A History of Transhumanist Thought

Nick Bostrom
Faculty of Philosophy, Oxford University
www.nickbostrom.com

(2005)

[Originally published in *Journal of Evolution and Technology* - Vol. 14 Issue 1 - April 2005; reprinted (in its present slightly edited form) in *Academic Writing Across the Disciplines*, eds. Michael Rectenwald & Lisa Carl (New York: Pearson Longman, 2011)]

1. Cultural and philosophical antecedents

The human desire to acquire new capacities is as ancient as our species itself. We have always sought to expand the boundaries of our existence, be it socially, geographically, or mentally. There is a tendency in at least some individuals always to search for a way around every obstacle and limitation to human life and happiness.

Ceremonial burial and preserved fragments of religious writings show that prehistoric humans were disturbed by the death of loved ones. Although the belief in a hereafter was common, this did not preclude efforts to extend one's earthly life. In the Sumerian *Epic of Gilgamesh* (approx. 1700 B.C.), a king sets out on a quest for immortality. Gilgamesh learns that there exists a natural means – an herb that grows at the bottom of the sea.[1] He successfully retrieves the plant, but a snake steals it from him before he can eat it. In later times, explorers sought the Fountain of Youth, alchemists labored to concoct the Elixir of Life, and various schools of esoteric Taoism in China strove for physical immortality by way of control over or harmony with the forces of nature. The boundary between mythos and science, between magic and technology, was blurry, and almost all conceivable means to the preservation of life were attempted by somebody or other. Yet while explorers made many interesting discoveries and alchemists invented some useful things, such as new dyes and improvements in metallurgy, the goal of life extension proved elusive.

The quest to transcend our natural confines has long been viewed with ambivalence, however. Reining it in is the concept of *hubris*: that some ambitions are off-limits and will backfire if pursued. The ancient Greeks exhibited this ambivalence in their mythology. Prometheus stole fire from Zeus and gave it to humans, thereby permanently improving the human condition. Yet for this act he was severely punished by Zeus. The gods are repeatedly

[1] (Mitchell 2004).

challenged, quite successfully, by Daedalus, the clever engineer and artist, who uses non-magical means to extend human capabilities. In the end, disaster ensues when his son Icarus ignores paternal warnings and flies too close to the sun, causing the wax in his wings to melt.

Medieval Christians held similarly conflicted views about the pursuits of the alchemists, who were attempting to transmute substances, create homunculi in test tubes, and invent a panacea. Some scholastics, following the anti-experimentalist teachings of Augustine, believed that alchemy was an ungodly activity. There were allegations that it involved the invocation of daemonic powers. But other theologians, such as Albertus Magnus and Thomas Aquinas, defended the practice.[2]

The otherworldliness and stale scholastic philosophy that dominated Europe during the Middle Ages gave way to a renewed intellectual vigor in the Renaissance. The human being and the natural world again became legitimate objects of study. Renaissance humanism encouraged people to rely on their own observations and their own judgment rather than to defer in every matter to religious authorities. Renaissance humanism also created the ideal of the well-rounded person, one who is highly developed scientifically, morally, culturally, and spiritually. A landmark of the period is Giovanni Pico della Mirandola's *Oration on the Dignity of Man* (1486), which proclaims that man does not have a ready-made form and is responsible for shaping himself:

> We have made you a creature neither of heaven nor of earth, neither mortal nor immortal, in order that you may, as the free and proud shaper of your own being, fashion yourself in the form you may prefer. It will be in your power to descend to the lower, brutish forms of life; you will be able, through your own decision, to rise again to the superior orders whose life is divine.[3]

The Age of Enlightenment is often said to have started with the publication of Francis Bacon's *Novum Organum*, "the new tool" (1620), which proposes a scientific methodology based on empirical investigation rather than a-priori reasoning. [4] Bacon advocated the project of "effecting all things possible," by which he meant using science to achieve mastery over nature in order to improve the living condition of human beings. The heritage from the Renaissance combines with the influence of Isaac Newton, Thomas Hobbes, John Locke, Immanuel Kant, the Marquis de Condorcet, and others to form the basis for rational

[2] See e.g. (Newman 2004).

[3] (Pico della Mirandola 1956).

[4] (Bacon 1620).

humanism, which emphasizes empirical science and critical reason – rather than revelation and religious authority – as ways of learning about the natural world and our place within it and of providing a grounding for morality. Transhumanism has roots in rational humanism.

In the 18th and 19th centuries we catch glimpses of the idea that humans themselves can be developed through the application of science. Condorcet speculated about extending human life span by means of medical science:

> Would it be absurd now to suppose that the improvement of the human race should be regarded as capable of unlimited progress? That a time will come when death would result only from extraordinary accidents or the more and more gradual wearing out of vitality, and that, finally, the duration of the average interval between birth and wearing out has itself no specific limit whatsoever? No doubt man will not become immortal, but cannot the span constantly increase between the moment he begins to live and the time when naturally, without illness or accident, he finds life a burden?[5]

Benjamin Franklin longed wistfully for suspended animation, foreshadowing the cryonics movement:

> I wish it were possible... to invent a method of embalming drowned persons, in such a manner that they might be recalled to life at any period, however distant; for having a very ardent desire to see and observe the state of America a hundred years hence, I should prefer to an ordinary death, being immersed with a few friends in a cask of Madeira, until that time, then to be recalled to life by the solar warmth of my dear country! But... in all probability, we live in a century too little advanced, and too near the infancy of science, to see such an art brought in our time to its perfection.[6]

After the publication of Darwin's *Origin of Species* (1859), it became increasingly plausible to view the current version of humanity not as the endpoint of evolution but rather as an early phase.[7] The rise of scientific physicalism might also have contributed to the belief that technology might well improve the human organism. For example, a simple kind of materialist view was boldly proposed in 1750 by the French physician and materialist philosopher Julien Offray de La Mettrie in *L'Homme Machine*, in which he argued that "man

[5] (Condorcet 1979).

[6] (Franklin et al. 1956), pp. 27-29.

[7] (Darwin 2003).

is but an animal, or a collection of springs which wind each other up."[8] If human beings are constituted of matter obeying the same laws of physics that operate outside us, then it should in principle be possible to learn to manipulate human nature in the same way that we manipulate external objects.

The Enlightenment is said to have expired as the victim of its own excesses. It gave way to Romanticism and to latter-day reactions against the rule of instrumental reason and the attempt to rationally control nature, such as can be found in some postmodernist writings, the New Age movement, deep environmentalism, and parts of the anti-globalization movement. However, the Enlightenment's legacy, including a belief in the power of human rationality and science, is still an important shaper of modern culture. In his famous 1784 essay "What Is Enlightenment?", Kant summed it up as follows:

> Enlightenment is man's leaving his self-caused immaturity. Immaturity is the incapacity to use one's own understanding without the guidance of another. Such immaturity is self-caused if its cause is not lack of intelligence, but by lack of determination and courage to use one's intelligence without being guided by another. The motto of enlightenment is therefore: *Sapere aude!* Have courage to use your own intelligence![9]

It might be thought that a major inspiration for transhumanism was Friedrich Nietzsche, famous for his doctrine of *der Übermensch*:

> I teach you the overman. Man is something that shall be overcome. What have you done to overcome him? All beings so far have created something beyond themselves; and do you want to be the ebb of this great flood and even go back to the beasts rather than overcome man?[10]

What Nietzsche had in mind, however, was not technological transformation but a kind of soaring personal growth and cultural refinement in exceptional individuals (who he thought would have to overcome the life-sapping "slave-morality" of Christianity). Despite some surface-level similarities with the Nietzschean vision, transhumanism – with its Enlightenment roots, its emphasis on individual liberties, and its humanistic concern for the welfare of all humans (and other sentient beings) – probably has as much or more in

[8] (La Mettrie 1996).
[9] (Kant 1986).
[10] (Nietzsche 1908).

common with Nietzsche's contemporary the English liberal thinker and utilitarian John Stuart Mill.

2. Speculation, science fiction, and twentieth-century totalitarianism

In 1923, the noted British biochemist J. B. S. Haldane published the essay *Daedalus; or, Science and the Future*, in which he argued that great benefits would come from controlling our own genetics and from science in general. He predicted a wealthier society, with abundant clean energy, where genetics would be employed to make people taller, healthier, and smarter and where ectogenesis (gestating fetuses in artificial wombs) would be commonplace. He also commented on what has in recent years become known as the "yuck factor":

> The chemical or physical inventor is always a Prometheus. There is no great invention, from fire to flying, which has not been hailed as an insult to some god. But if every physical and chemical invention is a blasphemy, every biological invention is a perversion. There is hardly one which, on first being brought to the notice of an observer from any nation which has not previously heard of their existence, would not appear to him as indecent and unnatural.[11]

Haldane's essay became a bestseller and set off a chain of future-oriented discussions, including *The World, the Flesh and the Devil*, by J. D. Bernal (1929)[12], which speculated about space colonization and bionic implants as well as mental improvements arising from advanced social science and psychology; the works of Olaf Stapledon, a philosopher and science fiction author; and the essay "Icarus: the Future of Science" (1924) by Bertrand Russell.[13] Russell took a more pessimistic view, arguing that without more kindliness in the world, technological power would mainly serve to increase our ability to harm one another. Science fiction authors such as H. G. Wells and Stapledon got many people thinking about the future evolution of the human race.

Aldous Huxley's *Brave New World*, published in 1932, has had an enduring impact on debates about human technological transformation[14] matched by few other works of fiction (possibly excepting Mary Shelley's *Frankenstein*[15]). Huxley describes a dystopia where psychological conditioning, promiscuous sexuality, biotechnology, and the opiate drug "soma" keep the population placid in a static, conformist caste society governed by ten

---

[11] (Haldane 1924).
[12] (Bernal 1929[1969]).
[13] (Russell 1924)
[14] (Huxley 1932).
[15] (Shelley 1818).

world controllers. Children are manufactured in fertility clinics and artificially gestated. The lower castes are chemically stunted or deprived of oxygen during their maturation process, in order to limit their physical and intellectual development. From birth, members of every caste are indoctrinated during their sleep, by recorded voices repeating the slogans of the official "Fordist" religion, and conditioned to believe that their own caste is the best one to belong to. The society depicted in *Brave New World* is often likened to another influential 20th century dystopia, that of George Orwell's *Nineteen Eighty-Four*.[16] *Nineteen Eighty-Four* features a more overt form of oppression, including ubiquitous surveillance by "Big Brother" and brutal police coercion. Huxley's world controllers, by contrast, rely on less blatant means (bio-engineered predestination, psychological conditioning, soma) to prevent people from *wanting* to think for themselves. Herd mentality and promiscuity are promoted, while high art, individuality, knowledge of history, and romantic love are discouraged. It should be noted that in neither *Nineteen Eighty-Four* nor *Brave New World* is technology employed to increase human capacities; rather, society is set up to repress the full development of humanity. Both dystopias curtail scientific and technological exploration for fear of upsetting the social equilibrium. Nevertheless, *Brave New World* in particular has become an emblem of the dehumanizing potential of the use of technology to promote social conformism and shallow contentment.

In the early decades of the 20th century, not only racists and right-wing ideologues but also a number of left-leaning social progressives became concerned about the effects of medicine and social safety nets on the quality of the human gene pool. They believed that modern society enabled many "unfit" individuals to survive—individuals who would in earlier ages have perished—and they worried that this would lead to a deterioration of the human stock. As a result, many countries (including the USA, Canada, Australia, Sweden, Denmark, Finland, and Switzerland) implemented state-sponsored eugenics programs, which infringed in various degree on individual rights. In the United States, between 1907 and 1963 some 64,000 individuals were forcibly sterilized under eugenics laws. The principal victims of the American program were the mentally disabled, but the deaf, the blind, the epileptic, the physically deformed, orphans, and the homeless were also sometimes targeted. But even such widespread compulsory sterilization pales in comparison with the German eugenics program, which resulted in the systematic murder of millions of people regarded as "inferior" by the Nazis.

The holocaust left a scar on the human psyche. Determined not to let history repeat itself, most people developed an instinctive revulsion to all ideas appearing to have any kind of association with Nazi ideology. (And yet, it must be remembered, history did repeat itself,

[16] (Orwell 1949).

e.g. in the Rwandan genocide of 1994, in which the world did nothing but wring its hands as 800,000 Africans were slaughtered.) In particular, the eugenics movement as a whole, in all its forms, became discredited because of the terrible crimes that had been committed in its name, although some of the milder eugenics programs continued for many years before they were finally scrapped. These programs are all now almost universally condemned. The goal of creating a new and better world through a centrally imposed vision became passé. The Stalinist tyranny, too, underscored the dangers of totalitarian utopianism.

In the postwar era, many optimistic futurists who had become suspicious of collectively orchestrated social change found a new home for their hopes in scientific and technological progress. Space travel, medicine, and computers seemed to offer a path to a better world. The shift of attention also reflected the breathtaking pace of development in these fields. Science had begun to catch up with speculation. Transhumanist themes during this period were discussed and analyzed chiefly in the science fiction literature. Authors such as Arthur C. Clarke, Isaac Asimov, Robert Heinlein, and Stanislaw Lem explored how technological development could come to profoundly alter the human condition.

The word "transhumanism" appears to have been first used by Aldous Huxley's brother, Julian Huxley, a distinguished biologist (who was also the first director-general of UNESCO and a founder of the World Wildlife Fund). In *New Bottles for New Wine* (1957), he wrote:

> The human species can, if it wishes, transcend itself – not just sporadically, an individual here in one way, an individual there in another way – but in its entirety, as humanity. We need a name for this new belief. Perhaps *transhumanism* will serve: man remaining man, but transcending himself, by realizing new possibilities of and for his human nature.[17]

3. Technological genies: AI, the singularity, nanotech, and uploading

Human-like automata have always fascinated the human imagination. Mechanical engineers since the early Greeks have constructed clever self-moving devices.

In Judaic mysticism, a "golem" refers to an animated being crafted from inanimate material. In the early golem stories, a golem could be created by a holy person who was able to share some of God's wisdom and power (although the golem, not being able to speak, was never more than a shadow of God's creations). Having a golem servant was the ultimate symbol of

[17] (Huxley 1957).

wisdom and holiness. In the later stories, which were influenced by the more Islamic concern about humanity**'s** getting too close to God, the golem became a creation of overreaching mystics who were inevitably punished for their blasphemy. The story of the Sorcerer's Apprentice is a variation of this theme: The apprentice animates a broomstick to fetch water but is unable to make the broom stop – like *Frankenstein*, it is, a story of technology out of control. The word "robot" was coined by the Czech writer Karel Čapek, in his dark play *R.U.R.* (1921), in which a robot labor force destroys its human creators.[18] With the invention of the electronic computer, the idea of human-like automata graduated from the kindergarten of mythology to the school of science fiction (Asimov, Lem, Clarke) and eventually to the college of technological prediction.

Could continued progress in artificial intelligence lead to the creation of machines that think in the same general way as human beings? Alan Turing gave an operational definition to this question in his classic "Computing Machinery and Intelligence" (1950) and predicted that computers would eventually pass what came to be known as the Turing Test. In the Turing Test, a human experimenter interviews a computer and another human via a text interface**;** the computer succeeds if the interviewer cannot reliably distinguish it from the human.)[19] Much ink has been spilt in debates on whether this test furnishes a necessary and sufficient condition for a computer being able to think, but what matters more from a practical perspective is whether and, if so, *when* computers will be able to match human performance in tasks involving general reasoning ability. With the benefit of hindsight, we can say that many of the early AI researchers turned out to be overoptimistic about the timescale for this hypothetical development. Of course, the fact that we have not yet reached human-level artificial intelligence does not mean that we never will, and a number of people, e.g. Marvin Minsky, Hans Moravec, Ray Kurzweil, and Nick Bostrom have put forward reasons for taking seriously the possibility that this could happen within the first half of this century.[20]

In a 1958 tribute, the Polish mathematician Stanislaw Ulam, referring to a meeting with his late colleague John von Neumann, wrote:

> One conversation centered on the ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue.[21]

---

[18] (Capek 2004).

[19] (Turing 1950).

[20] (Minsky 1994; Moravec 1999; Bostrom 1998, 2002; Kurzweil 1999).

[21] (Ulam 1958).

The rapidity of technological change in recent times leads naturally to the idea that continued technological innovation will have a profound effect on humanity in the decades ahead. This prediction is strengthened if one believes that some of those variables that currently exhibit exponential growth will continue to do so and that they will be among the main drivers of change. Gordon E. Moore, co-founder of Intel, noticed in 1965 that the number of transistors on a chip exhibited exponential growth. This led to the formulation of "Moore's law", which states (roughly) that computing power doubles every eighteen months to two years.[22] More recently, Kurzweil has documented similar exponential growth rates in a number of other technologies. It is interesting to note that the world economy, a general index of humanity's productive capacity, has doubled about every fifteen years in modern times.

The singularity hypothesis, to which von Neumann seemingly alludes in the quoted passage above, holds that these changes will lead to some kind of discontinuity. But nowadays it often refers to a more specific prediction**:** namely**,** that the creation of self-improving artificial intelligence will at some point result in radical changes within a very short time span. This hypothesis was first clearly stated in 1965 by the statistician I. J. Good:

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an 'intelligence explosion,' and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make.[23]

Vernor Vinge discussed this idea in a little more detail in his influential 1993-paper "Technological Singularity", in which he predicted:

> Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended.[24]

Transhumanists today hold diverging views about the singularity: some see it as a likely scenario, others believe that it is more probable that there will never be any very sudden and dramatic changes as the result of progress in artificial intelligence.

---

[22] (Moore 1965).

[23] (Good 1965).

[24] (Vinge 1993).

The singularity idea also comes in a somewhat different eschatological version, which traces its lineage to the writings of Pierre Teilhard de Chardin, a paleontologist and Jesuit theologian who saw an evolutionary telos in the development of an encompassing noosphere (a global consciousness) – via physicist Frank Tipler, who argued that advanced civilizations might come to have a defining influence on the future evolution of the cosmos, and, in the final moments of the Big Crunch, might manage to extract an infinite number of computations by harnessing the sheer energy of the collapsing matter.[25,26] However, while these ideas might appeal to those who fancy a marriage between mysticism and science, they have not caught on either among transhumanists or the larger scientific community. Current cosmological theories indicate that the universe will continue to expand forever (falsifying Tipler's prediction). But the more general point that the transhumanist might make in this context is that we need to learn to think about "big-picture questions" without resorting to wishful thinking or mysticism. Big-picture questions, including ones about our place in the world and the long-term fate of intelligent life *are* part of transhumanism; however, these questions should be addressed in a sober, disinterested way, using critical reason and our best available scientific evidence. One reason why such questions are of transhumanist interest is that their answers might affect what outcomes we should expect from our own technological development, and therefore – indirectly – what policies it makes sense for humanity to pursue.

In 1986, Eric Drexler published *Engines of Creation*, the first book-length exposition of molecular manufacturing.[27] (The possibility of nanotechnology had been anticipated by Nobel laureate physicist Richard Feynman in his famous after-dinner address in 1959 entitled "There is Plenty of Room at the Bottom".[28]) In this seminal work, Drexler not only argued for the feasibility of assembler-based nanotechnology but also explored its consequences and began charting the strategic challenges posed by its development. Drexler's later book *Nanosystems* (1992) supplied a more technical analysis that seemed to confirm his original conclusions.[29] To prepare the world for nanotechnology and work towards its safe implementation, he founded the Foresight Institute together with his then wife, Christine Peterson, in 1986.

---

[25] (Teilhard de Chardin 1964).
[26] (Tipler 1994).
[27] (Drexler 1985).
[28] (Feynman 1960).
[29] (Drexler 1992).

In the last several years, nanotechnology has become big business, with worldwide research funding amounting to billions of dollars. Yet little of this work fits Drexler's ambitious vision of nanotechnology as an assembler-based, near-universal, construction technology. The mainstream nanotechnology community has sought to distance itself from Drexler's claims. The chemist Richard Smalley (another Nobel laureate) has debated Drexler, asserting that non-biological molecular assemblers are impossible.[30] To date, however, no technical critique of Drexler's work in the published literature has found any significant flaws in his reasoning. If molecular nanotechnology is indeed physically possible, as Drexler maintains, the question becomes just how difficult it will be to develop it, and how long it will take. These issues are very difficult to settle in advance.

If molecular nanotechnology could be developed as Drexler envisions it, it would have momentous ramifications:

> Coal and diamonds, sand and computer chips, cancer and healthy tissue: throughout history, variations in the arrangement of atoms have distinguished the cheap from the cherished, the diseased from the healthy. Arranged one way, atoms make up soil, air, and water arranged another, they make up ripe strawberries. Arranged one way, they make up homes and fresh air; arranged another, they make up ash and smoke.[31]

Molecular nanotechnology would enable us to transform coal into diamonds, sand into supercomputers, and to remove pollution from the air and tumors from healthy tissue. In its mature form, it could help us abolish most disease and aging, make possible the reanimation of cryonics patients, enable affordable space colonization, and – more ominously – lead to the rapid creation of vast arsenals of lethal or non-lethal weapons.

Another hypothetical technology that would have a revolutionary impact is uploading, the transfer of a human mind to a computer. This would involve the following steps: First, create a sufficiently detailed scan of a particular human brain, perhaps by deconstructing it with nanobots or by feeding thin slices of brain tissues into powerful microscopes for automatic image analysis. Second, from this scan, reconstruct the neuronal network that the brain implemented, and combine this with computational models of the different types of neurons. Third, emulate the whole computational structure on a powerful supercomputer. If successful, the procedure would result in the original mind, with memory and personality intact, being transferred to the computer where it would then exist as software; and it could

---

[30] (Drexler and Smalley 1993).

[31] (Drexler 1985), p. 3.

either inhabit a robot body or live in a virtual reality.[32] While it is often thought that, under suitable circumstances, the upload would be conscious and that the original person would have survived the transfer to the new medium, individual transhumanists take different views on these philosophical matters.

If either superintelligence, or molecular nanotechnology, or uploading, or some other technology of a similarly revolutionary kind is developed, the human condition could clearly be radically transformed. Even if one believed that the probability of this happening any time soon is quite small, these prospects would nevertheless merit serious attention in view of their extreme impact. However, transhumanism does not depend on the feasibility of such radical technologies. Virtual reality; preimplantation genetic diagnosis; genetic engineering; pharmaceuticals that improve memory, concentration, wakefulness, and mood; performance-enhancing drugs; cosmetic surgery; sex change operations; prosthetics; anti-aging medicine; closer human-computer interfaces: these technologies are already here or can be expected within the next few decades. The combination of these technological capabilities, as they mature, could profoundly transform the human condition. The transhumanist agenda, which is to make such enhancement options safely available to all persons, will become increasingly relevant and practical in the coming years as these and other anticipated technologies come online.

4. The growth of grassroots

Benjamin Franklin wished to be preserved in a cask of Madeira and later recalled to life, and regretted that he was living too near the infancy of science for this to be possible. Since then, science has grown up a bit. In 1962, Robert Ettinger published the book, *The Prospect of Immortality*, which launched the idea of cryonic suspension.[33] Ettinger argued that as medical technology seems to be constantly progressing, and since science has discovered that chemical activity comes to a complete halt at low-enough temperatures, it should be possible to freeze a person today (in liquid nitrogen) and preserve the body until a time when technology is advanced enough to repair the freezing damage and reverse the original cause of deanimation. Cryonics, Ettinger believed, offered a ticket to the future.

Alas, the masses did not line up for the ride. Cryonics has remained a fringe alternative to more traditional methods of treating the diseased, such as cremation and burial. The practice of cryonics was not integrated into the mainstream clinical setting and was instead conducted on the cheap by a small number of enthusiasts. Two early cryonics organizations

[32] (Bostrom 2003).

[33] (Ettinger 1964).

went bankrupt, allowing their patients to thaw out. At that point, the problem of massive cellular damage that occurs when ice crystals form in the body also became more widely known. As a result, cryonics acquired a reputation as a macabre scam. The media controversy over the suspension of baseball star Ted Williams in 2002 showed that public perception of cryonics has not changed much over the past decades.

Despite its image problem and its early failures of implementation, the cryonics community continues to be active and it counts among its members several eminent scientists and intellectuals. Suspension protocols have been improved, and the infusion of cryoprotectants prior to freezing to suppress the formation of ice crystals has become standard practice. The prospect of nanotechnology has given a more concrete shape to the hypothesized future technology that could enable reanimation. There are currently two organizations that offer full-service suspension, the Alcor Life Extension Foundation (founded in 1972) and the Cryonics Institute (founded in 1976). Alcor has recently introduced a new suspension method, which relies on a process known as "vitrification", which further reduces micro-structural damage during suspension.

In a later work, *Man into Superman* (1972), Ettinger discussed a number of conceivable technological improvements of the human organism, continuing the tradition started by Haldane and Bernal.[34]

Another early transhumanist was F. M. Esfandiary, who later changed his name to FM-2030. One of the first professors of future studies, FM taught at the New School for Social Research in New York in the 1960s and formed a group of optimistic futurists known as the UpWingers.

> Who are the new revolutionaries of our time? They are the geneticists, biologists, physicists, cryonologists, biotechnologists, nuclear scientists, cosmologists, radio astronomers, cosmonauts, social scientists, youth corps volunteers, internationalists, humanists, science-fiction writers, normative thinkers, inventors… They and others are revolutionizing the human condition in a fundamental way. Their achievements and goals go far beyond the most radical ideologies of the Old Order.[35]

In his book *Are you a transhuman?* (1989), FM described what he regarded as the signs of the emergence of the "transhuman".[36] In FM's terminology, a transhuman is a "transitional

[34] (Ettinger 1972).
[35] (Esfandiary 1970).
[36] (FM-2030 1989).

human," someone who by virtue of their technology usage, cultural values, and lifestyle constitutes an evolutionary link to the coming era of posthumanity. The signs that FM saw as indicative of transhuman status included prostheses, plastic surgery, intensive use of telecommunications, a cosmopolitan outlook and a globetrotting lifestyle, androgyny, mediated reproduction (such as in vitro fertilization), absence of religious belief, and a rejection of traditional family values. However, it was never satisfactorily explained why somebody who, say, rejects family values, has a nose job, and spends a lot of time on jet planes is in closer proximity to posthumanity than the rest of us.

In the 1970s and 1980s, many organizations sprang up that focused on a particular topic such as life extension, cryonics, space colonization, science fiction, and futurism. These groups were often isolated from one another, and whatever shared views and values they had did not yet amount to any unified worldview. Ed Regis's *Great Mambo Chicken and the Transhuman Condition* (1990) took a humorous look at these proto-transhumanist fringes, which included eccentric and otherwise intelligent individuals trying to build space rockets in their backyards or experimenting with biofeedback machines and psychedelic drugs, as well as scientists pursuing more serious lines of work but who had imbibed too deeply of the Californian spirit.[37]

In 1988, the first issue of the *Extropy Magazine* was published by Max More and Tom Morrow, and in 1992 they founded the Extropy Institute (the term "extropy" being coined as a metaphorical opposite of entropy). The Institute served as a catalyst that brought together disparate groups of people with futuristic ideas and facilitated the formation of novel memetic compounds. The Institute ran a series of conferences, but perhaps most important was the extropians mailing list, an online discussion forum where new ideas were shared and debated. In the mid-nineties, many got first exposure to transhumanist views from the Extropy Institute's listserve.

More had immigrated to California from Britain after changing his name from Max O'Connor. Of his new name, he said:

> It seemed to really encapsulate the essence of what my goal is: always to improve, never to be static. I was going to get better at everything, become smarter, fitter, and healthier. It would be a constant reminder to keep moving forward.[38]

[37] (Regis 1990).
[38] (Regis 1994).

Max More wrote the first definition of transhumanism in its modern sense, and created his own distinctive brand of transhumanism, "extropianism," which emphasized the principles of "boundless expansion," "self-transformation," "dynamic optimism," "intelligent technology," and "spontaneous order". Originally, extropianism had a clear libertarian flavor, but in later years More has distanced himself from this ingredient, replacing "spontaneous order" with "open society," a principle that opposes authoritarian social control and promotes decentralization of power and responsibility.[39]

Natasha Vita-More is the Extropy Institute's current president. She is an artist and designer, and has over the years issued a number of manifestos on transhumanist and extropic art.[40]

The Extropy Institute's conferences and mailing list also served as a hangout place for some people who liked to discuss futuristic ideas but who were not necessarily joiners. Those who were around in the mid-nineties will remember individuals such as Anders Sandberg, Alexander "Sasha" Chislenko, Hal Finney, and Robin Hanson from among the more thoughtful regulars in the transhumanist milieu at the time. An enormous amount of discussion about transhumanism has taken place on various email lists in the past decade. The quality of postings has been varied (putting it mildly). Yet at their best, these online conversations explored ideas about the implications of future technologies that were, in some respects, far advanced over what could be found in printed books or journals. The Internet played an important role in incubating modern transhumanism by facilitating these meetings of minds – and perhaps more indirectly, too, via the "irrational exuberance" that pervaded the dot-com era?

The World Transhumanist Association was founded in early 1998 by Nick Bostrom and David Pearce, to provide a general organizational basis for all transhumanist groups and interests, across the political spectrum. The aim was also to develop a more mature and academically respectable form of transhumanism, freed from the "cultishness" which, at least in the eyes of some critics, had afflicted some of its earlier convocations. The two founding documents of the WTA were the *Transhumanist Declaration* (see appendix), and the *Transhumanist FAQ* (v. 1.0).[41] The Declaration was intended as a concise consensus statement of the basic principle of transhumanism. The FAQ was also a consensus or near-consensus document, but it was more ambitious in its philosophical scope in that it developed a number of themes that had previously been, at most, implicit in the movement. More than fifty people contributed comments on drafts of the FAQ. The document was produced by

[39] (More 2003).
[40] (Vita-More 2002).
[41] (WTA 2002).

Bostrom but major parts and ideas were also contributed by several others, including the British utilitarian thinker David Pearce, Max More, the American feminist and disability rights activist Kathryn Aegis, and the walking encyclopedia Anders Sandberg, who was at the time a neuroscience student in Sweden.

David Pearce has also developed his own distinctive flavor of transhumanism based on an ethic of hedonistic utilitarianism. Pearce argues, in *The Hedonistic Imperative*, for an ambitious program to eliminate suffering in both human and non-human animals by means of advanced neuro-technology (in the short term pharmaceuticals, in the longer term perhaps genetic engineering).[42] In parallel with this negative effort to abolish suffering, he proposes a positive program of "paradise engineering" in which sentient beings would be redesigned to enable everybody to experience of unprecedented levels of well-being. In Pearce's utopia, our motivation system would run on "gradients of bliss" instead of the current pleasure-pain axis.

The WTA's membership grew rapidly, and local chapters mushroomed around the world. Activities focused mainly on Internet discussion, development of documents, representation in the media, organizing of an annual TransVision conference, and publication of the scholarly online Journal of Transhumanism (later renamed to "Journal of Evolution and Technology").

In the first few years of its existence, the WTA was a very loosely and informally organized structure. It entered its next phase after a meeting in 2001 between James Hughes (a sociologist at Trinity College in Hartford Connecticut), Mark Walker (a philosopher at the University of Toronto, then the editor of the Journal of Transhumanism), and Bostrom (who was at the time teaching at Yale). Hughes was elected Secretary and turned his organizing skills and energy to the task. Within short order, the WTA adopted a constitution, incorporated as a non-profit, and began building up a vigorous international network of local groups and volunteers. Currently, the WTA has approximately 3,000 members from more than 100 countries, and it pursues a wide range of activities, all volunteer-driven.

A number of related organizations have also cropped up in recent years, focusing more narrowly on particular transhumanist issues, such as life-extension, artificial intelligence, or the legal implications of "converging technologies" (nano-bio-info-neuro technologies). The Institute for Ethics and Emerging Technologies, a non-profit think tank, was established in 2004, to "promote the ethical use of technology to expand human capacities".

---

[42] (Pearce 2004).

5. The academic frontier

Over the past couple of decades, academia has picked up the ball and started to analyze various "transhumanist matters," both normative and positive. The contributions are far too many to comprehensively describe here, so we will pick out just a few threads, beginning with ethics.

For most of its history, moral philosophy did not shy away from addressing practical problems. In the early and mid-parts of the twentieth century, during heydays of logical positivism, applied ethics became a backwater as moral philosophers concentrated on linguistic or meta-ethical problems. Since then, however, practical ethics has reemerged as a field of academic inquiry. The comeback started in medical ethics. Revelations of the horrific experiments that the Nazis had conducted on human subjects in the name of science led to the adoption of the Nuremberg code (1947) and the Declaration of Helsinki (1964), which laid down strict safeguards for medical experimentation, emphasizing the need for patient consent.[43,44] But the rise of the modern health care system spawned new ethical dilemmas – turning off life-support, organ donation, resource allocation, abortion, advance directives, doctor-patient relationships, protocols for obtaining informed consent and for dealing with incompetent patients. In the 1970s, a broader kind of enquiry began to emerge, stimulated particularly by developments in assisted reproduction and genetics. This field became known as bioethics. Many of the ethical issues most directly linked to transhumanism would now fall under this rubric, although other normative discourses are also involved, e.g. population ethics, meta-ethics, political philosophy, and bioethics' younger sisters – computer ethics, engineering ethics, environmental ethics.

Bioethics was from the beginning an interdisciplinary endeavor, dominated by theologians, legal scholars, physicians, and, increasingly, philosophers, with occasional participation by representatives of patients' rights groups, disability advocates, and other interested parties.[45] Lacking a clear methodology, and operating on a plain often swept by the winds of political or religious controversy, the standard of scholarship has frequently been underwhelming. Despite these difficulties, bioethics burgeoned. A cynic might ascribe this accomplishment to the ample fertilization that the field received from a number of practical imperatives: absolving doctors of moral dilemmas, training medical students to behave, enabling hospital boards to trumpet their commitment to the highest ethical standards of care, providing sound bites for the mass media, and allowing politicians to cover their behinds by delegating controversial issues to ethics committees. But a kinder gloss is

[43] (Office 1949).

[44] (World_Medical_Organization 1996).

[45] See (Jonsen 1998).

possible: decent people recognized that difficult moral problems arose in modern biomedicine, that these problems needed to be addressed, and that having some professional scholars trying to clarify these problems in some sort of systematic way might be helpful. While higher-caliber scholarship and a more robust methodology would be nice, in the meantime we make the most of what we have.

Moral philosophers have in the last couple of decades made many contributions that bear on the ethics of human transformation, and we must limit ourselves to a few mentions. Derek Parfit's classic *Reasons and Persons* (1984) discussed many relevant normative issues.[46] In addition to personal identity and foundational ethical theory, this book treats population ethics, person-affecting moral principles, and duties to future generations. Although Parfit's analysis takes place on an idealized level, his arguments elucidate many moral considerations that emerge within the transhumanist program.

Jonathan Glover's *What Sort of People Should there Be?* (1984) addressed technology-enabled human-transformation at a somewhat more concrete level, focusing especially on genetics and various technologies that could increase social transparency. Glover gave a clear and balanced analytic treatment of these issues that was well ahead of its time. His general conclusion is that

> not just any aspect of present human nature… is worth preserving. Rather it is especially those features which contribute to self-development and self-expression, to certain kinds of relationships, and to the development of our consciousness and understanding. And some of these features may be extended rather than threatened by technology.[47]

Several people have argued for principles that assert some kind of ethical equivalence between environmental and genetic interventions. For example, Peter Singer has proposed the "preventive principle":

> For any condition X, if it would be a form of child abuse for parents to inflict X on their child soon after birth, then it must, other things being equal, at least be permissible to take steps to prevent one's child having that condition.[48]

---

[46] (Parfit 1984).

[47] (Glover 1984).

[48] (Singer 2003).

Julian Savulescu has argued for a principle of Procreative Beneficence, according to which prospective parents should select the child, of the possible children they could have, who would have the best life, based on the relevant, available information (where the "should" is meant to indicate that persuasion is justified, but not coercion).[49] This principle does not presuppose that all lives can be placed in a definite ranking with respect to their well-being, only that pair-wise comparisons are possible in at least some cases. For instance, if a couple is having IVF and must select one of two embryos which are genetically identical except that one of them has one defective gene that predisposes to asthma, then Procreative Beneficence suggests they ought to choose the healthy embryo for implantation.

In *From Chance to Choice* (2000), Allen Buchanan, Dan W. Brock, Norman Daniels, and Daniel Wikler, examined how advances in genetic engineering should affect our understanding of distributive justice, equal opportunity, our rights and obligations as parents, the meaning of disability, and the concept of human nature in ethical theory and practice.[50] They developed a framework inspired by John Rawls's work in an attempt to answer some of these questions.

Greg Stock, John Harris, Gregory Pence, and Eric Juengst, among others, have also discussed the ethics of genetic engineering from a broadly transhumanist perspective.[51] Mark Walker has argued from a perfectionist standpoint that we have a duty to use technology to improve ourselves. Walker has also argued that one reason to pursue cognitive enhancements is that it could help us solve philosophical problems.[52] Nick Bostrom and several others have drawn attention to the distinction between enhancements that offer only positional advantages (e.g. an increase in height), which are only advantages insofar as others lack them, and enhancements that provide either intrinsic benefits or net positive externalities (such as a better immune system or improvement of cognitive functioning). We ought to promote enhancements of the second kind, but not enhancements that are merely positional.[53]

Bostrom has suggested that we have a reason to develop means to explore the "larger space of possible modes of being" that is currently inaccessible to us because of our biological limitations, on the ground that we might find that it contains extremely worthwhile modes of being – ways of living, thinking, feeling, and relating.[54] Along with many other

[49] (Savulescu 2001).
[50] (Buchanan et al. 2002).
[51] E.g. (Stock 2002; Harris 1992; Pence 1998; Parens 1998).
[52] (Walker 2002).
[53] (Bostrom 2003).
[54] (Bostrom 2004).

transhumanist writers, Bostrom has argued for the moral urgency of developing means to slow or reverse the aging process.[55] He has also proposed a broader conception of human dignity which can accommodate "posthuman dignity".[56] A recent joint paper by Bostrom and Toby Ord proposes a heuristic for eliminating "status quo" bias in bioethics, a bias which, they claim, afflicts many of our moral intuitions.[57]

Eliezer Yudkowsky (an independent scholar) has probed the ethics of superintelligence and has tried to develop a theory of how to program a human-friendly AI, a challenge that could take on life-and-death significance once we become capable of creating such a machine. Yudkowsky argues that simple rule-based injunctions (such as Isaac Asimov's "three laws of robotics") would produce deadly unintended consequences. He conceives of a superintelligence as an enormously powerful optimization process, and the central task is to specify the mental architecture and goal-structure of the AI in such a way that it realizes desirable outcomes. Rather than creating a list of specific goals, Yudkowsky argues that we need to take a more indirect approach and choose the AI's initial conditions so that it would use its superior intellectual powers to derive the specific goals and extrapolate our decisions if we were better calibrated, better informed, and better able to reflect on the forces influencing our decisions. Yudkowsky also wishes to specify an AI that would use its initial rules for extrapolation to extrapolate smarter human decisions about extrapolation rules; in effect, a set of initial rules for extrapolation would "renormalize" themselves.[58]

Aside from normative questions, there are also positive questions to be asked, about the nature and timing of transforming technologies and their consequences. Hans Moravec's 1989-book *Mind Children* explored the ramifications of possible future advances in robotics and uploading.[59] A later Moravec book, *Robot* (1999), and Ray Kurzweil's best-selling *Age of Spiritual Machines* (1999) introduced these ideas to a wider audience.[60,61] As we have seen, Eric Drexler was trying to anticipate the consequences of molecular nanotechnology back in the 80s, an endeavor in which he has since been joined by several other researchers such as Robert Freitas, who has studied potential medical application of nanotechnology in great detail, and Ralph Merkle who has collaborated with Freitas to study the kinematics of self-replicating systems and the technical steps towards crude molecular assemblers.[62] All these

[55] (Bostrom 2005).
[56] (Bostrom 2005).
[57] (Bostrom and Ord 2005).
[58] (Yudkowsky 2004).
[59] (Moravec 1989).
[60] (Moravec 1999).
[61] (Kurzweil 1999).
[62] (Freitas and Merkle 2005).

authors recognize that technologies as potent as superintelligence or molecular nanotechnology are not without serious risks of accidents or deliberate misuse.

Bostrom (2002) introduced the concept of an "existential risk", defined as "one where an adverse outcome would either annihilate Earth-originating intelligent life or permanently and drastically curtail its potential", and created a catalogue of what he saw as the most probable existential risks.[63] Both nanotechnology- and superintelligence-related risks attain high ranks on that list. In a much-discussed popular article, "Why the Future Doesn't Need Us" (2000), Bill Joy argued that we ought to relinquish developments in AI, nanotechnology, and genetics because of the risks that will eventually emerge from these disciplines.[64] Several people, reacting to Joy, argued against such bans on grounds that they are unrealistic, would deprive us of great benefits, and might increase rather than decrease risk if development were driven underground or to less hesitant regions of the world. John Leslie, Martin Rees, and Richard Posner have also investigated threats to human survival in the 21st century – all of them have rated the risk as highly significant.[65]

Robin Hanson has analyzed several topics of relevance to human transformation, including the consequences of uploading in an unregulated economy, the social-signaling function of beliefs, the sources and epistemological status of disagreements of opinion, the dynamics of a space colonization race, and information markets as a system for aggregating information and guiding policy.[66] Related to Hanson's work on upload competition and colonization races, Bostrom has explored how dystopian outcomes could result in some future evolutionary scenarios.[67] Drawing on his earlier work on observation selection effects, he also formulated the Simulation argument, which purports to show that it follows from some fairly weak assumptions that

> at least one of the following propositions is true: (1) the human species is very likely to go extinct before reaching a "posthuman" stage; (2) any posthuman civilization is extremely unlikely to run a significant number of simulations of their evolutionary history (or variations thereof); (3) we are almost certainly living in a computer simulation. It follows that the belief that there is a significant chance that we will one day become posthumans who run ancestor-simulations is false, unless we are currently living in a simulation.[68]

---

[63] (Bostrom 2002).

[64] (Joy 2000).

[65] (Leslie 1996; Rees 2003; Posner 2004).

[66] E.g. (Hanson 1994, 1995, 1998).

[67] (Bostrom 2005).

[68] (Bostrom 2003).

We do not know what will happen, but several subtle constraints enable us to narrow down the range of tenable views about humanity's future and our place in the universe. These constraints derive from a variety of sources, including analysis of the capacities of possible technologies based on physical or chemical simulations; economic analysis; evolution theory; probability theory; game theory and strategic analysis; and cosmology. Partly because of the interdisciplinary and sometimes technical nature of these considerations, they are not widely understood. Yet any serious attempt to grapple with the long-term implications of technological development should take them into account.

6. 21st century biopolitics: the transhumanist-bioconservative dimension

James Hughes has argued that biopolitics is emerging as a fundamental new dimension of political opinion. In Hughes' model, biopolitics joins with the more familiar dimensions of cultural and economic politics, to form a three-dimensional opinion-space. We have already seen that in the early 90s, the extropians combined liberal cultural politics and laissez-faire economic politics with transhumanist biopolitics. In *Citizen Cyborg* (2004), Hughes sets forward what he terms "democratic transhumanism," which mates transhumanist biopolitics with social democratic economic politics and liberal cultural politics.[69] He argues that we will achieve the best posthuman future when we ensure that technologies are safe, make them available to everyone, and respect the right of individuals to control their own bodies. The key difference between extropian transhumanism and democratic transhumanism is that the latter accords a much bigger role for government in regulating new technologies for safety and ensuring that the benefits will be available to all, not just a wealthy or tech-savvy elite.

In principle, transhumanism can be combined with a wide range of political and cultural views, and many such combinations are indeed represented, e.g. within the membership of the World Transhumanist Association. One combination that is not often found is the coupling of transhumanism to a culture-conservative outlook. Whether this is because of an irresolvable tension between the transformative agenda of transhumanism and the cultural conservative's preference for traditional arrangements is not clear. It could instead be because nobody has yet seriously attempted to develop such a position. It is possible to imagine how new technologies could be used to reinforce some culture-conservative values. For instance, a pharmaceutical that facilitated long-term pair bonding could help protect the traditional family. Developing ways of using our growing technological powers to help

[69] (Hughes 2004).

people realize widely held cultural or spiritual values in their lives would seem a worthwhile undertaking.

This is not, however, the route for which cultural conservatives have so far opted. Instead, they have gravitated towards transhumanism's opposite, bioconservatism, which opposes the use of technology to expand human capacities or to modify aspects of our biological nature. People drawn to bioconservatism come from groups that traditionally have had little in common. Right-wing religious conservatives and left-wing environmentalists and anti-globalists have found common causes, for example in their opposition to the genetic modification of humans.

The different strands of contemporary bioconservatism can be traced to a multifarious set of origins: ancient notions of taboo; the Greek concept of hubris; the Romanticist view of nature; certain religious (anti-humanistic) interpretations of the concept of human dignity and of a God-given natural order; the Luddite workers' revolt against industrialization; Karl Marx's analysis of technology under capitalism; various Continental philosopher's critiques of technology, technocracy, and the rationalistic mindset that accompanies modern technoscience; foes of the military-industrial complex and multinational corporations; and objectors to the consumerist rat-race. The proposed remedies have ranged from machine-smashing (the original Luddites), to communist revolution (Marx), to buying "organic", to yoga (José Ortega y Gasset), – but nowadays it commonly emanates in calls for national or international bans on various human enhancement technologies (Fukuyama, Annas, etc.).

Feminist writers have come down on both sides of the debate. Ecofeminists have suspected biotechnology, especially its use to reshape bodies or control reproduction, of being an extension of traditional patriarchal exploitation of women, or, alternatively, have seen it as a symptom of a control-obsessed, unemphatic, gadget-fixated, body-loathing mindset. Some have offered a kind of psychoanalysis of transhumanism, concluding that it represents an embarrassing rationalization of self-centered immaturity and social failure. But others have welcomed the libratory potential of biotechnology. Shulamith Firestone argued in the feminist classic *The Dialectic of Sex* (1971) that women will be fully liberated only when technology has freed them from having to incubate children.[70] Cyberfeminist Donna Haraway proclaims that she would "rather be a cyborg than a goddess" and argues against the dualistic view that associates men with culture and technology and women with nature.[71]

[70] (Firestone 1970).
[71] (Haraway 1991).

Perhaps the most prominent bioconservative voice today is that of Leon Kass, chairman of President Bush's Council on Bioethics. Kass acknowledges an intellectual debt to three other distinguished bioconservatives: Protestant theologian Paul Ramsey, Christian apologist C. S. Lewis, and German-born philosopher-theologian Hans Jonas (who studied under Martin Heidegger).[72] Kass's concerns center on human dignity and the subtle ways in which our attempts to assert technological mastery over human nature could end up dehumanizing us by undermining various traditional "meanings" such as the meaning of the life cycle, the meaning of sex, the meaning of eating, and the meaning of work. Kass is well known for his advocacy of "the wisdom of repugnance" (which echoes Hans Jonas's "heuristics of fear"). While Kass stresses that a gut feeling of revulsion is not a moral argument, he nevertheless insists that the yuck factor merits our respectful attention:

> In crucial cases … repugnance is the emotional expression of deep wisdom, beyond reason's power to fully articulate … we intuit and feel, immediately and without argument, the violation of things we rightfully hold dear … To pollution and perversion, the fitting response can only be horror and revulsion; and conversely, generalized horror and revulsion are prima facie evidence of foulness and violation.[73]

Francis Fukuyama, another prominent bioconservative and member of the President's Council, has recently identified transhumanism as "the world's most dangerous idea".[74] For Fukuyama, however, the chief concern is not about the subtle undermining of "meanings" but the prospect of violence and oppression. He argues that liberal democracy depends on the fact that all humans share an undefined "Factor X", which grounds their equal dignity and rights. The use of enhancing technologies, he fears, could destroy Factor X.[75]

Bioethicists George Annas, Lori Andrews, and Rosario Isasi have proposed legislation to make inheritable genetic modification in humans a "crime against humanity", like torture and genocide. Their rationale is similar to Fukuyama's:

> The new species, or "posthuman," will likely view the old "normal" humans as inferior, even savages, and fit for slavery or slaughter. The normals, on the other hand, may see the posthumans as a threat and if they can, may engage in a

[72] (Kass 2002).

[73] (Kass 1997).

[74] (Fukuyama 2004). For a response, see (Bostrom 2004).

[75] (Fukuyama 2002).

> preemptive strike by killing the posthumans before they themselves are killed or enslaved by them. It is ultimately this predictable potential for genocide that makes species-altering experiments potential weapons of mass destruction, and makes the unaccountable genetic engineer a potential bioterrorist.[76]

There is some common ground between Annas et al. and the transhumanists: they agree that murder and enslavement, whether of humans by posthumans or the other way around, would be a moral atrocity and a crime. Transhumanists deny, however, that this is a likely consequence of germ-line therapy to enhance health, memory, longevity, or other similar traits in humans. If and when we develop the capability to create some singular entity that could potentially destroy the human race, such as a superintelligent machine, then we could indeed regard it as a crime against humanity to proceed without a thorough risk analysis and the installation of adequate safety features. As we saw in the previous section, the effort to understand and find ways to reduce existential risks has been a central preoccupation for some transhumanists, such as Eric Drexler, Nick Bostrom, and Eliezer Yudkowsky.

There are other commonalities between bioconservatives and transhumanists. Both agree that we face a realistic prospect that technology could be used to substantially transform the human condition in this century. Both agree that this imposes an obligation on the current generation to think hard about the practical and ethical implications. Both are concerned with medical risks of side-effects, of course, although bioconservatives are more worried that the technology might succeed than that it might fail. Both camps agree that technology in general and medicine in particular have a legitimate role to play, although bioconservatives tend to oppose many uses of medicine that go beyond therapy to enhancement. Both sides condemn the racist and coercive state-sponsored eugenics programs of the 20th century. Bioconservatives draw attention to the possibility that subtle human values could be eroded by technological advances, and transhumanists should perhaps learn to be more sensitive to these concerns. On the other hand, transhumanists emphasize the enormous potential for genuine improvements in human well-being and human flourishing that are attainable only via technological transformation, and bioconservatives could try to be more appreciative of the possibility that we could realize great values by venturing beyond our current biological limitations.[77]

---

[76] (Annas, Andrews, and Isasi 2002).

[77] I'm grateful to Anders Sandberg and Sara Lippincott for comments, and to Lovro Furjanić for a correction.

Appendix

The Transhumanist Declaration
(Version of March 2009)

(1) Humanity stands to be profoundly affected by science and technology in the future. We envision the possibility of broadening human potential by overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth.

(2) We believe that humanity's potential is still mostly unrealized. There are possible scenarios that lead to wonderful and exceedingly worthwhile enhanced human conditions.

(3) We recognize that humanity faces serious risks, especially from the misuse of new technologies. There are possible realistic scenarios that lead to the loss of most, or even all, of what we hold valuable. Some of these scenarios are drastic, others are subtle. Although all progress is change, not all change is progress.

(4) Research effort needs to be invested into understanding these prospects. We need to carefully deliberate how best to reduce risks and expedite beneficial applications. We also need forums where people can constructively discuss what should be done, and a social order where responsible decisions can be implemented.

(5) Reduction of existential risks, and development of means for the preservation of life and health, the alleviation of grave suffering, and the improvement of human foresight and wisdom should be pursued as urgent priorities, and heavily funded.

(6) Policymaking ought to be guided by responsible and inclusive moral vision, taking seriously both opportunities and risks, respecting autonomy and individual rights, and showing solidarity with and concern for the interests and dignity of all people around the globe. We must also consider our moral responsibilities towards generations that will exist in the future.

(7) We advocate the well-being of all sentience, including humans, non-human animals, and any future artificial intellects, modified life forms, or other intelligences to which technological and scientific advance may give rise.

(8) We favor allowing individuals wide personal choice over how they enable their lives. This includes use of techniques that may be developed to assist memory, concentration, and mental energy; life extension therapies; reproductive choice technologies; cryonics procedures; and many other possible human modification and enhancement technologies.

References

Annas, G., L. Andrews, and R. Isasi (2002), "Protecting the Endangered Human: Toward an International Treaty Prohibiting Cloning and Inheritable Alterations", *American Journal of Law and Medicine* 28 (2&3):151-178.

Bacon, F. (1620), *Novum Organum*. Translated by R. L. Ellis and J. Spedding. Robertson, J. ed, *The Philosophical Woeks of Francis Bacon, 1905*. London: Routledge.

Bernal, J. D. (1969), *The world, the flesh & the devil; an enquiry into the future of the three enemies of the rational soul*. Bloomington: Indiana University Press.

Bostrom, N. (1998), "How Long Before Superintelligence?" *International Journal of Futures Studies* 2.

——— (2002), "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards", *Journal of Evolution and Technology* 9.

——— (2002), "When Machines Outsmart Humans", *Futures* 35 (7):759-764.

——— (2003), "Are You Living in a Computer Simulation?" *Philosophical Quarterly* 53 (211):243-255.

——— (2003), "Human Genetic Enhancements: A Transhumanist Perspective", *Journal of Value Inquiry* 37 (4):493-506.

——— *The Transhumanist FAQ: v 2.1*. World Transhumanist Association 2003. http://transhumanism.org/index.php/WTA/faq/.

——— (2004), "Transhumanism - The World's Most Dangerous Idea?" *Betterhumans* 10/19/2004.

——— (2004), "Transhumanist Values", in Fredrick Adams (ed.), *Ethical Issues for the 21st Century*: Philosophical Documentation Center Press.

——— (2005), "The Fable of the Dragon-Tyrant", *Journal of Medical Ethics* 31 (5):273-277.

——— (2005), "The Future of Human Evolution", in Charles Tandy (ed.), *Death and Anti-Death*: Ria University Press.

——— (2005), "In Defence of Posthuman Dignity", *Bioethics* forthcoming.

Bostrom, N., and T. Ord (2005), "Status Quo Bias in Bioethics: The Case for Cognitive Enhancement", in Nick Bostrom and Julian Savulescu (eds.), *Improving Humans*, Oxford: Oxford University Press.

Buchanan, A., D. W. Brock, N. Daniels, and D. Wikler (2002), *From Chance to Choice: Genetics and Justice*: Cambridge University Press.

Capek, K. (2004), *R.U.R. (Rossum's universal robots)*, *Penguin classics*. London: Penguin Books.

Condorcet, J.-A.-N. d. C. (1979), *Sketch for a historical picture of the progress of the human mind*. Westport, Conn.: Greenwood Press.

Darwin, C. (2003), *The origin of the species, Barnes & noble classics*. New York, NY: Fine Creative Media.

Drexler, E., and R. Smalley (1993), "Nanotechnology: Drexler and Smalley make the case for and against 'molecular assemblers'", *Chemical & Engineering News* 81 (48):37-42.

Drexler, K. E. (1985), *Engines of Creation: The Coming Era of Nanotechnology*. London: Forth Estate.

——— (1992), *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. New York: John Wiley & Sons, Inc.

Esfandiary, F. M. (1970), *Optimism one; the emerging radicalism*. New York: Norton.

Ettinger, R. (1964), *The prospect of immortality*. New York: Doubleday.

Ettinger, R. C. W. (1972), *Man into superman; the startling potential of human evolution--and how to be part of it*. New York: St. Martin's Press.

Feynman, R. (1960), "There is Plenty of Room at the Bottom", *Engineering and Science* Feb.

Firestone, S. (1970), *The dialectic of sex; the case for feminist revolution*. New York,: Morrow.

FM-2030 (1989), *Are you a transhuman?: monitoring and stimulating your personal rate of growth in a rapidly changing world*. New York, NY: Warner Books.

Franklin, B., et al. (1956), *Mr. Franklin: a selection from his personal letters*. New Haven: Yale University Press.

Freitas, R., and R. Merkle (2005), *Diamond Surfaces and Diamond Mechanosynthesis*. Georgetown, TX: Landes Bioscience.

Fukuyama, F. (2002), *Our Posthuman Future: Consequences of the Biotechnology Revolution*: Farrar, Straus and Giroux.

——— (2004), "Transhumanism", *Foreign Affairs* September/October.

Glover, J. (1984), *What Sort of People Should There Be?*: Pelican.

Good, I. J. (1965), "Speculations Concerning the First Ultraintelligent Machine", *Advances in Computers* 6:31-88.

Haldane, J. B. S. (1924), *Daedalus; or, Science and the future*. London,: K. Paul, Trench, Trubner & co., ltd.

Hanson, R. (1994), "What If Uploads Come First: The Crack of a Future Dawn", *Extropy* 6 (2).

——— (1995), "Could Gambling Save Science? Encouraging an Honest Consensus", *Social Epistemology* 9:1:3-33.

——— *Burning the Cosmic Commons: Evolutionary Strategies for Interstellar Colonization* 1998. http://hanson.gmu.edu/filluniv.pdf.

Haraway, D. (1991), "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century", in, *Simians, Cyborgs and Women: The Reinvention of Nature*, New York: Routledge, 149-181.

Harris, J. (1992), *Wonderwoman and Superman: the ethics of human biotechnology, Studies in bioethics*. Oxford: Oxford University Press.

Hughes, J. (2004), *Citizen Cyborg: why democratic societies must respond to the redesigned human of the future*. Cambridge, MA: Westview Press.
Huxley, A. (1932), *Brave New World*. London: Chatto & Windus.
Huxley, J. (1957), *New Bottles for New Wine*. London: Chatto & Windus.
Jonsen, A. R. (1998), *The birth of bioethics*. New York: Oxford University Press.
Joy, B. (2000), "Why the future doesn't need us", *Wired* 8.04.
Kant, I. (1986), *Philosophical writings, The German library; v. 13*. New York: Continuum.
Kass, L. (1997), "The Wisdom of Repugnance", *The New Republic* 2 June 1997:22.
——— (2002), *Life, liberty, and the defense of dignity: the challenge for bioethics*. 1st ed. San Francisco: Encounter Books.
Kurzweil, R. (1999), *The Age of Spiritual Machines: When computers exceed human intelligence*. New York: Viking.
La Mettrie, J. O. d. (1996), *Machine man and other writings, Cambridge texts in the history of philosophy*. Cambridge: Cambridge University Press.
Leslie, J. (1996), *The End of the World: The Science and Ethics of Human Extinction*. London: Routledge.
Minsky, M. (1994), "Will Robots Inherit the Earth?" *Scientific American*.
Mitchell, S. (2004), *Gilgamesh: a new English version*. New York: Free Press.
Moore, G. E. (1965), "Cramming more components onto integrated circuits", *Electronics* 38 (8).
Moravec, H. (1989), *Mind Children*. Harvard: Harvard University Press.
——— (1999), *Robot: Mere Machine to Transcendent Mind*. New York: Oxford University Press.
More, M. *Principles of Extropy, Version 3.11* 2003. http://www.extropy.org/principles.htm.
Newman, W. R. (2004), *Promethean ambitions: alchemy and the quest to perfect nature*. Chicago: University of Chicago Press.
Nietzsche, F. W. (1908), *Also sprach Zarathustra: ein Buch fèur alle und keinen*. Leipzig: Insel-Verlag.
Office, U. S. G. P. (1949), "Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10", 2:181-182.
Orwell, G. (1949), *Nineteen eighty-four, a novel*. New York: Harcourt.
Parens, E. (1998), *Enhancing human traits: ethical and social implications, Hastings Center studies in ethics*. Washington, D.C.: Georgetown University Press.
Parfit, D. (1984), *Reasons and Persons*. Oxford: Clarendon Press.
Pearce, D. *The Hedonistic Imperative* 2004. http://www.hedweb.com/hedab.htm.
Pence, G. E. (1998), *Who's afraid of human cloning?* Lanham: Rowman & Littlefield.
Pico della Mirandola, G. (1956), *Oration on the dignity of man*. Chicago: Gateway Editions.
Posner, R. (2004), *Catastrophe*. Oxford: Oxford University Press.
Rees, M. (2003), *Our Final Hour: A Scientist's Warning: How Terror, Error, and Environmental Disaster Threaten Humankind's Future in This Century - On Earth and Beyond*: Basic Books.

Regis, E. (1990), *Great mambo chicken and the transhuman condition: science slightly over the edge*. Reading, Mass.: Addison-Wesley.

——— (1994), "Meet the Extropians", *Wired* 2 (10).

Russell, B. (1924), *Icarus; or The future of science*. London: K. Paul, Trench, Trubner & Co., ltd.

Savulescu, J. (2001), "Procreative Beneficence: Why We Should Select the Best Children", *Bioethics* 15 (5-6):413-426.

Shelley, M. W. (1818), *Frankenstein; or, The modern Prometheus*. London,: Printed for Lackington, Hughes, Harding, Mavor, & Jones.

Singer, P. (2003), "Shopping at the Genetic Supermarket", in SY Song, YM Koo and DRJ. Macer (eds.), *Bioethics in Asia in the 21st Century*: Eubios Ethics Institute, 143-156.

Stock, G. (2002), *Redesigning Humans: Our Inevitable Genetic Future*: Houghton Mifflin Company.

Teilhard de Chardin, P. (1964), *The future of man*. New York: Harper & Row.

Tipler, F. (1994), *The Physics of Immortality*. New York: Doubleday.

Turing, A. (1950), "Computing machinery and intelligence", *Mind* 59:433-460.

Ulam, S. (1958), "John von Neumann 1903-1957", *Bulletin of the American Mathematical Society* (May).

Vinge, V. (1993), "The Coming Technological Singularity", *Whole Earth Review* Winter issue.

Vita-More, N. *Transhumanist Arts Statement* 2002. http://www.transhumanist.biz/transart.htm.

Walker, M. (2002), "Prolegomena to Any Future Philosophy", *Journal of Evolution and Technology* 10.

World Medical Organization (1996), "Declaration of Helsinki", *British Medical Journal* 313 (7070):1448-1449.

WTA *The Transhumanist Declaration* 2002. http://transhumanism.org/index.php/WTA/declaration/.

Yudkowsky, E. *Collective Volition* 2004. http://www.singinst.org/friendly/collective-volition.html.