Chapter 8
What is the Good of Transhumanism?[1]

Charles T. Rubin

8.1 Introduction

Broadly speaking, transhumanism is a movement seeking to advance the cause of post-humanity. It advocates using science and technology for a reconstruction of the human condition sufficiently radical to call into question the appropriateness of calling it "human" anymore. While there is not universal agreement among transhumanists as to the best path to this goal, the general outline is clear enough. Advances in genetic engineering, artificial intelligence, robotics and nanotechnology will make possible the achievement of the Baconian vision of "the relief of man's estate," as they allow us to conquer disease, eliminate unhappiness, end scarcity and postpone, perhaps indefinitely, death itself. But fulfilling such long-standing dreams is only the beginning of what our new powers will make possible. Left to itself, the present trajectory of technological development necessarily aims at a future incomprehensible to beings such as we are – at no distant date an evolutionary leap in the way intelligence is embodied, and what it can accomplish. Transhumanism seeks to make sure no atavistic scruple obstructs this momentum, and to maximize its benefits and minimize its admitted risks.

While there is no lack of illuminating print works advocating transhumanism, that its public face should be on the World Wide Web is as much a matter of course as once would have been the use by similar movements of the printed broadside or the public lecture. On the websites of the World Transhumanist Association (www. transhumanism.org) and the Extropy Institute (www.extropy. org) – premier among transhumanism's many organizations – one finds authoritative statements explaining and justifying the transhumanist project. If transhumanism were primarily an academic school or a professional association, it would not be entirely fair to turn to these admittedly popular presentations for a critical look at the transhumanist vision. But as these are the documents by which transhumanism presents itself to the public as a movement, and through which it hopes to gain adherents, it is legitimate to make

them the primary, though not exclusive, focus of this analysis of transhumanism's vision of the way things ought to be.

This chapter will argue that however rhetorically effective it might be for transhumanism to present its opponents as obscurantists, the real debate between transhumanism and its thoughtful critics is not about further developments in science and technology *per se* but about the substantive goals for which science and technology will be employed. While at first glance transhumanism appears to aim at increasing health and wealth and extending life, a deeper look shows that the promise to reconstruct humanity necessarily will change the meaning behind these familiar aims in ways we cannot necessarily now comprehend. That ignorance is covered over by transhumanism's belief in diversity, a belief that proves to be perfectly consistent with human extinction. But the willingness to support diverse modes of not being human over being human ultimately illustrates a nihilistic aspect of transhumanist norms.

8.2 Progress, Competition and Restraint

Transhumanism advocates progress in scientific research and technological development, and reason as a foundation for both. Max More's "Principles of Extropy (3.11)" explains, "Extropy entails strongly affirming the value of science and technology" (More 2003: Section 4). The second of the substantive six parts of Nick Bostrom's "The Transhumanist FAQ: A General Introduction (2.1)" is about "Technologies and Projections," the ways in which cryonics, nanotechnology, genetic engineering, artificial intelligence and virtual reality will advance transhumanist goals (Bostrom 2003b: 7–19). The human condition can be improved through "applied reason" (Bostrom 2003b: 4). Thus, following Francis Bacon's early lead, the knowledge achieved through the use of reason is valued as a means to further ends. We want to know what these up and coming technologies are "good for" (Bostrom 2003b: 7).

The instrumental good of reason has important consequences with respect to the other side of the coin – the critical stance transhumanism takes against those it often calls "Bioluddites," people who make efforts to restrict developments in certain areas of science and technology because of fears about how they will be used (Hughes 2004: *passim*). While the indignation deployed against such efforts might make one think that transhumanists were standing up on behalf of reason and knowledge for its own sake, and therefore under all circumstances, that is not simply true. Transhumanists acknowledge that scientific and technological reason could produce frightening outcomes; "the gravest existential risks facing us in the coming decades," the FAQ say, "will be of our own making" (Bostrom 2003b: 22). But transhumanism makes what amounts to a prudential judgement that the best way of dealing with such risks is by anticipating them and creating the proper conditions for their avoidance or minimization. Generally that means that the very technologies that will pose the risks are the ones we will also have to rely on to

reduce them. Thus, for example, we can anticipate that genetic engineering might create the possibility of weaponised disease outbreaks. But we will need to rely on genetic engineering to produce cures or prophylactic measures. If "we" (however defined) choose to restrain relevant research, that will only leave us vulnerable to those who choose not to restrain themselves. Indeed, "we" had best be sure that we are well ahead of "them" in our abilities. Similarly, if nanotechnology can be imagined to have attractive commercial possibilities, "we" will lose market share to "them," or encourage a black market, if "we" operate under restrictions that "they" don't (Naam 2005: 39).

This arms race logic is central to transhumanism's effort to invest so far only imagined technologies and their consequences with an aura of inevitability. It is a powerful argument, and in the real world cannot be ignored. Yet while it certainly suggests the difficulty of control and restraint of technological development, it does not prove its impossibility or undesirability. It abstracts from difficult but necessary questions about restricting access to information and techniques even in a world where, on balance, we want research and development to be reasonably free. As a result, if it proves anything it proves too much, as Ray Kurzweil – a major intellectual ally of transhumanism – must have realized when he was moved to write a joint editorial with Bill Joy – a major critic – against the decision to publish the genome of the virus that caused the 1918 flu pandemic (Kurzweil & Joy: 2005).

Freedom of research and development does not always have to mean the widest dissemination of all results; that something has a market does not always mean it should be freely marketed. Restraint of development is more plausible the larger the "we" among whom there is a consensus grows and the more there is effective enforcement, social and/or legal, of that norm. This point is acknowledged when the FAQ speak on behalf of "expanding the rule of law to the international plane" (Bostrom 2003b: 33). Bostrom adds, "Global security is the most fundamental and nonnegotiable requirement of the transhumanist project" (Bostrom 2003a: Section 4). While such statements are quite vague (if the transhumanist project must wait on global security it could wait a very long time indeed), they suggest that at least all transhumanists are not advocating the anarchy under which the arms race logic is most compelling.

While some will use enforcement costs and lack of complete success at enforcing restraint as an argument for removing it altogether, that is an argument that can be judged on its particular merits – even when the risks of enforcement failures are extremely great. The fact that nuclear non-proliferation efforts have not been entirely successful has not yet created a powerful constituency for putting plans for nuclear weapons on the Web, and allowing free sale of the necessary materials. In the event, transhumanists, like "Bioluddites," want to make distinctions between legitimate and illegitimate uses of "applied reason," even if as we will see they want to minimize the number of such distinctions because, as we will note later, they see diversity as a good. Of course, those who want to restrict some technological developments likewise look to some notion of the good. This disagreement about goods is the important one, untouched by "Bioluddite" name-calling. The mom-and-apple-pie defense of reason, science and technology one finds in transhumanism is

rhetorically useful, within the framework of modern societies which have already bought into this way of looking at the world, to lend a sense of familiarity and necessity to arguments that are designed eventually to lead in very unfamiliar directions. But it is secondary to ideas of what these enterprises are good for, to which we now turn, and ultimately to questions about the foundation on which transhumanist ideas of the good are built.

8.3   Health, Happiness and Longevity

Transhumanism sees the good of scientific research and technological development in their proven ability to facilitate wealthier, healthier, longer and happier lives. "Principles of Extropy" says, "Science and technology are essential to eradicate constraints on lifespan, intelligence, personal vitality, and freedom" (More 2003: Section 1). As stated in the FAQ, transhumanism is "The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to eliminate aging and to greatly enhance human intellectual, physical, and psychological capacities" (Bostrom 2003b: 4). As "Principles of Extropy" puts it, "Pursuing extropy means seeking continual improvement in ourselves, our cultures, and our environments. Perpetual progress involves improving ourselves physically, intellectually, and psychologically" (More 2003: Section 1). It means, "Living vigorously, effectively, and joyfully" (More 2003: Section 3).

On the surface, we are again seeing little more than a restatement and elaboration of the fundamental Baconian project of "the relief of man's estate" that has been so definitive for the creation of the modern world. And so too there is something familiar about the transhumanist diagnosis of the roadblocks in the way of achieving such goals: nature, tradition and religion. The FAQ note how "Changing nature for the better is a noble and glorious thing for humans to do." They acknowledge that "the qualification 'for the better'" is crucial (Bostrom 2003b: 35) – but it seems that for transhumanism that is not a hard standard to reach. Hitting on all three of the roadblocks, the FAQ describe how, "the pre-industrial age was anything but idyllic. It was a life of poverty, misery, disease, heavy manual toil from dawn to dusk, superstitious fears, and cultural parochialism" (Bostrom 2003b: 29). "Principles of Extropy" likewise argues that "Perpetual progress calls for us to question traditional assertions that we should leave human nature fundamentally unchanged in order to conform to 'God's will' or to what is considered 'natural'". Or again, "Valuing perpetual progress is incompatible with acquiescing in the undesirable aspects of the human condition. Continuing improvements means challenging natural and traditional limitations on human possibilities" (More 2003: Section 1).

Despite this rejection of so many traditional sources for grounding or deriving values, the FAQ suggest that "it is perfectly possible to be a transhuman – or, for that matter, a transhumanist – and still embrace most traditional values and principles of personal conduct" (Bostrom 2003b: 7). Since people hold in their heads all kinds

of contradictory beliefs at the same time, doubtless there is that much truth to this careful formulation. But David Pearce, a cofounder of the World Transhumanist Association presents in his "hedonistic imperative" a more consistent picture of the relationship between transhumanism and traditional values and principles. He draws conclusions that, while doubtless not universally accepted among transhumanists, open the door to seeing some problems within the broader argument about the meaning of "for the better."

"Nature is barbarous and futile beyond belief," Pearce believes.

> Warfare, rape, famine, pestilence, infanticide and child-abuse have existed since time immemorial. They are quite 'natural,' whether from a historical, cross-cultural or sociobiological perspective. The implicit, and usually highly selective, equation of the 'natural' with the morally good is dangerously facile and simplistic. The popular inclination to ascribe some kind of benign wisdom to an anthropomorphized Mother Nature serves, in practice, only to legitimate all manner of unspeakable cruelties (Pearce 1998: 4.6).

So Nature has dealt human beings a pretty bad hand. But Pearce believes that advances in understanding and manipulating the brain chemically and genetically will eventually make it possible for people to be happy all the time. Indeed, people will be so unimaginably (to us) happy that previous human psychology, based as it was on the naturally given, will be seen in this hedonic future as tragic, perhaps incomprehensible, mental illness. On the basis of a particular variant of utilitarianism, Pearce concludes that we have a moral obligation to seek out this euphoria for ourselves, and indeed to reconstruct the natural order entirely so that any other beings capable of suffering will likewise be happy. "It is not, needless to say, the fault of cats that they are prone to torturing mice; but then, given the equations of physics, it isn't the fault of Nazis they try to persecute Jews. This is no reason to let them continue to do so" (Pearce 1998: 1.10). But more: "For ethically it is imperative that the sort of unspeakable suffering characteristic of the last few hundred million years on earth should never recur elsewhere. If such horror might exist anywhere else in the cosmos, presumably in the absence of practical intelligence sufficiently evolved to eliminate its distal roots, then this suffering too must be systematically sought out. It needs to be extirpated just as hell-states will have been on earth" (Pearce 1998: 4.13).

Imagine this rescue fleet arriving on *our* doorstep; it seems likely that its promise of the complete reconstruction of human psychology and terrestrial ecology would be greeted with alarm. For Pearce that atavistic reaction is merely an indicator of the raw deal we have from Darwinian evolution, which has not selected for the prevalence of brain states that we call happiness. It is also built on an unwillingness to confront the fact, as he presents it, that happiness as we experience it is entirely a matter of brain chemistry. If I am happy because I'm in love with a biological person or with a virtual person whose computer program is stimulating the appropriate nerve centers in my brain, or born with a sunny disposition, or properly medicated or genetically enhanced, it is essentially the same thing as far as what goes on in the brain is concerned. Conditions in our brain such as that which we label happiness are products of the laws of physics, and as such open to deliberate and self-conscious manipulation.

   Three observations need to be made about this effort to revise our understand-
ings of the human good. First, Pearce promises us freedom from the *particulars* of
the naturally given (the way our brains happen to have evolved) while at the same
time completely subsuming the human into the naturally given *in general* (the
chemistry and physics of the brain that make us what we are). That would seem to
accord with the view of the FAQ: "There is no fundamental dichotomy between
humanity and the rest of the world. One could say that nature has, in humanity,
become conscious and self-reflective" (Bostrom 2003b: 38). So the Blind Watchmaker
has by chance created not just a watch but a Sighted Watchmaker; intelligence and
intentionality replace chance. Yet even the Sighted Watchmaker remains bound by
the fundamental constraints through which the Blind Watchmaker makes watches.
Transhumanism generally is committed to the same proposition, essentially accepting
the Baconian dictum that "nature to be commanded must be obeyed." This position
suggests that our sense of freedom may always have been an illusion of consciousness,
or based on ignorance. What we think of as our choice to conquer nature is dictated
to us by nature.
   Hence (in the second place), the theme of overcoming nature requires more
attention. Pearce may be at an extreme in his willingness to present such a thor-
oughgoing moral condemnation of the naturally given. The FAQ seem to stand in
contrast, sometimes speaking with a more conventional voice of environmental
concern. ("Not only are transhumanist technologies ecologically sound, they may
be the only environmentally viable option for the long term" (Bostrom 2003b: 38)).
But ecological soundness and environmental viability do not have to be taken as
phrases that could only characterize present or historically given ecosystems or
environments. "Nature" in this respect could be modified along with human beings.
There is no reason to think the FAQ suggest any more serious scruples about the
macrocosm than it does about the microcosm, however much lip service must be
paid to the environmental awareness necessary for coalition building.
   Still, it is certainly true that nature often falls short of human moral aspirations,
just as it is true that human beings can exhibit a depravity that is hard to see in the
rest of nature. How to understand the human relationship to nature, which arises as
a question out of the fact that we are the only beings we know to have arisen out of
the naturally given with the ability to significantly alter that given, is a question too
important to be allowed to run to the extremes such as Pearce's thoroughgoing
condemnation, or more conventional environmentalist veneration.[2]
   Yet third, despite his ability to speak the language of happiness – and after all,
who does not want to be happy? – Pearce knows he has an uphill battle to get his
readers on board with statements like, a "symbiotic union of biologically pro-
grammed euphoria and mature virtual reality software engineering, however, is an
awesomely good prospect" (Pearce 1998: 3.5). "Accounts like this," he admits,
"inevitably sound cold, technocratic and Brave New Worldish. It should be recalled

---

[2] Rejection of/reliance on nature is not unique to transhumanism and is of a piece with understand-
ing nature within the limits of the framework set by modern assumptions – a particularly illumi-
nating instance of which is Mill 1969: 373–402.

that the developments they describe should avert suffering on a scale which a single mind cannot possibly comprehend; and make a lot of people blissfully well" (Pearce 1998: 3.3). "Blissfully well" here means not achieving the mean or being saintly, for example; reasoning teaches Pearce that it means rather possessed of a brain genetically engineered to produce certain chemical states beyond what the unengineered brain produces while being electronically stimulated by a computer in such a way as to be experiencing sensations that do not correspond to the actual location or activity of the body. Indeed, so "awesomely good" is this prospect that Pearce uses it to suggest why in fact the alien happiness rescue fleet has not done its moral duty and arrived on our doorstep from a far more advanced civilization that has already taken the step he contemplates. As the "motivational incentives to choose the inconvenient kinds of experience involved in (non-virtual) space-exploration etc. are somewhat diminished…the very possibility of vulgar physical star hopping may just never arise" (Pearce 1998: 3.5). Forget "Star Trek" or "Firefly," with their conflicted and often unhappy people; while Pearce thinks he has arguments that suggest why constant happiness might lead to ever greater human achievement, his premises are always drawing him back to a "wirehead" future.

Pearce's transhumanism suggests in a particularly dramatic way how the transhumanist promise of wealth, health, happiness and longevity has specific content which will not necessarily correspond to people's pre-existing understandings of what makes for a good life. Of course longer, healthier and more productive lives, are going to command wide endorsement as good things. Yet our existing understandings of these goods do not comprehend how we can use the laws of nature to free ourselves from norms drawn from natural, traditional or religious constraints that up to the present have defined this goodness. At the same time, they tend not to be built on the strong determinism about the source and meaning of our choices such as is found in Pearce's doubtless intentionally shocking statement (even given Art Spiegelman's *Maus*) equating cats and Nazis, instead presupposing some degree of moral freedom.

Again, not all transhumanists would draw Pearce's particular consequences so sharply. But despite the fact that the FAQ and "Principles of Extropy" do not enter into discussions of their underlying principles so deeply as Pearce, it is hard to see how some form of the same premises one finds in Pearce could fail to inform transhumanism's picture of happiness more generally. Outside of the FAQ, Bostrom acknowledges that transhumanism "has its roots in secular humanist thinking" (Bostrom 2003a: Section 1). The manner in which science, technology and reason are described and endorsed would certainly lead one to think that transhumanism is committed to the modern, scientific materialism that sees nature, including human nature, as malleable; happiness will be defined by the manipulation of stuff whether inside or outside the brain. Otherwise, the whole discussion of overcoming limits via technology would hardly make sense. But as a consequence, a good many traditional views must fall by the wayside – any that depend on the existence of a soul, for example, or divine revelation or on a given human nature, or even on existent social constellations. Among "traditional" norms, then, which could *consistently* be combined with transhumanist aspirations as readily as some sort of pragmatic utilitarianism of the sort Pearce adduces?

For a "loosely defined movement" seeking to expand its base of adherents, it would not necessarily serve to press such consistency too hard. When it comes to presenting its vision of the good, one often finds in transhumanism a rhetoric of compassion to supplement its rhetoric of development and discovery. One finds an initial emphasis on the techniques that will allow healing of those suffering from specific disability and disease, apparently accepting the existence of something like a norm of health (Naam 2005: *passim*). But this acceptance is only provisional, because the same techniques that will cure disease and disability will also prove useful to enhance and reconstruct. We may create direct brain/machine interfaces in order to help restore the ability to move and manipulate to paralyzed people; the more we know the more such linkages will become commonplace for all. Eventually, to fail to have such a hookup may be regarded as a disability.

Broadly speaking, then, what is promised is not the same as what is delivered. Transhumanism stands foursquare behind the now traditional modern project for improvement of the human condition. The FAQ, for example, pose the question "Shouldn't we concentrate on current problems such as improving the situation of the poor, rather than putting our efforts into planning for the 'far' future?" The answer, of course, is "We should do both" (Bostrom 2003b: 27). But the paths it chooses to such goals are determined by its vision of the technologically imaginable, rather than by some determinate idea of a good life, because it advocates overcoming the existing limitations that define the human condition, and a world that is in some way unimaginable to the merely human.

To put it another way, what "healthier," "wealthier," "happier," and even "longer" lives look like, what the very terms themselves mean, becomes a new problem once we start to move from transhumans to posthumans, as the bars of achievement and ability rise to unprecedented heights. Of course, there have always been disagreements about what makes a good life. Transhumanism is not the first, even secular, movement to put the force of necessity behind projections of an end to human ills, nor to imagine beyond that some unimaginable future destiny. It is even not unique in rejecting nature, religion, tradition or history as relevant standards by which a good life may be defined. But such extreme open-endedness creates a problem. This principled uncertainty about the meaning of the posthuman is relevant to the judgement of cases today; it influences essential elements of the transhumanist case for what we should be doing now. For already to some extent today and all the more so even in the near term, all kinds of things are or will be possible that, from the non-transhumanist point of view, don't look like good ideas. Objections to cloning or electrically induced happiness can't only be answered by claims that we *must* develop these technologies; technological might makes right no more than any other variety. Yet justification in terms of achieving some ineffable posthuman condition is not exactly a reasoned defense, and the very incomprehensibility of the posthuman makes it look like allegiance to it is an article of faith. In light of that problem, transhumanism makes a virtue of a necessity, and celebrates a future built up from the maximization of choice, a future of diversity and inclusion. It aspires to the higher selfishness, a world in which "it looks good to me" commands moral respect and is a key principle of social organization.

8.4   Diversity, Choice and Inclusion

"Self-direction" is one of the seven keys "Principles of Extropy"; technology creates new possibilities of choice that must be appreciated without the burden of ideas inapplicable to these new circumstances.

> Self-direction calls on us to rise above the surrender of independent judgement that we see – especially in religion, politics, morals, and relationships. Directing our lives asks us to determine for ourselves our values, purposes, and actions. New technologies offer more choices not only over what we do but also over who we are physically, intellectually, and psychologically. By taking charge of ourselves we can use these new means to advance ourselves according to our personal values (More 2003: Section 6).

Hence "Each individual should be free and responsible for deciding for themselves in what ways to change or to stay the same… Pursuing extropy means vigorously resisting coercion from those who try to impose their judgements of safety and effectiveness of various means of self-experimentation" (More 2003: Section 6).

By using the phrase "safety and effectiveness," More likely aims this last shot at the United States Food and Drug Administration, for it is tasked by statute precisely with making those judgements. In any case, for More doubtless any regulatory regime so tasked is inappropriate when it comes to transhuman enhancements. This position is disputed by less libertarian transhumanist advocates, who see an ongoing role for government regulation in such matters. Indeed, there is some tension within the "Principles of Extropy" themselves, evident in the assertion that "Coercion of mature, sound minds outside of the realm of self-protection, whether for the purported 'good of the whole' or for the paternalistic protection of the individual, is unacceptable" (More 2003: Section 6). For when is coercion for self-protection not paternalistic? But the basic preference for maximization of free choice and individual responsibility for such choice is clear enough – choice, however, whose maturity is presumably indicated by the extent to which it is "rational," "reflective" and "informed."

"Since self-direction applies to everyone, this principle requires that we respect the self-direction of others … Appreciating that other persons have their own lives, purposes, and values implies seeking win-win cooperative solutions rather than trying to force our interests at the expense of others" (More 2003: Section 6). So there is no illusion here about the diversity of choices likely to be made and the potential for conflict contained therein, but a hope that such situations can be met with a "benevolent disposition" which "embodies more emotional stability, resilience, and vitality than cynicism, hostility or meanness," approaching others in a spirit of "friendship cooperation and pleasure" (More 2003: Section 6).

The end result is an "open society," another of the main points of the "Principles of Extropy." The concept of

> Open societies avoids utopian plans for 'the perfect society,' instead appreciating the diversity in values, lifestyle preferences, and approaches to solving problems. In place of the static perfection of a utopia, we might imagine a dynamic 'extropia' — an open, evolving

framework allowing individuals and voluntary groupings to form the institutions and social forms they prefer. Even where we find some of those choices mistaken or foolish, open societies affirm the value of a system that allows all ideas to be tried with the consent of those involved (More 2003: Section 5).

Of course, libertarians have espoused ideals such as this since long before transhumanism came on the scene. But to those who have, in light of their reading of human nature or history, hitherto greeted them with slack-jawed amazement, Extropy has the reply that now we can talk about such arrangements in light of a complete human redesign. Benevolent dispositions can be manufactured. Still, there is some tension lurking between allowing all ideas to be tried as a mechanism to achieve posthumanity, and requiring posthumanity to allow all ideas to be tried, a tension that again speaks to the confusing freedom that transhumanism promises. If I am born benevolent by someone's clever redesign, then a whole realm of moral choice is denied to me. If, as is imagined by James Hughes, executive director of the World Transhumanist Association, I can turn benevolence on and off as one among a variety of programmable dispositions, we are back to square one, wondering where the disposition to flick that switch as opposed to others Hughes imagines ("Kohlberg's Stage Six, Islamic Sharia, or Ayn Randian selfishness" (Hughes 2004: 255)) is coming from – since of course even without such enhancements one is able to choose to be benevolent. Or not, if Pearce is correct, in which case this freedom promised by transhumanism is as illusory as any other.

The FAQ, while somewhat less libertarian in emphasis, further complicate the picture of what diversity will mean as a practical matter, suggesting that it will have to accommodate challenges from two directions. "Transhumanists reject speciesism, the (human racist) view that moral status is strongly tied to membership in a particular biological species, in our case homo sapiens" (Bostrom 2003b: 31). That means that on the one hand, animals that may not be sufficiently in our moral circle will have to be included, since "all beings that can experience pain have some moral status" (Bostrom 2003b: 31). On the other hand, "posthuman" creations will likewise need to be included; indeed, there are already serious discussions even of the rights and legal status of sentient computers. (A further complication: if Pearce is correct, will the euphoric beings of the future feel pain at all, or as we do – since there will no longer be any reason for it to produce unhappiness?)

How will societies deal with this expanded range of moral inclusion? From the side of the posthuman it is difficult to say because "we must bear in mind that we are likely to base our expectations on the experiences, desires, and psychological characteristics of humans. Many of these expectations may not hold true of posthuman persons. When human nature changes, new ways of organizing a society may become feasible" (Bostrom 2003b: 32). It may be that transhumanists recognize the echoes of 20th century totalitarianism in this promise of what becomes possible if only we can change human nature. Perhaps as a way of avoiding the same terrible results, the FAQ suggest that:

> The ideal social organization may be one that includes the possibility for those who so wish to form independent societies voluntarily secluded from the rest of the world, in order to pursue traditional ways of life or to experiment with new forms of communal living.

> Achieving an acceptable balance between the rights of such communities for autonomy, on the one hand, and the security concerns of outside entities and the just demands for protection of vulnerable and oppressed individuals inside these communities on the other hand, is a delicate task and a familiar challenge in political philosophy (Bostrom 2003b: 32).

The state of the world today may suggest that this challenge can be "familiar" without having been definitively met, despite the efforts of political philosophy hitherto. One might speculate that it is a problem that worsens as human power increases. Yet the modes of political and social organization that in practice seem to work relatively well in a highly diverse nation such as the United States work within the framework of "human racist" assumptions about the meaning of equality such as are suggested in the Declaration of Independence. Furthermore, contrary to the hopes of Extropy, these modes of organization do not by and large depend decisively on benevolence and niceness but rather assume, with James Madison, that men are not angels. Indeed, the somewhat glib protestations about new social forms more than anything else call attention to the possibility that posthumans will be as little concerned for those mere humans choosing "traditional ways of life" as those humans have been about other animals.

That this possibility is real is acknowledged by transhumanists. Speaking of the possibility of super-intelligent, sentient computers, the FAQ note that:

> The would-be creator of a new life form with such surpassing capabilities would have an obligation to ensure that the proposed being is free from psychopathic tendencies and, more generally, that it has humane inclinations. For example, a superintelligence should be built with a clear goal structure that has friendliness to humans as its top goal. Before running such a program, the builders of a superintelligence should be required to make a strong case that launching it would be safer than alternative courses of action (Bostrom 2003b: 34).

FAQ Version 1 was more open on the topic than FAQ 2, for in FAQ 1 we find explicit admission that "if the posthumans are not bound by human-friendly laws and they don't have a moral code that says it would be wrong, they might then decide to take actions that would entail the extinction of the human species" (Hughes 2004: 247). FAQ 2, on the other hand, professes to find the concern about such conflict overblown.

> It is a common theme in fiction because of the opportunities for dramatic conflict, but that is not the same as social, political, and economic plausibility in the real world. It seems more likely that there would be a continuum of differently modified or enhanced individuals, which would overlap with the continuum of as-yet unenhanced humans. The scenario in which 'the enhanced' form a pact and then attack 'the naturals' makes for exciting science fiction but is not necessarily the most plausible outcome. Even today, the segment containing the tallest 90 percent of the population could, in principle, get together and kill or enslave the shorter decile. That this does not happen suggests that a well-organized society can hold together even if it contains many possible coalitions of people sharing some attribute such that, if they unified under one banner, would make them capable of exterminating the rest (Bostrom 2003b: 33).

Here again we have the familiar Madisonian notion that societies with a great degree of social and economic diversity can limit the formation of dangerous factions. Yet that idea works best on the basis of an ideal of human equality which was itself rather hard won in the face of a range of actual inequalities that pale in comparison

with those which will be introduced by trans- and posthumanity. So it is extremely curious that in the context of a criticism of fiction for being unrealistic, FAQ 2 pass over the known history of relations between human beings and animals, between human beings at vastly different levels of technological development, indeed between human groups that are in any manner "strange" to one another and instead has us celebrate the social accomplishment that today taller people have not killed or enslaved shorter people.[3]

While it is strictly speaking true, given our complete ignorance of the constraints that will operate on posthuman beings, that efforts on their part to bring about human extinction is "not necessarily the most plausible outcome" one could with equal or greater truth say it is "not necessarily the most implausible outcome."

Such concern is justified because all the characteristics by which we understand and admire diversity in the world we now know are, as the transhumanists are fully aware, human characteristics, based on a human given that limits our potential for good or ill. Since it is just that given which transhumanism proposes to eliminate, all bets about the resulting moral universe are off. The FAQ want to avoid this consequence by distinguishing between being human and being humane:

> If there is value in being human, it does not comes from being 'normal' or 'natural,' but from having within us the raw material for being humane: compassion, a sense of humor, curiosity, the wish to be a better person. Trying to preserve 'humanness,' rather than cultivating humaneness, would idolize the bad along with the good. One might say that if 'human' is what we are, then 'humane' is what we, as humans, wish we were. Human nature is not a bad place to start that journey, but we can't fulfill that potential if we reject any progress past the starting point (Bostrom 2003b: 36).

Here humane attributes are being treated as abstractions only accidentally connected to our humanity. But the positive characteristics mentioned are positive precisely because we are the kind of being that we are. Compassion and empathy, for example, are positive because we do not have a hive mind, and are separated by our bodies, and can suffer. Or again, we admire "the wish to be a better person" because it is hard to be better; we can't just buy upgrades and we have to fight against passions and interests that do not make us better.

So Hughes is quite right to wonder whether it "is possible to imagine a 'liberty-respecting' policy that discourages misanthropy among posthumans" (Hughes 2004: 248). Will the mature, rational and informed choices of humans look the same way to transhumans or posthumans, if as expected their capacities in all these areas are superior to ours? And if not, what respect will they grant them? The

---

[3] Using the persistence of a short minority as an example of social tolerance is odd given that height is well documented to convey advantage even in "well-organized societies"; the shortest are allowed to survive, but at a significant disadvantage. Indeed, the point becomes downright bizarre upon recollection that one of the main areas today in which enhancement of children is already being practiced is providing them with human growth hormone. While by definition there will always be a shortest decile, in fact there are societies already open to making it as tall as possible, "eliminating" the category as defined by today's measurements.

respect that parents give to the choices of children? That humans give to the choices of their pets? That humans give to the choices of nuisance animals?

8.5   Enhancement, Identity and Extinction

While the transhumanists use the traditional language of libertarian inclined progressivism to discuss the good of transhumanism, there is really no way to dispute that they are leading us into completely uncharted moral waters. In that context, it is not foolish to be concerned about the fate of humanity, and not only out of the conventional worry that highly advanced beings might find their precursors an embarrassing nuisance, or that we may fall prey to their incomprehensible projects or conflicts, or that we might be useful to them in some degrading way, or that great power might easily coexist with great malevolence, or that we will simply be outcompeted in an evolutionary struggle. An underappreciated source for human extinction might be found in a corollary to Arthur Clarke's law "any sufficiently advanced technology is indistinguishable from magic": any sufficiently advanced act of benevolence is indistinguishable from malevolence (Rubin 1996: 168). Of course I am doing the right thing by not giving the pan-handler money to buy drugs with – but perhaps he does not see it quite the same way. One need only recall the aforementioned arrival of the fleet of alien ships intent on making all sentient life happy all the time to see the problem here. Indeed, Clarke himself wrote the definitive book on the subject, *Childhood's End*, in which a benevolent race comes to shepherd humanity to the next evolutionary level, literally destroying the world and all remaining human beings as the *successful* result of their mission.

Yet even if, as the FAQ would have us believe, such inter-specific problems can be solved or are overblown, it remains the case that transhumanism is in effect promising the end of humanity. For if they are correct about the appeal of the possibilities inherent in transhumanity and posthumanity, what mature, rational decision could be made to remain human? As transhumanists tend to portray those who oppose them as in thrall to the irrational, as bio-Luddites, as racists, as death-lovers they are in effect saying that they can imagine no good reasons why people would not enhance themselves to the maximum extent possible. But of course if anyone *wants* to decay and die, transhumanism would not, as is supposed of its opponents, impose its views and prevent it.

Or would it? An illuminating point comes up when the FAQ attempt to stand forthrightly against eugenics and for reproductive freedom. Yet:

> Beyond this, one can argue that parents have a moral responsibility to make use of these methods, assuming they are safe and effective. Just as it would be wrong for parents to fail in their duty to procure the best available medical care for their sick child, it would be wrong not to take reasonable precautions to ensure that a child-to-be will be as healthy as possible. This, however, is a moral judgment that is best left to individual conscience rather than imposed by law. Only in extreme and unusual cases might state infringement of procreative liberty be justified. If, for example, a would-be parent wished to undertake a genetic modification that would be clearly harmful to the child or would drastically curtail

its options in life, then this prospective parent should be prevented by law from doing so. This case is analogous to the state taking custody of a child in situations of gross parental neglect or child abuse (Bostrom 2003b: 21).

Since already today failure to provide necessary blood transfusions can count as "gross parental neglect," is it so difficult to imagine a day when a parental choice *not* to make a generally accepted genetic modification in their child would trigger "state infringement"? In a world where technological enhancement is expected to widen greatly "options in life," will not a parental decision to accept the current norm more and more look like the choice of tragic yet preventable disability?

Nor does the existence posited in the FAQ of a "continuum of differently modified or enhanced individuals" (Bostrom 2003b: 33) between the human and the posthuman really act as intended to change the anti-human dynamic of the argument. First of all, is this continuum really consistent with the supposed necessity on which transhumanism depends that drives scientific and technological development in the first place? Will not these competitive forces move people to seek maximum available enhancement, not to speak of the supposed advantages of so doing in living healthier, happier, wealthier and longer lives? People will surely not always make the same choices of particular enhancements, but if the transhumanist program prevails, they will surely tend to maximize enhancement to the limit of the constantly changing desires and capabilities, about which more below. An analogy: at present, relatively few people in the US have a big plasma TV and the widest cable selection of channels, but nearly everyone has some sort TV, and of those who don't only a tiny fraction are holding out in principle. What from one point of view is a continuum of TV possibilities is from another an isolated minority of non-TV watchers.

Furthermore, the TV continuum is created in part by price discrimination. Transhumanists acknowledge this effect as a short-term issue; early adopters of new possibilities will pay a high price for them. But they also point to the powerful tendency for technology prices to go down over time, and/or call for public support for the provision of otherwise too costly technological benefits (Naam 2005: 63–66). This effort to deal with wealth based inequalities that historically tend to leave the have-nots at the mercy of the haves would be more convincing if it were not for the "perpetual progress" element of transhumanism, which would seem to imply that there will always be an exploitable, expensive advantage to be found at the cutting edge. Of course, here again it is doubtless an error to apply human-based thinking to the transformed beings advocated by transhumanism. Maybe they will find a way to make us free in the face of competitive necessity, and equal in our radical inequality.

The more one appreciates the gap that transhumanism would have us believe will exist between the human and the posthuman, the more it appears that the "ideal," continuum of possibilities or not, will indeed be the opportunity for those wishing to maintain their "traditional" humanity to be voluntarily "secluded" from the rest of the world. The survival of the voluntary element in this seclusion will surely be tested the moment mere humans begin to look like a danger to themselves or to posthumans if left in a mixed milieu.

The foregoing thoughts cannot prove that successful transhumanism will result in human extinction or even a bad deal for the merely human. But they can remind that the advocacy of free choice and the protection of diversity which substitute for a substantive picture of the human good in transhumanism are themselves instrumental to creating a posthuman world which, on assumption, we cannot know to have any room for human values like choice or diversity. In any case, transhumanists appropriately seem clearer that the story of intelligence will pass out of human hands. "The arrival of superintelligence will clearly deal a heavy blow to anthropocentric worldviews. Much more important than its philosophical implications, however, would be its practical effects. Creating superintelligence may be the last invention that humans will ever need to make, since super intelligences could themselves take care of further scientific and technological development. They would do so more effectively than humans. Biological humanity would no longer be the smartest life form on the block" (Bostrom 2003b: 13).

What transhumanism does, then, is to dangle before us the glorious possibilities of human inventiveness, with the ultimate expectation that human inventiveness will be superceded. If what defines our humanity is self-transcendence as a species (Kurzweil 2005: 9), will those who remain human yet uninventive be in effect less than human? Here in about the most artless form imaginable we are being offered a Faustian bargain. "Fulfill your deepest material hopes and dreams – but by the way, the price is human obsolescence." One could call this bargain devilish were it not made with such illuminating, if occasional, candor, as if by some imp under compulsion periodically to blurt out the truth (Todorov 2002: 3). Yet the promise compels nevertheless by the very familiarity of the good things it promises, by the imagination of ring of Gyges-like powers of wish fulfillment to come, by intimation of presently unimaginable pleasures, by the ultimate hope of staving off our mortality.

But to whom exactly is the bargain being offered? One might think from the transhumanist emphasis on individual choice and social diversity, from their effort to claim the moral high-grounds of compassion, benevolence and cooperation, that the choice is being offered to a free moral being. But we have already seen how that is hardly the case. In asserting that Nazi and cat alike act faultlessly out of the "equations of physics," Pearce is simply making explicit the consequences of the scientific materialism on which transhumanism seems committed to building its understanding of human things. From this point of view, we are not free, and what we call our subjectivity is a manipulable object.[4] To think that a perduring "I" is freely accepting or declining enhancement and modification is to fall prey to the error of thinking there is a "ghost in the machine" when in fact such a decision is the in-principle calculable result of determinate bodily biology/chemistry/physics. "The reality is we are constantly changing" (Naam 2005: 59).

---

[4] The issue of what kind of human freedom, if any, is consistent with scientific materialism is of course not unique to transhumanism – but that does not mean it ought to be elided. There are efforts to try to reconcile scientific materialism with genuine human freedom. See, for example, Wolfram 2002: 750–53 or Hameroff and Penrose 1996.

But there is a tension in transhumanist thinking in this connection. The FAQ talk about "uploads," the prospect of creating posthuman beings by transferring the contents of brains into a computer. The resulting being could be embodied in robot form, or it could inhabit virtual realities. In any case, "it matters little whether you are implemented on a silicon chip inside a computer or in that gray, cheesy lump inside your skull, assuming both implementations are conscious" (Bostrom 2003b: 18). But of course it matters a great deal; the FAQ also admit that downloads raise "philosophical, legal and ethical challenges" galore: "if we imagine that several similar copies are made of your uploaded mind. Which one of them are you? Are they all you, or are none of them you?" (Bostrom 2003b: 18) Such questions do not arise with human beings; they are indicative of a profound change in the meaning of the self, and hence of the moral universe the self inhabits.

It may be that scientific materialism already undermines notions of a unified self, but transhumanism absolutely seeks to explode them (Todorov 2002: 4–5). Not only does the I that chooses become thereby a different I, but in effect it does not choose, the outcome being determined by the temporary physical state of the I prior to the change to a new state. So it is not quite right to suggest that "In a world where we can sculpt our own emotions and personalities, people will no longer be able to say 'I can't help it, that's just the way I am'" (Naam 2005: 60). For there is just as little an I free to sculpt itself in the new world as there is an I free to resolve to change in present world.

8.6   Uncertainty, Faith and Nihilism

Starting from transhumanism's materialism, then, it is the purest conjecture, pure faith, that would make one think that qualities of humaneness derived from the human would have any meaning in a world bent on overcoming the human. Nick Bostrom attempts to defend against this conclusion that transhumanism commits unknowable selves to unknowable values. On the one hand, he notes that not all enhancements need be so radical as to create discontinuity of personal identity, that some posthuman values may be things we already value now, but only with an incomplete understanding of them, and that an "incremental exploration of the post-human realm may be indispensable for understanding posthuman values" (Bostrom 2003a: Section 3). But what is noteworthy is the relentless pressure that the logic of transhumanism places on such moderate (relatively speaking) formulations. For on the other hand, (a) "Preservation of personal identity, especially if this notion is given a narrow construal, is not everything" (b) "we may favor future people being posthuman rather than human, if the posthumans would lead lives more worthwhile than the alternative humans would" (c) "Transhumanism promotes the quest to develop further so that we can explore *hitherto inaccessible realms of value*" [emphasis added]. So "worthwhile" may well mean "worthwhile in terms of some now inaccessible realm of value". Finally, (d) "if the mode of being of a posthuman being is radically different from that of a human being, then we may doubt whether

a posthuman being could be the same person as a human being, even if the posthuman being originated from a human being" (Bostrom 2003a: Section 3).

In connection with this last point, Bostrom makes what he doubtless feels is a telling *tu quoque* argument. "Depending on what our views are about what constitutes personal identity, it could be that certain modes of being, while possible, are not possible for us, because any being of such a kind would be so different from us that they could not be us. Concerns of this kind are familiar from theological discussions of the afterlife." In "Christian theology," souls only enter heaven after a period of purification. "Skeptics may doubt that the resulting minds would be sufficiently similar to our current minds for it to be possible for them to be the same person" (Bostrom 2003a: Section 3).

Leave aside the equivocation about the corrosive effect of scientific materialism on concepts of personal identity, the confusion between mind and soul, and the theological argument that the soul becomes more essentially itself as it is purified. Note instead that Bostrom tacitly admits a generic likeness between transhumanism, which we thought was a product of reasoning, and arguments based on faith. This insight is important. When the goal of transhumanism becomes the incomprehensible posthuman, discussion of it, as of the afterlife, becomes a matter of faith.

Not simply faith, of course. If the transhumanists are correct, their faith will be incrementally justified by works – scientific and technological achievements – until such time that intelligence ascends to levels inaccessible to those who are left behind. Those works will doubtless fall out along a continuum, from those which are said to be good from the point of view of the mundane human, to others (technically speaking, no longer on a continuum) which will appear good only to those whose faith allows them to see from *sub specie* the posthuman apotheosis. Since the very notions of transhumanity and posthumanity point so firmly to this extreme, we may once again wonder what it means to define a good in terms of the unknowable.

Biblical theologies that point to an equally unknowable end will not help illuminate this question, because in them the route to that end is through defined, humanly comprehensible acts of self-discipline called forth by a providential order in which human power is vanishingly small. Transhumanism seeks to liberate the diverse wills of the choosing individuals to remake their worlds, even as it turns those free choices into outcomes determined by laws of nature – but also only limited by laws of nature. So where Biblical theologies face the unknowable in an attitude of submission, transhumanism uses it to liberate what Thomas Hobbes long ago called a "restless desire of power after power that ceaseth only in death" (Hobbes 1968: 161), assuming death cannot be indefinitely delayed. How do we judge the extent of our power? By our ability to negate what is given by nature or tradition, to be sure, but also power to negate *whatever* is for the moment given in the name of possibilities foreseen or unforeseen. That is the deeper meaning behind the bland pragmatism in the "Principles of Extropy" discussion of "perpetual progress" ("conserving what works for as long as it works and altering that which can be improved"): "no mysteries are sacrosanct, no limits unquestionable" (More 2003: Section 1).

Negation can be good. To negate suffering and disease is well worth attempting; the effort to negate death is at least understandable. But finally, transhumanism

does not seek to negate suffering and disease in the name of human happiness and health, but rather in the name of the willful achievement of any imaginable and unimaginable alternative. Pious caveats about safety and humaneness and equality of access will ultimately have no traction in the face of this assumption. Promoting discussion of the future and "an interdisciplinary approach to understanding and evaluating the opportunities for enhancing the human condition and the human organism" (Bostrom 2003a: Section 1) seem a weak reed with which to hold off the will to power (Garreau 2005: 241).

So without firm ground to stand on, there is more than a hint of nihilism in transhumanist negation. The rationality which it values so highly is itself moored to nothing but the will of the individual reasoner. According to the "Principles of Extropy" reason helps us "understand the Universe" and "advance our knowledge," but its essentially critical function cannot allow even such claims to stand unqualified: we have to "remain wary of the human propensity to settle for and defend any comfortable explanation" (More 2003: Section 7). We have gone far beyond Socratic doubt, here. "Rational thinkers accept no final intellectual authorities. No individual, no institution, no book, and no single principle can serve as the source or standard of truth. All beliefs are fallible and must be open to testing and challenging. Rational thinkers do not accept revelation, authority, or emotion as reliable sources of knowledge. Rational thinkers place little weight on claims that cannot be checked. In thinking rationally, we rely on the judgement of our own minds while continually re-examining our own intellectual standards and skills" (More 2003: Section 7). In making everything perpetually provisional, the reason of transhumanism means we can do anything we want so long as we maintain a critical distance from the doing of it – a utopia of irony.

Transhumanists are doubtless decent folk, struggling to do right in the world as they see the right. So the conclusion that the good of transhumanism is finally "everything is permitted" will probably be objectionable to many, who would never dream of themselves offending against the "values, standards and principles" of their milieus. But how could it be otherwise when the very goal of the movement is overcoming the constraints of the human, constraints which define the character of our moral world, a goal which constantly pushes it to extremes? When nature is shorn of goals or purposes that would imbue it with moral significance? When tradition, religion, and custom all are to be tried at the bar of the willful individual reasoner? When it must have faith in the desirability of a future that it admits we cannot understand?

8.7 Conclusion

The transhumanist program to rebuild humanity leaves it morally adrift. Despite the fact that it begins from the familiar goals of healthier, happier, wealthier and longer lives, removing those goals from their human context makes their meaning and, to the extent some of them are instrumental, their purposes, uncertain. The libertarian effort to substitute the goal of diversity for this uncertainty means transhumanism

can only have faith that the future it advocates, a future that may well have no room for human beings, will be desirable. To the extent that transhuman and posthuman diversity is achieved by negation of whatever is given, it appears in fact to represent a variety of nihilism.

The problems inherent in transhumanism should make us look again at what it means to help people lead better human lives, and question whether this project has in fact become obsolete in the face of our latest scientific and technical achievements. The supposed necessity, in light of our powers, that we take the path to posthumanity is really a result of a failure to make moral distinctions based on a substantive picture of the human good, and to think that such distinctions can matter. In his brilliant science fiction novel *The Diamond Age*, Neal Stephenson imagines a world where nano- and information technology open the door to all kinds of transhuman possibilities – some of which are indeed exploited. He departs from the transhumanists, however, in a key premise of the book: as "nearly anything" becomes possible because of the new technology in his fictional world, the "cultural role in deciding what *should* be done with it had become far more important than imagining what *could* be done" (Stephenson 1995: 31). The most successful societies in this world cultivate rather than reject restraint. Transhumanism is too entranced by the "could" to pay serious attention to the "should" beyond assertions that because this transformation is going to happen we better talk about ways to deal with it. But because culture is all about making distinctions between what should and should not be done, Stephenson's science *fiction* is more realistic than the transhumanist *science* fiction about a posthuman world. Transhumanists may be correct that we are on a slippery slope to a new world, but a choice can still be made about joining them in pouring on more oil.

References

Bostrom N (2003a) Transhumanist Values. http://www.transhumanism.org/index.php/WTA/more/transhumanist-values Last accessed 5/23/06

Bostrom N (2003b) The Transhumanist FAQ: A General Introduction (Version 2.1). http://www.transhumanism.org/resources/FAQv21.pdf. Last accessed 12/1/05

Garreau J (2005) Radical Evolution: The Promise and Peril of Enhancing Our Minds, Our Bodies–And What It Means to Be Human Doubleday, New York

Hameroff S, Penrose R (1996) Orchestrated Objective Reduction of Quantum Coherence in Brain Microtubules: The 'Orch OR' Model for Consciousness. www.quantumconsciousness.org/penrose-hameroff/orchOR.html Last accessed 12/12/05

Hobbes T (1968) Leviathan.Penguin Books, Baltimore, MD

Hughes J (2004) Citizen Cyborg: Why Democratic Societies Must Respond to the Redesigned Human Future. Westview, Cambridge

Kurzweil R (2005) The Singularity Is Near. Viking, New York

Kurzweil R, Joy B (2005) Recipe for Destruction. The New York Times, October 17: A23

Mill J S (1969) Nature. In: Essays on Ethics, Religion and Society. University of Toronto Press, Toronto

More M (2003) Principles of Extropy Version 3.11. http://www.extropy.org/principles.htm Last accessed 5/23/06

Naam R (2005) More Than Human: Embracing the Promise of Biological Enhancement. Broadway Books, New York

Pearce D (1998) The Hedonic Imperative. http://www.hedweb.com/hedethic/hedonist.htm Last accessed 5/23/06

Rubin C T (1996) First contact: Copernican moment or nine day's wonder? In: Kingsley S A, Lemarchand G A (eds) The Search for Extraterrestrial Intelligence (SETI) in the Optical Spectrum II. The International Society for Optical Engineering, Bellingham, WA

Stephenson N (1995) The Diamond Age or, A Young Lady's Illustrated Primer. Bantam Books, New York

Todorov T (2002) Imperfect Garden: The Legacy of Humanism. Princeton University Press, Princeton, NJ

Wolfram S (2002) A New Kind of Science. Wolfram Media Champaign, Champaign, Ilinois