

15

WHY I WANT TO BE A POSTHUMAN WHEN I GROW UP¹

Nick Bostrom

CHAPTER SUMMARY

In this chapter, Nick Bostrom discusses the possibility that extreme human enhancement could result in “posthuman” modes of being. After offering some definitions and conceptual clarifications, he argues for two theses. First, there are posthuman modes of being – including some related to healthspan, cognition, and emotion – that would be very worthwhile. Second, it could be very good for human beings to become posthuman in those ways. He then considers and responds to objections to his theses, including several raised by the President’s Council on Bioethics in *Beyond Therapy* – for example, that personal identity could not be maintained through posthuman enhancement and that it constitutes a failure to be open to the gifted nature of life.

RELATED READINGS

Introduction: Form of Life (2.5); Autonomy, Authenticity, and Identity (2.6.4); Intrinsic Concerns (2.7)

Other Chapters: Walter Glannon, *Psychopharmacology and Functional Neurosurgery* (Ch. 12); President’s Council on Bioethics, *Beyond Therapy* (Ch. 14); Thomas Douglas, *Moral Enhancement* (Ch. 16)

1. SETTING THE STAGE

The term “posthuman” has been used in very different senses by different authors.² I am sympathetic to the view that the word often causes more confusion than clarity, and that we might be better off replacing it with some alternative vocabulary. However, as the purpose of this paper is not to propose terminological reform but to argue for certain substantial normative theses (which one would

¹ This chapter is excerpted from Nick Bostrom (2008) ‘Why I want to be a Posthuman When I Grow Up,’ in *Medical Enhancement and Posthumanity*, eds. B. Gordijn and R. Chadwick (Springer). It appears here by permission of Springer and the author.

² The definition used here follows in the spirit of Bostrom (2003). A completely different concept of “posthuman” is used in e.g. Hayles (1999).

naturally search for in the literature under the label “posthuman”), I will instead attempt to achieve intelligibility by clarifying the meaning that I shall assign to the word. Such terminological clarification is surely a minimum precondition for having a meaningful discussion about whether it might be good for us to become posthuman.

I shall define a posthuman as a being that has at least one posthuman capacity. By a posthuman capacity, I mean a general central capacity greatly exceeding the maximum attainable by any current human being without recourse to new technological means. I will use general central capacity to refer to the following:

healthspan – the capacity to remain fully healthy, active, and productive, both mentally and physically

cognition – general intellectual capacities, such as memory, deductive and analogical reasoning, and attention, as well as special faculties such as the capacity to understand and appreciate music, humor, eroticism, narration, spirituality, mathematics, etc.

emotion – the capacity to enjoy life and to respond with appropriate affect to life situations and other people

In limiting my list of general central capacities to these three, I do not mean to imply that no other capacity is of fundamental importance to human or posthuman beings. Nor do I claim that the three capacities in the list are sharply distinct or independent. Aspects of emotion and cognition, for instance, clearly overlap. But this short list may give at least a rough idea of what I mean when I speak of posthumans, adequate for present purposes.

In this paper, I will be advancing two main theses. The first is that some possible posthuman modes of being would be very good. I emphasize that the claim is not that *all* possible posthuman modes of being would be good. Just as some possible human modes of being are wretched and horrible, so too are some of the posthuman possibilities. Yet it would be of interest if we can show that there are some posthuman possibilities that would be very good. We might then, for example, specifically aim to realize those possibilities.

The second thesis is that it could be very good *for us* to become posthuman. It is possible to think that it could be good to be posthuman without it being good *for us* to become posthuman. This second thesis thus goes beyond the first. When I say “good for us”, I do not mean to insist that for every single current human individual there is some posthuman mode of being such that it would be good for that individual to become posthuman in that way. I confine myself to making a weaker claim that allows for exceptions. The claim is that for *most* current human beings, there are possible posthuman modes of being such that it could be good for these humans to become posthuman in one of those ways.

I am setting aside issues of feasibility, costs, risks, side-effects, and social consequences. While those issues are obviously important when considering what we have most reason to do all things considered, they will not be addressed here.

In the next three sections we will look in a little more detail at each of the three general central capacities that I listed in the introduction section. I hope to show that the claim that it could be very good to be posthuman is not as radical as it might appear to some. In fact, we will find that individuals and society already in some ways seem to be implicitly placing a very high value on posthuman

capacities – or at least, there are strong and widely accepted tendencies pointing that way. I therefore do not regard my claim as in any strong sense revisionary. On the contrary, I believe that the denial of my claim would be strongly revisionary in that it would force us to reject many commonly accepted ethical beliefs and approved behaviors. I see my position as a conservative extension of traditional ethics and values to accommodate the possibility of human enhancement through technological means.

2. HEALTHSPAN

It seems to me fairly obvious why one might have reason to desire to become a posthuman in the sense of having a greatly enhanced capacity to stay alive and stay healthy.³ I suspect that the majority of humankind already has such a desire implicitly.

People seek to extend their healthspan, i.e. to remain healthy, active, and productive. This is one reason why we install air bags in cars. It may also explain why we go to the doctor when we are sick, why higher salaries need to be paid to get workers to do physically dangerous work, and why governments and charities give money to medical research.⁴ Instances of individuals sacrificing their lives for the sake of some other goal, whether suicide bombers, martyrs, or drug addicts, attract our attention precisely because their behavior is unusual. Heroic rescue workers who endanger their lives on a dangerous mission are admired because we assume that they are putting at risk something that most people would be very reluctant to risk, their own survival.

For some three decades, economists have attempted to estimate individuals' preferences over mortality and morbidity risk in labor and product markets. While the tradeoff estimates vary considerably between studies, one recent meta-analysis puts the median value of the value of a statistical life for prime-aged workers to about \$7 million in the United States (Viscusi and Aldy, 2003). A study by the EU's Environment Directorates-General recommends the use of a value in the interval €0.9 to €3.5 million (Johansson, 2002). Recent studies by health economists indicate that improvements in the health status of the U.S. population over the 20th century have made as large a contribution to raising the standards of living as all other forms of consumption growth combined (Murphy and Topel, 2003; Nordhaus, 2003). While the exact numbers are debatable, there is little doubt that most people place a very high value on their continued existence in a healthy state.

Admittedly, a desire to extend one's healthspan is not necessarily a desire to become posthuman. To become posthuman by virtue of healthspan extension, one would need to achieve the capacity for a healthspan that greatly exceeds

³ Having such a capacity is compatible with also having the capacity to die at any desired age. One might thus desire a capacity for greatly extended healthspan even if one doubts that one would wish to live for more than, say, 80 years. A posthuman healthspan capacity would give one the option of much longer and healthier life, but one could at any point decide no longer to exercise the capacity.

⁴ Although on the last item, see Hanson (2000) for an alternative view.

the maximum attainable by any current human being without recourse to new technological means. Since at least some human beings already manage to remain quite healthy, active, and productive until the age of 70, one would need to desire that one's healthspan were extended greatly beyond this age in order that it would count as having a desire to become posthuman.⁵

Many people will, if asked about how long they would wish their lives to be, name a figure between 85 and 90 years (Cohen and Langer, 2005). In many cases, no doubt, this is because they assume that a life significantly longer than that would be marred by deteriorating health – a factor from which we must abstract when considering the desirability of healthspan extension. People's stated willingness to pay to extend their life by a certain amount does in fact depend strongly on the health status and quality of that extra life (Johnson et al., 1998). Since life beyond 85 is very often beset by deteriorating health, it is possible that this figure substantially underestimates how long most people would wish to live if they could be guaranteed perfect health.

It is also possible that a stated preference for a certain lifespan is hypocritical. Estimates based on revealed preferences in actual market choices, such as fatality risk premiums in labor markets or willingness to pay for health care and other forms of fatality risk reduction might be more reliable. It would be interesting to know what fraction of those who claim to have no desire for healthspan extension would change their tune if they were ever actually handed a pill that would reliably achieve this effect. My conjecture would be that when presented with a real-world choice, most would choose the path of prolonged life, health, and youthful vigor over the default route of aging, disease, and death.

One survey asked: "Based on your own expectations of what old age is like, if it were up to you, how long would you personally like to live – to what age?" Only 27% of respondents said they would like to live to 100 or older (Cohen and Langer, 2005). A later question in the same survey asked: "Imagine you could live to 100 or older, but you'd have to be very careful about your diet, exercise regularly, not smoke, avoid alcohol, and avoid stress. Would it be worth it, or not?" To this, 64% answered in the affirmative! Why should *more* people want to live beyond 100 when restrictions on activity are imposed? Is it because it frames the question more as if it were a real practical choice rather than as an idle mind game? Perhaps when the question is framed as a mind game, respondents tend to answer in ways which they believe expresses culturally approved attitudes, or which they think signal socially desirable personal traits (such as having "come to terms" with one's own mortality), while this tendency is diminished when the framing suggests a practical choice with real consequences. We do not know for sure, but this kind of anomaly suggests that we should not take people's stated "preferences" about how long they would wish to live too seriously, and that revealed preferences might be a more reliable index of their guiding values.

⁵ At least one human, Jeanne Calment, lived to 122. But although she remained in relatively fair health until close to her death, she clearly suffered substantial decline in her physical (and presumably mental) vigor compared to when she was in her twenties. She did not retain the capacity to be *fully* healthy, active, and productive for 122 years.

It is also worth noting that only a small fraction of us commit suicide, suggesting that our desire to live is almost always stronger than our desire to die.⁶ Our desire to live, *conditional on our being able to enjoy full health*, is even stronger. This presumption in favor of life is in fact so strong that if somebody wishes to die soon, even though they are seemingly fully healthy, with a long remaining healthy life expectancy, and if their external circumstances in life are not catastrophically wretched, we would often tend suspect that they might be suffering from depression or other mental pathology. Suicidal ideation is listed as a diagnostic symptom of depression by the American Psychiatric Association.⁷

Even if a stated preference against healthspan extension were sincere, we would need to question how well-considered and informed it is. It is of relevance that those who know most about the situation and are most directly affected by the choice, namely the elderly, usually prefer life to death. They usually do so when their health is poor, and overwhelmingly choose life when their health is at least fair. Now one can argue that a mentally intact 90-year-old is in a better position to judge how their life would be affected by living for another year than she was when she was 20, or 40. If most healthy and mentally intact 90-year-olds prefer to live for another year (at least if they could be guaranteed that this extra year would be one of full health and vigor), this would be evidence against the claim that it would be better for these people that their lives end at 90.⁸ Similarly, of course, for people of even older age.

One can compare this situation with the different case of somebody becoming paraplegic. Many able-bodied people believe that their lives would not be worth living if they became paraplegic. They claim that they would prefer to die rather than continuing life in a paraplegic state. Most people who have actually become paraplegic, however, find that their lives are worth living.⁹ People who are paraplegic are typically better judges of whether paraplegic lives are worth continuing than are people who have never experienced what it is like to be paraplegic. Similarly, people who are 90 years old are in a better position to judge whether their lives are worth continuing than are younger people (including themselves at any earlier point in their lives).¹⁰

⁶ For some, the reluctance to commit suicide might reflect a desire not to kill oneself rather than a desire not to die, or alternatively a fear of death rather than an authentic preference not to die.

⁷ DSM-IV (American Psychiatric Association, 2000).

⁸ This is a kind of Millian best-judge argument. However, if fear of death were irrational, one could argue that people who are closer to death are on average worse judges of the value for them of an extra year of life, because their judgments would tend to be more affected by irrational fear.

⁹ This basic result is reflected in many chronic disease conditions (Ubel et al., 2003). The discrepancy of attitudes seems to be due to non-patient's failure to realize the extent to which patients psychologically adapt to their condition (Damschroder et al., 2005).

¹⁰ The analogy with paraplegia is imperfect in at least one respect: when the issue is healthspan extension, we are considering whether it would be worth living an extended life in perfect health and vigor. If anything, this discrepancy strengthens the conclusion, since it is more worth continuing living in perfect health than in poor health, not less worth it.

One study assessed the will to live among 414 hospitalized patients aged 80 to 98 years, presumably representing the frailer end of the distribution of the “old old”. 40.8% of respondents were unwilling to exchange any time in their current state of health for a shorter life in excellent health, and 27.8% were willing to give up at most 1 month of 12 in return for excellent health.¹¹ (Patients who were still alive one year later were even less inclined to give up life for better health, but with continued large individual variations in preferences.) The study also found that patients were willing to trade significantly less time for a healthy life than their surrogates assumed they would.

Research shows that life-satisfaction remains relatively stable into old age. One survey of 60,000 adults from 40 nations discovered a slight upward trend in life-satisfaction from the 20s to the 80s in age (Diener and Suh, 1998). Life satisfaction showed this upward trend even though there was some loss of positive affect. Perhaps life-satisfaction would be even higher if positive affect were improved (a possibility we shall discuss in a later section). Another study, using a cross-sectional sample (age range 70–103 years), found that controlling for functional health constraints reversed the direction of the relationship between age and positive affect and produced a negative association between age and negative affect (Kunzmann et al., 2000). These findings suggest that some dimensions of subjective well-being, such as life-satisfaction, do not decline with age but might actually increase somewhat, and that the decline in another dimension of subjective well-being (positive affect) is not due to aging per se but to health constraints.

Most people reveal through their behavior that they desire continued life and health,¹² and most of those who are in the best position to judge the value of continued healthy life, at any age, judge that it is worth having. This constitutes *prima facie* support for the claim that extended life is worth having even when it is not fully healthy. The fact that this holds true at all currently realized ages suggests that it is not a strongly revisionary view to hold that it could be good for many people to become posthuman through healthspan extension. Such a view might already be implicitly endorsed by many.

3. COGNITION

People also seem to be keen on improving cognition. Who wouldn't want to remember names and faces better, to be able more quickly to grasp difficult abstract ideas, and to be able to “see connections” better? Who would seriously object to being able to appreciate music at a deeper level? The value of optimal cognitive functioning is so obvious that to elaborate the point may be unnecessary.¹³

¹¹ Tsevat et al. (1998). See also McShine et al. (2000). For a methodological critique, see Arnesen and Norheim (2003).

¹² This is fully consistent with the fact that many people knowingly engage in risky behaviors such as smoking. This might simply mean that they are unable to quit smoking, or that they desire the pleasure of smoking more than they desire a longer healthier life. It does not imply that they do not desire longer healthier life.

This verdict is reflected in the vast resources that society allocates to education, which often explicitly aims not only to impart specific items of knowledge but also to improve general reasoning abilities, study skills, critical thinking, and problem solving capacity.¹⁴ Many people are also keen to develop various particular talents that they may happen to have, for example musical or mathematical, or to develop other specific faculties such as aesthetic appreciation, narration, humor, eroticism, spirituality etc. We also reveal our desire for improving our cognitive functioning when take a cup of coffee to increase our alertness or when we regret our failure to obtain a full night's sleep because of the detrimental effects on our intellectual performance.

Again, the fact that there is a common desire for cognitive improvement does not imply that there is a common desire for becoming posthuman. To want to become posthuman through cognitive improvement, one would have to want a great deal of cognitive improvement. It is logically possible that each person would only want to become slightly more intelligent (or musical, or humorous) than he or she currently is and would not want any very large gain. I will offer two considerations regarding this possibility.

First, it seems to me (based on anecdotal evidence and personal observations) that people who are already endowed with above-average cognitive capacities are at least as eager, and, from what I can tell, actually *more* eager to obtain further improvements in these capacities than are people who are less talented in these regards. For instance, someone who is musically gifted is likely to spend more time and effort trying to further develop her musical capacities than is somebody who lacks a musical ear; and likewise for other kinds of cognitive gifts.

This phenomenon may in part reflect the external rewards that often accrue to those who excel in some particular domain. An extremely gifted musician might reap greater rewards in terms of money and esteem from a slight further improvement in her musicality than would somebody who is not musically gifted to begin with. That is, the difference in external rewards is sometimes greater for somebody who goes from very high capacity to outstandingly high capacity than it is for somebody who goes from average capacity to moderately high capacity. However, I would speculate that such differences in external rewards are only part of the explanation and that people who have high cognitive capacities are usually also more likely (or at least no less likely) to desire further increases in those capacities than are people of lower cognitive capacities even when only the intrinsic benefits of capacities are considered. Thus, if we imagine a group of people placed in solitary confinement for the remainder of their lives, but with access to books, musical instruments, paints and canvasses, and other prerequisites for the exercise of capacities, I would hypothesize that those with the highest pre-existing capacity in a given domain would be more likely (or at least not less likely) to work hard to further develop their capacities in that domain, for the sake of the intrinsic benefits that the possession and exercise of those capacities bestow, than would those with lower pre-existing capacities

¹³ One might even argue that a desire for cognitive improvement is a constitutive element of human rationality, but I will not explore that hypothesis here.

¹⁴ U.S. public expenditure on education in 2003 was 5.7% of its GDP (World Bank, 2003).

in the same domain.¹⁵ While \$100 brings vastly less utility to a millionaire than to a pauper, the marginal utility of improved cognitive capacities does not seem to exhibit a similar decline.

These considerations suggest that there are continuing returns in the “intrinsic” (in the sense of non-instrumental, non-positional) utility of gains in cognitive capacities, at least within the range of capacity that we find instantiated within the current human population.¹⁶ It would be implausible to suppose that the current range of human capacity, in all domains, is such that while increments of capacity within this range are intrinsically rewarding, yet any further increases outside the current human range would lack intrinsic value. Again, we have a *prima facie* reason for concluding that enhancement of cognitive capacity to the highest current human level, and probably beyond that, perhaps up to and including the posthuman level, would be intrinsically desirable for the enhanced individuals. We get this conclusion if we assume that those who have a certain high capacity are generally better judges of the value of having that capacity or of a further increment of that capacity than are those who do not possess the capacity in question to the same degree.

4. EMOTION

It is straightforward to determine what would count as an enhancement of healthspan. We have a clear enough idea of what it means to be healthy, active, and productive, and the difference between this state and that of being sick, incapacitated, or dead. An enhancement of healthspan is simply an intervention that prolongs the duration of the former state. It is more difficult to define precisely what would count as a cognitive enhancement because the measure of cognitive functioning is more multifaceted, various cognitive capacities can interact in complex ways, and it is a more normatively complex problem to determine what combinations of particular cognitive competences are of value in different kinds of environments. For instance, it is not obvious what degree of tendency to forget certain kinds of facts and experiences is desirable. The answer might depend on a host of contextual factors. Nevertheless, we do have some general idea of how we might value various increments or decrements in many aspects of our cognitive functioning – a sufficiently clear idea, I suggest, to make it intelligible without much explanation what one might mean by phrases like “enhancing musical ability”, “enhancing abstract reasoning ability,” etc.

It is considerably more difficult to characterize what would count as emotional enhancement. Some instances are relatively straightforward. Most would readily

¹⁵ Complication: if high capacity were solely a result from having spent a lot of effort in developing that capacity, then the people with high capacity in some domain might be precisely those that started out having an unusually strong desire for having a strong capacity in that domain. It would then not be surprising that those with high capacity would have the strongest desire for further increases in capacity. Their stronger desire for higher capacity might then not be the result of more information and better acquaintance with what is at stake, but might instead simply reflect a prior inclination.

¹⁶ It would be more difficult to determine whether the marginal intrinsic utility of gains in capacity are constant, or diminishing, or increasing at higher levels of capacity, and if so by what amount.

agree that helping a person who suffers from persistent suicidal depression as the result of a simple neurochemical imbalance so that she once again becomes capable of enjoyment and of taking an interest in life would be to help her improve her emotional capacities. Yet beyond cases involving therapeutic interventions to cure evident psychopathology it is less clear what would count as an enhancement. One's assessment of such cases often depends on the exact nature of one's normative beliefs about different kinds of possible emotional constitutions and personalities.

It is correspondingly difficult to say what would constitute a "posthuman" level of emotional capacity. Nevertheless, people often do strive to improve their emotional capacities and functionings. We may seek to reduce feelings of hate, contempt, or aggression when we consciously recognize that these feelings are prejudiced or unconstructive. We may take up meditation or physical exercise to achieve greater calm and composure. We may train ourselves to respond more sensitively and empathetically to those we deem deserving of our trust and affection. We may try to overcome fears and phobias that we recognize as irrational, or we may wrestle with appetites that threaten to distract us from what we value more. Many of us expend life-long effort to educate and ennoble our sentiments, to build our character, and to try to become better people. Through these strivings, we seek to achieve goals involving modifying and improving our emotional capacities.

An appropriate conception of emotional capacity would be one that incorporates or reflects these kinds of goal, while allowing perhaps for there being a wide range of different ways of instantiating "high emotional capacity", that is to say, many different possible "characters" or combinations of propensities for feeling and reacting that could each count as excellent in its own way. If this is admitted, then we could make sense of emotional enhancement in a wide range of contexts, as being that which makes our emotional characters more excellent. A posthuman emotional capacity would be one which is much more excellent than that which any current human could achieve unaided by new technology.

One might perhaps question whether there are possible emotional capacities that would be *much* more excellent than those attainable now. Conceivably, there might be a maximum of possible excellence of emotional capacity, and those people who currently have the best emotional capacities might approach so closely to this ideal that there is not enough potential left for improvement to leave room for a posthuman realm of emotional capacity. I doubt this, because aside from the potential for fine-tuning and balancing the various emotional sensibilities we already have, I think there might also be entirely new psychological states and emotions that our species has not evolved the neurological machinery to experience, and some of these sensibilities might be ones we would recognize as extremely valuable if we became acquainted with them.

It is difficult intuitively to understand what such novel emotions and mental states might be like. This is unsurprising, since by assumption we currently lack the required neurological bases. It might help to consider a parallel case from within the normal range of human experience. The experience of romantic love is something that many of us place a high value on. Yet it is notoriously difficult for a child or a prepubescent teenager to comprehend the meaning of romantic love or why adults should make so much fuss about this experience. Perhaps we are all currently

in the situation of children relative to the emotions, passions, and mental states that posthuman beings could experience. We may have no idea of what we are missing out on until we attain posthuman emotional capacities.

One dimension of emotional capacity that we can imagine enhanced is subjective well-being and its various flavors: joy, comfort, sensual pleasures, fun, positive interest and excitement. Hedonists claim that pleasure is the only intrinsic good, but one need not be a hedonist to appreciate pleasure as one important component of the good. The difference between a bleak, cold, horrid painful world and one that is teeming with fun and exciting opportunities, full of delightful quirks and lovely sensations, is often simply a difference in the hedonic tone of the observer. Much depends on that one parameter.

It is an interesting question how much subjective well-being could be enhanced without sacrificing other capacities that we may value. For human beings as we are currently constituted, there is perhaps an upper limit to the degree of subjective well-being that we can experience without succumbing to mania or some other mental unbalance that would prevent us from fully engaging with the world if the state were indefinitely prolonged. But it might be possible for differently constituted minds to have experiences more blissful than those that humans are capable of without thereby impairing their ability to respond adequately to their surroundings. Maybe for such beings, gradients of pleasure could play a role analogous to that which the scale ranging between pleasure and pain has for us (Pearce, 2004). When thinking the possibility of *posthumanly happy* beings, and their psychological properties, one must abstract from contingent features of the human psyche. An experience that would consume us might perhaps be merely “spicy” to a posthuman mind.

It is not necessary here to take a firm stand on whether posthuman levels of pleasure are possible, or even on whether posthuman emotional capacities more generally are possible. But we can be confident that, at least, there is vast scope for improvements for most of individuals in these dimensions because even within the range instantiated by currently existing humans, there are levels of emotional capacities and degrees of subjective well-being that, for most of us, are practically unattainable to the point of exceeding our dreams. The fact that such improvements are eagerly sought by many suggests that if posthuman levels were possible, they too would be viewed as highly attractive.¹⁷

5. STRUCTURE OF THE ARGUMENT AND FURTHER SUPPORTING REASONS

It might be useful to pause briefly to reflect on the structure of the argument presented so far. I began by listing three general central capacities (healthspan, cognition, and emotion), and I defined a posthuman being as one who has at

¹⁷ The quest for subjective well-being, in particular, seems to be a powerful motivator for billions of people even though arguably none of the various means that have been attempted in this quest has yet proved very efficacious in securing the goal (Brickman and Campbell, 1971).

least one of these capacities in a degree unattainable by any current human being unaided by new technology.

I offered some plausibility arguments suggesting that it could be highly desirable to have posthuman levels of these capacities. I did this partly by clarifying what having the capacities would encompass and by explaining how some possible objections would not apply because they rely on a misunderstanding of what is proposed. Furthermore, I tried to show that for each of the three capacities we find that many individuals actually desire to develop the capacities to higher levels and often undertake great effort and expense to achieve these aims. This desire is also reflected in social spending priorities, which devote significant resources to e.g. healthspan-extending medicine and cognition-improving education. Significantly, at least in the cases of healthspan extension and cognitive improvement, the persons best placed to judge the value and desirability of incremental improvements at the high end of the contemporary human capacity distribution seem to be especially likely to affirm the desirability of such additional improvements of capacity. For many cognitive faculties, it appears that the marginal utility of improvements *increases* with capacity levels. This suggests that improvements beyond the current human range would also be viewed as desirable when evaluated by beings in a better position to judge than we currently are.

6. PERSONAL IDENTITY

Supposing the previous sections have succeeded in making it plausible that being a posthuman could be good, we can now turn to a further question: whether becoming posthuman could be good *for us*. It may be good to be Joseph Haydn. Let us suppose that Joseph Haydn had a better life than Joe Bloggs so that in some sense it is better to be Haydn and living the life that Haydn lived than to be Bloggs and living Bloggs' life. We may further suppose that this is so from Bloggs' evaluative standpoint. Bloggs might recognize that on all the objective criteria which he thinks makes for a better mode of being and a better life, Haydn's mode of being and life are better than his own. Yet it does not follow that it would be good for Bloggs to "become" Haydn (or to become some kind of future equivalent of Haydn) or to live Haydn's life (or a Haydn-like life). There are several possible reasons for this which we need to examine.

First, it might not be possible for Bloggs to become Haydn without ceasing to be Bloggs. While we can imagine a thought experiment in which Bloggs' body and mind are gradually transformed into those of Haydn (or of a Haydn-equivalent), it is not at all clear that personal identity could be preserved through such a transformation. If Bloggs' personal identity is essentially constituted by some core set of psychological features such as his memories and dispositions, then, since Haydn does not have these features, the person Bloggs could not become a Haydn-equivalent. Supposing that Bloggs has a life that is worth living, any transformation that causes the person Bloggs to cease to exist might be bad for Bloggs, including one that transforms him into Haydn.

Could a current human become posthuman while remaining the same person, or is the case like the one of Bloggs becoming Haydn, the person Bloggs necessarily ceasing to exist in the process? The case of becoming posthuman is

different in an important respect. Bloggs would have to lose all the psychological characteristics that made him person Bloggs in order to become Haydn. In particular, he would have to lose all his memories, his goals, his unique skills, and his entire personality would be obliterated and replaced by that of Haydn. By contrast, a human being could retain her memories, her goals, her unique skills, and many important aspects of her personality even as she becomes posthuman. This could make it possible for personal identity to be preserved during the transformation into posthuman.¹⁸

It is obvious that personal identity could be preserved, at least in the short run, if posthuman status is achieved through radical healthspan enhancement. Suppose that I learnt that tonight after I go to bed, a scientist will perform some kind of molecular therapy on my cells while I'm sleeping to permanently disable the aging processes in my body. I might worry that I would not wake up tomorrow because the surgery might go wrong. I would not worry that I might not wake up tomorrow because the surgery succeeded. Healthspan enhancement would help preserve my personal identity.

The case that personal identity could be preserved is perhaps less clear-cut with regard to radical cognitive or emotional enhancement. Could a person become radically smarter, more musical, or come to possess much greater emotional capacities without ceasing to exist? Here the answer might depend more sensitively on precisely which changes we are envisaging, how those changes would be implemented, and on how the enhanced capacities would be used. The case for thinking that both personal identity and narrative identity would be preserved is arguably strongest if we posit that (a) the changes are in the form of addition of new capacities or enhancement of old ones, without sacrifice of preexisting capacities; and (b) the changes are implemented gradually over an extended period of time; (c) each step of the transformation process is freely and competently chosen by the subject; and (d) the new capacities do not prevent the preexisting capacities from being periodically exercised; (e) the subject retains her old memories and many of her basic desires and dispositions; (f) the subject retains many of her old personal relationships and social connections; and (g) the transformation fits into the life narrative and self-conception of the subject. Posthuman cognitive and emotional capacities could in principle be acquired in such a way that these conditions are satisfied.

Even if not all the conditions (a)–(g) were fully satisfied in some particular transformation process, the normatively relevant elements of a person's (numerical or narrative) identity could still be *sufficiently* preserved to avoid raising any fundamental identity-based objection to the prudence of undergoing such a transformation. We should not use a stricter standard for technological self-transformation than for other kinds of human transformation, such as migration, career change, or religious conversion.

¹⁸ See, also, DeGrazia (2005). DeGrazia argues that identity-related challenges to human enhancement largely fails, both ones based on considerations of personal identity and ones based on narrative identity (authenticity), although he mainly discusses more moderate enhancements than those I focus on in this paper.

Consider again a familiar case of *radical* human transformation: maturation. You currently possess vastly greater cognitive capacities than you did as an infant. You have also lost some capacities, e.g. the ability to learn to speak a new language without an accent. Your emotional capacities have also changed and developed considerably since your babyhood. For each concept of identity which we might think has relevant normative significance – personal (numerical) identity, narrative identity, identity of personal character, or identity of core characteristics – we should ask whether identity in that sense has been preserved in this transformation.

The answer may depend on exactly how we understand these ideas of identity. For each of them, on a sufficiently generous conception of the identity criteria, identity was completely or in large part preserved through your maturation. But then we would expect that identity in that sense would also be preserved in many other transformations, including the ones that are *no more profound* as that of a child growing into an adult; and this would include transformations that would make you posthuman. Alternatively, we might adopt conceptions that impose more stringent criteria for the preservation of identity. On these conceptions, it might be impossible to become posthuman without wholly or in large part disrupting one form of identity or another. However, on such restrictive conceptions, identity would also be disrupted in the transformation of child into adult. Yet we do not think it is bad for a child to grow up. Disruptions of identity in those stringent senses form part of a normal life experience and they do not constitute a disaster, or a misfortune of any kind, for the individual concerned.

Why then should it be bad for a person to continue to develop so that she one day matures into a being with posthuman capacities? Surely it is the other way around. If this had been our usual path of development, we would have easily recognized the failure to develop into a posthuman as a misfortune, just as we now see it as a misfortune for a child to fail to develop normal adult capacities.

Many people who hold religious beliefs are already accustomed to the prospect of an extremely radical transformation into a kind of posthuman being, which is expected to take place after the termination of their current physical incarnation. Most of those who hold such a view also hold that the transformation *could* be very good for the person who is transformed.

7. BRIEF SKETCHES OF SOME OBJECTIONS AND REPLIES

Objection: One might think that it would be bad for a person to be the only posthuman being since a solitary posthuman would not have any equals to interact with.

Reply: It is not necessary that there be only one posthuman.

Objection: The accumulated cultural treasures of humanity might lose their appeal to somebody whose capacities greatly exceeded those of the humans who produced them. More generally, challenges that seemed interesting to the person while she was still human might become trivial and therefore uninteresting to her when she acquires posthuman capacities. This could deprive posthumans of the good of meaningful achievements.

Reply: It is not clear why the ability to appreciate what is more complex or subtle should make it impossible to appreciate simpler things. Somebody who has

learnt to appreciate Schoenberg may still delight in simple folk songs, even bird songs. A fan of Cézanne may still enjoy watching a sunrise.

Even if it were impossible for posthuman beings to appreciate some simple things, they could compensate by creating new cultural riches. If some challenges become too easy for posthumans, they could take on more difficult challenges. One might argue that an additional reason for developing posthuman cognitive capacities is that it would increase the range of interesting intellectual challenges open to us. At least within the human range of cognitive capacity, it seems that the greater one's capacity, the more numerous and meaningful the intellectual projects that one can embark on. When one's mind grows, not only does one get better at solving intellectual problems – entirely new possibilities of meaning and creative endeavor come into view.

Objection: A sense of vulnerability, dependence, and limitedness can sometimes add to the value of a life or help a human being grow as a person, especially along moral or spiritual dimensions.

Reply: A posthuman could be vulnerable, dependent, and limited. A posthuman could also be able to grow as a person in moral and spiritual dimensions without those extrinsic spurs that sometimes necessary to affect such growth in humans. The ability to spontaneously develop in these dimensions could be seen as an aspect of emotional capacity.

Objection: The very desire to overcome one's limits by the use of technological means rather than through one's own efforts and hard work could be seen as expressive of a failure to open oneself to the unbidden, gifted nature of life, or as a failure to accept oneself as one is, or as self-hate.

Reply: This paper makes no claims about the expressive significance of a desire to become posthuman, or about whether having such a desire marks one as a worse person, whether necessarily or statistically. The concern here rather is about whether being posthuman could be good, and whether it could be good for us to become posthuman.

Objection: A capacity obtained through a technological shortcut would not have the same value as one obtained through self-discipline and sacrifice.

Reply: I have argued that the possession of posthuman capacities could be extremely valuable even were the capacities effortlessly obtained. It is consistent with what I have said that achieving a capacity through a great expenditure of blood, sweat, and tears would further increase its value. I have not addressed what would be the *best* way of becoming posthuman. We may note, however, that is unlikely that we *could* in practice become posthuman other than via recourse to advanced technology.

Objection: The value of achieving a goal like winning a gold medal in the Olympics is reduced and perhaps annulled if the goal is achieved through inappropriate means (e.g. cheating). The value of possessing a capacity likewise depends on how the capacity was acquired. Even though having posthuman capacities might be extremely valuable if the capacities had been obtained by appropriate means, there are no humanly possible means that are appropriate. Any means by which humans could obtain posthuman capacities would negate the value of having such capacities.

Reply: The analogy with winning an Olympic medal is misleading. It is in the nature of sports competitions that the value of achievement is intimately connected

with the process by which it was achieved. We may say that what is at stake in the analogy is not really the value of a medal, nor even the value of winning a medal, but rather (something like) winning the medal by certain specified means in a fair competition, in a non-fluke-like way, etc. Many other goods are not like this. When we visit the doctor in the hope of getting well, we do not usually think that the value of getting well is strongly dependent on the process by which health is achieved; health and the enjoyment of health are valuable in their own right, independently of how these states come about. Of course, we are concerned with the value of the means to getting well – the means themselves can have negative value (involving perhaps pain and inconvenience), and in evaluating the value of the consequences of an action, we take the value of the means into account as well as the value of the goal that they achieve. But usually, the fact that some means have negative value does not reduce the value of obtaining the goal state. The values that I have alleged could be derived from posthuman capacities are not like the value of an Olympic gold medal, but rather like the value of health... I am aware of no logical, metaphysical, or “in principle” reason why humans could not obtain posthuman capacities in ways that would avoid recourse to immoral means of the sort that would “taint” the outcome.

8. CONCLUSION

I have argued, first, that some posthuman modes of being would be extremely worthwhile; and, second, that it could be good for most human beings to become posthuman.

I have discussed three general central capacities – healthspan, cognition, and emotion – separately for most of this paper. However, some of my arguments are strengthened if one considers the possibility of combining these enhancements. A longer healthspan is more valuable when one has the cognitive capacity to find virtually inexhaustible sources of meaning in creative endeavors and intellectual growth. Both healthspan and cognition are more valuable when one has the emotional capacity to relish being alive and to take pleasure in mental activity.

It follows trivially from the definition of “posthuman” given in this paper that we are not posthuman at the time of writing. It does not follow, at least not in any obvious way, that a posthuman could not also remain a human being. Whether or not this is so depends on what meaning we assign to the word “human”. One might well take an expansive view of what it means to be human, in which case “posthuman” is to be understood as denoting a certain possible type of human mode of being – if I am right, an exceedingly worthwhile type.

WORKS CITED

- American Psychiatric Association. (2000). *Diagnostic criteria from DSM-IV-TR*. (Washington, D.C.: American Psychiatric Association).

- T. M. Arnesen, and O. F. Norheim (2003) 'Quantifying quality of life for economic analysis: time out for time trade off', *Medical Humanities* 29 (2): 81–86.
- T. L. Beauchamp, and J. F. Childress (2001) *Principles of biomedical ethics* (New York, N.Y.: Oxford University Press).
- N. Bostrom (2003). "The Transhumanist FAQ: v 2.1." World Transhumanist Association. Available from <http://transhumanism.org/index.php/WTA/faq/>.
- N. Bostrom (2004). Transhumanist Values, in *Ethical Issues for the 21st Century*. F. Adams, ed., Philosophical Documentation Center Press.
- N. Bostrom (2005). "The Fable of the Dragon-Tyrant." *Journal of Medical Ethics* 31(5): 273–277.
- P. Brickman, and D. T. Campbell (1971). Hedonic relativism and planning the good society, in *Adaptation-level theory: A symposium*. M. H. Apley, ed. New York, Academic Press: 287–301.
- J. Cohen and G. Langer (2005). "Most Wish for a Longer Life – Despite Broad Aging Concerns." ABC News/USA Today Poll. Available from <http://abcnews.go.com/images/Politics/995a1Longevity.pdf>.
- L. J. Damschroder, B. J. Zikmund-Fisher, et al. (2005). "The impact of considering adaptation in health state valuation." *Social Science & Medicine* 61(2): 267–277.
- D. DeGrazia (2005). "Enhancement Technologies and Human Identity." *Journal of Medicine and Philosophy* 30: 261–283.
- E. Diener and E. M. Suh (1998). "Subjective well-being and age: An international analysis." *Annual Review of Gerontology and Geriatrics* 17: 304–324.
- W. Glannon (2002). "Identity, prudential concern, and extended lives." *Bioethics* 16(3): 266–83.
- J. B. S. Haldane (1924). *Daedalus; or, Science and the future*. London, K. Paul, Trench, Trubner & co., ltd.
- R. Hanson (2000). "Showing That You Care: The Evolution of Health Altruism." Available from <http://hanson.gmu.edu/showcare.pdf>.
- N. K. Hayles (1999). *How we became posthuman: virtual bodies in cybernetics, literature, and informatics*. Chicago, Ill., University of Chicago Press.
- P. O. Johansson (2002). "On the definition and age-dependency of the value of a statistical life." *Journal of Risk and Uncertainty* 25(3): 251–263.
- F. R. Johnson, W. H. Desvousges, et al. (1998). "Eliciting stated health preferences: An application to willingness to pay for longevity." *Medical Decision Making* 18(2): S57–S67.
- U. Kunzmann, T. Little, et al. (2000). "Is age-related stability of subjective well-being a paradox? Cross-sectional and longitudinal evidence from the Berlin Aging Study." *Psychology and Aging* 15(3): 511–426.
- D. Lewis (1989). "Dispositional Theories of Value." *Proceedings of the Aristotelian Society, supp.* 63: 113–137.
- R. McShine, G. T. Lesser, et al. (2000). "Older Americans hold on to life dearly." *British Medical Journal* 320(7243): 1206–1207.
- G. E. Moore (1903). *Principia Ethica*. Cambridge, Cambridge University Press.

- K. Murphy and R. Topel (2003). 'The Economic Value of Medical Research,' *Measuring the gains from medical research: an economic approach*. Chicago, University of Chicago Press.
- W. Nordhaus (2003). The Health of Nations: The Contribution of Improved Health to Living Standards. *Measuring the gains from medical research: an economic approach*, in K. M. Murphy and R. H. Topel, eds. Chicago, University of Chicago Press: vi, 263 p.
- D. Pearce (2004). "The Hedonistic Imperative." Available from <http://www.hedweb.com/hedab.htm>.
- M. Sandel (2004). "The Case Against Perfection." *The Atlantic Monthly* 293(3).
- J. Tsevat, N. V. Dawson, et al. (1998). "Health values of hospitalized patients 80 years or older." *Jama-Journal of the American Medical Association* 279(5): 371–375.
- P. A. Ubel, G. Loewenstein, et al. (2003). "Whose Quality of Life? A Commentary exploring discrepancies between health state evaluations of patients and the general public." *Quality of Life Research* 12(6): 599–607.
- W. K. Viscusi and J. E. Aldy (2003). "The value of a statistical life: A critical review of market estimates throughout the world." *Journal of Risk and Uncertainty* 27(1): 5–76.
- World Bank. (2003). "EdStats – the World Bank's comprehensive Database of education statistics." 2006. Available from <http://www1.worldbank.org/education/edstats/index.html>.

DISCUSSION QUESTIONS

1. Bostrom argues that it could be good for a person to become posthuman with respect to cognitive ability, healthspan, and emotional capacity. Do you agree? Would you want to become posthuman in any of these respects (assuming the process of doing so is safe)? Why or why not?
2. Bostrom argues that personal identity could be preserved through a transition to being posthuman. What do you think constitutes personal identity? Given your understanding of personal identity, would an individual be the same person when they became posthuman?
3. Did you find Bostrom's response to the possible objections to his view persuasive? Why or why not?
4. Can you think of any objections to becoming posthuman that Bostrom does not consider?
5. If it is good for a person to become cognitively posthuman, as Bostrom argues, do you think that parents ought to technologically enhance their children to be cognitively posthuman? Would there be any ethical difference between what parents do now to help their children increase their cognitive capacities and the use of genetic and pharmacological technologies to help them do so?
6. This chapter focuses primarily on whether it would be good to become posthuman for the people who undergo the enhancement. Do you think it would be socially good – i.e. good for society as a whole – if there were widespread implementation of posthuman technological enhancement? Why or why not?