. In his political memoir, *A Journey*, Tony Blair insists that the less-than-perfect outcome of the invasion of Iraq could not have been predicted. He's surely right. However, recent accounts of the deliberative process leading up to war do seem to expose biases. Strategists were overly focused on the possibility that Saddam's dethroning would lead to a swift installation of democracy with corresponding consequences for neighboring dictatorships. This is something that might have happened. It would have been excessively pessimistic to dismiss it out of hand. But the optimistic scenario must be balanced with more pessimistic scenarios. It's hard to believe that adequate consideration was given to the possibility that some Iraqi people wouldn't view the foreign troops as liberators and that the invasion might encourage, rather than preventing, acts of terrorism.

Advocates of radical enhancement are much smarter than Donald Rumsfeld. But I do think that – notwithstanding – they're mesmerized by the potential upsides of indefinite life spans and intellects a billion times more powerful than the combined intelligence of all of Earth's early twenty-first century inhabitants.

In *Humanity's End*, I seek a certain kind of engagement with Kurzweil, de Grey, Bostrom, Hughes, and others. There's no shortage of opponents of human enhancement. Some of them have made important criticisms. But the dialogue between opponents and defenders has thus far not been particularly productive. Their arguments have altogether too many unshared premises. Defenders of life extension and enlarged minds suspect opponents of artfully concealing appeals to God. Would-be radical enhancers need opponents who challenge the details of their proposals rather than just expressing outrage at the very ideas of dramatically longer lives and much bigger brains. *Humanity's End* seeks to satisfy that need.

Are human-posthuman societies viable?

One example of precautionary reasoning applies to the societies made by radical enhancement. Hollywood gives frequent expression to fears about the viability of societies comprising unenhanced humans and enhanced posthumans. James Hughes' democratic transhumanism addresses these concerns (Hughes 2004).

I myself find democratic transhumanism to be an attractive view. If Hughes is collecting signatures for a petition on future political arrangements then he has mine. But that's not really the point. Truths about moral status offer protection only to the extent that they are incorporated into a society's dominant moral code. A society's dominant moral code is the collection of moral ideas and principles that guides behaviour in that society. It stands behind and justifies the society's justice system and the actions of its public officials. The patent moral wrongness of the Nazi genocide did little to protect its victims because the equal moral worth of humans did not find adequate expression in the dominant moral code of Hitler's Germany. The Nazis' enemies went to their deaths with the small consolation that they were morally blameless and their executioners morally deplorable.

How confident should we be that democratic transhumanism, or a view like it, will either constitute or significantly contribute to the dominant moral codes of human-posthuman societies? Answering this question requires us to predict the moral views of beings with radically enhanced intellects. We certainly shouldn't assume that these views will be identical to our own. According to current thinking, moral truths are partially determined by the natures of the beings to whom they apply (see for example, Smith 1994). So it's entirely possible that the different natures of posthumans will generate moral truths that differ from our own.

Pundits of SETI – the Search for Extraterrestrial Intelligence – speculate about the possible moral views of aliens (see, for example, Davies 2010). It would be nice to think that aliens with technology sufficiently sophisticated to carry them across galaxies would think it wrong to use this technology to obliterate us. But it's difficult to work out what we could base our predictions of alien morality on. We're in a somewhat better position in respect of posthuman morality. Posthumans will emerge from us, meaning that posthuman moralities are likely to be developments of our ideas about right and wrong. In *Humanity's End*, I nominate the social contract view and moral consequentialism as possible precursors of the posthuman morality. My interest in these theories is not the philosopher's traditional one of seeking to determine whether either of them can capture the full truth about human morality. Rather, I use them to help predict posthuman morality.

Consider the variety of consequentialism defended by the influential Australian philosopher Peter Singer. Singer makes a powerful case for the wrongness of our treatment of many nonhuman animals. Factory farming causes much suffering without sufficient moral compensation. For example, while it's true that many people enjoy eating meat, they could be just as effectively and satisfyingly nourished by vegetarian diets. Singer's view could inspire hope in advocates of radical enhancement. The popularity of steaks and Chicken McNuggets suggests that Singer's moral outlook has had a limited impact on our culture's dominant moral code. But animal welfarists are working hard on its behalf. If Singer's consequentialism can protect the interests of less cognitively able sheep and cattle against humans then perhaps a posthuman consequentialism can protect humans against the demands of radically enhanced posthumans.

I'm not so confident about this. The adjudication of trade-offs is one of morality's most important social functions. In an ideal world, meat-eaters would get their delicious steaks and cattle would get to keep their hindquarters. Unfortunately, that's not the world we inhabit. When there are conflicts we look to a moral theory to tell us which interest should prevail. Singer and the CEO of McDonalds have different views on which interest deserves priority. The question is: how will the dominant moral codes of human-posthuman societies prioritize the different and sometimes competing claims of differently enhanced citizens? Singer's consequentialism finds many of the current reasons humans give for inflicting suffering on nonhumans – the filling of hamburgers and the testing of cosmetics, for example – morally insufficient. But consequentialists are quite emphatic about there being interests of cognitively superior beings that permit the sacrifice of the cognitively inferior. Singer's consequentialism asserts the (in principle) greater moral importance of beings with more numerous and diverse experiences and preferences (see Singer 1993, 105-109). If torturing animals really is the only way to achieve some wonderful human good – a cure for cancer, the end to all war – then he's all in favor. Singer's main demand in these circumstances, is that we abandon our squeamishness about torturing humans with cognitive abilities on par with nonhuman test subjects.

Our confidence about the position of humans in human-posthuman societies depends on a prediction that posthumans won't have a morally sufficient reason to sacrifice the vital interests of humans. I think there are likely to be such reasons.

Here's a forecast that has inductive support as strong as Kurzweil's prediction of exponential improvements of enhancement technologies. There's a tendency for the variety of uses we can make of animate and inanimate parts of our environment to increase as our intelligence increases. Human technological progress is largely about making previously unusable parts of our environment useable. A quick scan of the periodic table reveals many ways of using parts of our environment that didn't exist a hundred years ago. Kurzweil gives a particularly vivid presentation of the terminal stages of this tendency. The super-intelligence produced from human intelligence by way of the law of accelerating returns will, according to him, make optimal use of the computational powers of matter's fundamental particles, converting it all into mind. Says Kurzweil, "Ultimately, the entire universe will become saturated with our intelligence. This is the destiny of the universe" (2005, 29). If he's right then there'll soon be little or no matter and energy left over for the unimproved remnants of humanity.

This forecast doesn't postulate homicidal posthumans. Instead it requires morally motivated posthumans who, according to their moral views, correctly place their significant interests ahead of the vital interests of unimproved humans.

The key point is that there is a way out of this problem. We can't make it the case that neighboring galaxies don't contain super-intelligent aliens bent on our destruction. But it is within our powers to prevent the existence of posthumans whose moral codes permit or require actions unfriendly to humans. We can choose not to create them in the first place. If posthumans never come into existence then their moral needs and interests never exist. There's no moral requirement to bring any kind of being into existence. We're no more required to bring posthumans into existence than we are to create vampires or Terminator robots. My precautionary reasoning and interest in my human children leads me to believe that refraining to do so is the best option.

Is there a skeleton in my closet?

In 2004 I published a book with the title *Liberal Eugenics*. It was a defense of genetic enhancement. So what's a defender of enhancement doing turning around and attacking enhancement. Did I get religion?

Earlier in this piece, I suggested that the debate about human enhancement should mature beyond a simple duel between its opponents and defenders. A realistic, scientifically-informed presentation enables us to discriminate morally between different varieties and degrees of human enhancement. It reveals enhancement to be a way of treating human beings that can be good if practiced in moderation but dangerous if taken to extremes. Many of the influences humans direct at themselves fall into this category – drinking alcohol, exposure to direct sunlight, exercising, consuming saturated fats, and so on. Too much sun substantially elevates the risk of skin cancer. A moderate amount furnishes the body with requisite vitamin D. Alcoholism is a disease that destroys lives. But moderate drinking offers enjoyable experiences, promotes certain forms of sociability, and may reduce the risk of heart disease.

The cover of *Liberal Eugenics* was emblazoned with a couple of super-muscled male torsos that suggested a somewhat fascist version of the future made by genetic enhancement. But the picture defended in the book's pages was more nuanced and (I hope) less scary than that. *Liberal Eugenics* differs from defences that present enhancement as rationally or morally obligatory (see, for example, Savulescu 2001). I place genetic enhancement in the context of a liberal political philosophy and argue that prospective parents should have a constrained freedom to choose genetic enhancements that conform to their particular values. No one is rationally or morally required to enhance. Some people will have reproductive preferences analogous to the preference of supermarket shoppers for organic food. They'll want human procreation to be as natural as possible. Their perfectly legitimate concept of enhancement presents unmodified human embryos as best.

Jürgen Habermas attacks liberal eugenics on the grounds that, were we to accept it, "decisions regarding the genetic composition of children should not be submitted to any regulation by the state, but rather should be left to the parents" (Habermas 2003, 76). The state would be powerless to block or to effectively discourage eccentric or sadistic enhancement agendas. This is false. A major focus of liberal thought is on the limits of choice. Compare – only the most extreme defenders of the freedom of speech think that there should be no limits on what one can say. Most liberals find it perfectly legitimate for the state to prevent and punish incitements to violence in racially divided communities. One thing that we can predict about the future made by enhancement technologies is that enhanced beings will be coexisting with unmodified humans. Those responsible for formulating a society's policy on enhancement must attend to their interests.

It seems to me that radical enhancement lies beyond the limits of choice that we should be permitting in the early twenty first century. It threatens the welfare of those who make the morally innocent choice of remaining human.

Liberal Eugenics and *Humanity's End* are, in truth, different parts of a single picture of our species' future. Enhancements fully compatible with our humanity would permit us to reach Jeanne Calment's 122 years and to make scientific discoveries on par with Albert Einstein's. There's plenty of room between these achievements and what most of us achieve, or legitimately expect to achieve, to satisfy the vast majority of human desires. There's nothing intrinsically "yucky" about moderate enhancement. Beyond these levels of attainment lie not just additional years and IQ points but dangers unacknowledged by Kurzweil, de Grey, Bostrom, Hughes, *et al.* If Leon Kass's wise repugnance does find a proper application to human enhancement, it is here (Kass 1997).

Enhancement technologies will have (and are having) a powerfully transformative effect on our species. Even if we successfully restrict ourselves to moderate enhancements – those that we judge to be compatible with our humanity – we will change together with our values. It's possible that our enhanced descendents and future selves will no longer value remaining human. But that doesn't mean that we don't currently. My job is to ensure that these values receive due respect in the debate about human enhancement.

Notes

1. Cambridge, MA: MIT Press, 2010.
2. See Savulescu 2009 for a rebuttal of "the human prejudice" – the claim that humans are morally justified in treating other members of their species preferentially.

Acknowledgments

I'm grateful to Russell Blackford, Ramon Das, Simon Keller, César Palacios, and Dan Weijers for comments on this paper.

References

- Agar, N. 2010. *Humanity's end: Why we should reject radical enhancement*. Cambridge MA: MIT Press.
- Agar, N. 2004. *Liberal eugenics: In defence of human enhancement*. Oxford: Blackwell.
- Bostrom, N. 2009. Why I want to be a posthuman when I grow up. In *Medical enhancement and posthumanity*, ed. B. Gordijn and R. Chadwick, 107-136. Dordrecht: Springer.
- Davies, P. *The eerie silence: Are we alone in the universe?* London: Allen Lane.
- de Grey, A., and M. Rae. 2007. *Ending aging: The rejuvenation breakthroughs that could reverse human aging in our lifetime*. New York: St Martin's Press.
- Habermas, J. 2003. *The future of human nature*. Cambridge: Polity Press.
- Hughes, J. 2004 *Citizen cyborg: Why democratic societies must respond to the redesigned human of the future*. Cambridge MA: Westview.
- Kass, L. 1997. The wisdom of repugnance: Why we should ban the cloning of humans. *New Republic*, June 2.
- Kurzweil, R. 1990. *The age of intelligent machines*. Cambridge MA: MIT Press.

Kurzweil, R. 2000. *The age of spiritual machines: When computers exceed human intelligence*. London: Penguin.

Kurzweil, R. 2005. *The singularity is near: When humans transcend biology*. London: Penguin.

Mayr, E. 1963. *Animal species and evolution*. Cambridge MA: Harvard University Press.

Smith, M. 1994. *The moral problem*. Oxford: Blackwell.

Savulescu, J. 2001. Procreative beneficence: Why we should select the best children. *Bioethics* 15: 413-426.

Savulescu, J. 2009. The human prejudice and the moral status of enhanced beings: What do we owe the gods? In *Human Enhancement*, ed. J. Savulescu and N. Bostrom, 211-247. Oxford: Oxford University Press.

Singer, P. 1993. *Practical ethics*. Cambridge: Cambridge University Press.

Starr, S. 2003. Science, risk and the price of precaution: The scientific community imagines what society would have lost, had the "precautionary principle" governed science in the past. *Spiked Online* 1 May 2003. URL=<http://www.spiked-online.com/Articles/0600006DD7A.htm>.

DR. RUPNATHJI (DR. RUPAK NATH)