The Philosophy of Transhumanism

Max More

I. The Philosophy

To write of "the" philosophy of transhumanism is a little daring. The growth of transhumanism as a movement and philosophy means that differing perspectives on it have formed. Despite all the varieties and interpretations we can still identify some central themes, values, and interests that give transhumanism its distinct identity. This coherence is reflected in the large degree of agreement between definitions of the philosophy from multiple sources.

According to my early definition (More 1990), the term refers to:

> Philosophies of life (such as extropian perspectives) that seek the continuation and acceleration of the evolution of intelligent life beyond its currently human form and human limitations by means of science and technology, guided by life-promoting principles and values.

According to the Transhumanist FAQ (Various 2003), 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.

A corollary definition (also from the FAQ) focuses on the activity rather than the content of transhumanism:

> The study of the ramifications, promises, and potential dangers of technologies that will enable us to overcome fundamental human limitations, and the related study of the ethical matters involved in developing and using such technologies.

Thus transhumanism is a life philosophy, an intellectual and cultural movement, and an area of study. In referring to it as a life philosophy, the 1990 definition places transhumanism in the company of complex worldviews such as secular humanism and Confucianism that have practical implications for our lives without basing themselves on any supernatural or physically transcendent belief. Transhumanism could be described by the term "eupraxsophy," coined by secular humanist Paul Kurtz, as a type of nonreligious philosophy of life that rejects faith, worship, and the supernatural, instead emphasizing a meaningful and ethical approach to living informed by reason, science, progress, and the value of existence in our current life.

What is the core content of this philosophy? A simple yet helpful way to grasp its nature is to think of transhumanism as "trans-humanism" plus "transhuman-ism." "Trans-humanism" emphasizes the philosophy's roots in Enlightenment humanism. From here comes the emphasis on progress (its possibility and desirability, not its inevitability), on taking personal charge of creating better futures rather than hoping or praying for them to be brought about by supernatural forces, on reason, technology, scientific method, and human creativity rather than faith.

While firmly committed to improving the human condition and generally optimistic about our prospects for doing so, transhumanism does not entail any belief in the inevitability of progress nor in a future free of dangers and downsides. The same powerful technologies that can transform human nature for the better could also be used in ways that, intentionally or unintentionally, cause direct damage or more subtly undermine our lives. The transhumanist concern with rationality and its concomitant acknowledgment of uncertainty implies recognizing and proactively warding off risks and minimizing costs.

"Trans-human" emphasizes the way transhumanism goes well beyond humanism in both means and ends. Humanism tends to rely exclusively on educational and cultural refinement to improve human nature whereas transhumanists want to apply technology to overcome limits imposed by our biological and genetic heritage. Transhumanists regard human nature not as an end in itself, not as perfect, and not as having any claim on our allegiance. Rather, it is just one point along an evolutionary pathway and we can learn to reshape our own nature in ways we deem desirable and valuable. By thoughtfully, carefully, and yet boldly applying technology to ourselves, we can become something no longer accurately described as human – we can become posthuman.

Becoming posthuman means exceeding the limitations that define the less desirable aspects of the "human condition." Posthuman beings would no longer suffer from disease, aging, and inevitable death (but they are likely to face other challenges). They would have vastly greater physical capability and freedom of form – often referred to as "morphological freedom" (More 1993; Sandberg 2001). Posthumans would also have much greater cognitive capabilities, and more refined emotions (more joy, less anger, or whatever changes each individual prefers). Transhumanists typically look to expand the range of possible future environments for posthuman life, including space colonization and the creation of rich virtual worlds. When transhumanists refer to "technology" as the primary means of effecting changes to the human condition, this should be understood broadly to include the design of organizations, economies, polities, and the use of psychological methods and tools.

As a philosophy, transhumanism does not intrinsically commend specific technologies. Even so, certain technologies and areas of current and projected future technological development clearly are especially relevant to transhumanist goals. These include information technology, computer science and engineering, cognitive science and the neurosciences, neural-computer interface research, materials science, artificial intelligence, the array of sciences and technologies

involved in regenerative medicine and life extension, genetic engineering, and nanotechnology. A genuine understanding of the goals and potentials of transhumanism requires taking an interdisciplinary view, integrating the physical and social sciences.

The first fully developed transhumanist philosophy was defined by the Principles of Extropy, the first version of which was published in 1990. The concept of "extropy" was used to encapsulate the core values and goals of transhumanism. Intended not as a technical term opposed to entropy but instead as a metaphor, extropy was defined as "the extent of a living or organizational system's intelligence, functional order, vitality, and capacity and drive for improvement."

The Principles were formulated to "use current scientific understanding along with critical and creative thinking to define a small set of principles or values that could help make sense of the confusing but potentially liberating and existentially enriching capabilities opening up to humanity" (More 2003). The goal was not to specify particular beliefs, technologies, or policies. The Principles of Extropy consist of a handful of principles (or values or perspectives) that codify proactive, life-affirming, and life-promoting ideals supportive of transhumanism.

Although the Principles of Extropy define a specific form of transhumanism, that document is both the first comprehensive and explicit statement of transhumanism and embodies several crucial elements shared by all extant forms of transhumanism. The 2003 (and still current) version included the principles of perpetual progress, self-transformation, practical optimism, intelligent technology, open society, self-direction, and rational thinking.

Perpetual progress is a strong statement of the transhumanist commitment to seek "more intelligence, wisdom, and effectiveness, an open-ended lifespan, and the removal of political, cultural, biological, and psychological limits to continuing development. Perpetually overcoming constraints on our progress and possibilities as individuals, as organizations, and as a species. Growing in healthy directions without bound." The individual element of this is expressed in the principle of self-transformation, which means "affirming continual ethical, intellectual, and physical self-improvement, through critical and creative thinking, perpetual learning, personal responsibility, proactivity, and experimentation. Using technology – in the widest sense to seek physiological and neurological augmentation along with emotional and psychological refinement." Both of these principles clearly express the implementation of transhumanism as being a continual process and *not* about seeking a state of perfection.

All transhumanists to date would likely also have no disagreement with the principles of intelligent technology, self-direction, or rational thinking. Intelligent Technology "means designing and managing technologies not as ends in themselves but as effective means for improving life. Applying science and technology creatively and courageously to transcend 'natural' but harmful, confining qualities derived from our biological heritage, culture, and environment." Self-direction means "valuing independent thinking, individual freedom, personal responsibility, self-direction, self-respect, and a parallel respect for others." And rational thinking means "favoring reason over blind faith and questioning over dogma. It means understanding, experimenting, learning, challenging, and innovating rather than clinging to beliefs."

An emphasis on "transhuman-ism" at the expense of "trans-humanism" might lead some transhumanists to reject the principle of open society, although so far it remains highly compatible with the vast majority of transhumanist views. That principle recommends "supporting social orders that foster freedom of communication, freedom of action, experimentation, innovation, questioning, and learning. Opposing authoritarian social control and unnecessary hierarchy and favoring the rule of law and decentralization of power and responsibility. Preferring bargaining over battling, exchange over extortion, and communication over

compulsion. Openness to improvement rather than a static utopia. Extropia ("ever-receding stretch goals for society") over utopia ("no place")."

All the slightly varied statements of transhumanism gravitate around some core values, goals, and commitments. Even so, as with any complex philosophy that is no longer completely new, it is always possible to find one or two representatives who will disagree with some aspects of the view as originally stated. One or two transhumanist writers have, for instance, questioned or challenged at least some interpretations of the perpetual progress and practical optimism principles (Verdoux 2009). A distinct yet highly concordant statement of transhumanism can be found in the Transhumanist Declaration (Various 2002).

In terms of the traditional areas of philosophy, we can inquire about the main transhumanist views of ethics, metaphysics, and epistemology. I will address ethics in the third section of this essay, since transhumanists disagree over the meta-ethical basis of transhumanist values far more than they differ over basic matters of metaphysics and epistemology.

It would not be accurate to speak of a universally accepted "transhumanist epistemology," if that is taken to mean a detailed theory of the acquisition and validation of knowledge. However, transhumanists over the last almost quarter-century have practically always identified themselves as strong rationalists. A healthy legacy of the humanist roots of transhumanism is its commitment to scientific method, critical thinking, and openness to revision of beliefs. A remarkably large number of transhumanists are preoccupied with understanding and attempting to avoid the cognitive biases and deficient cognitive shortcuts to which the human brain is inherently vulnerable. That preoccupation is especially evident on the Overcoming Bias and Less Wrong community blogs.

Except for a widespread commitment to rationalism and a self-critical drive to overcome endemic human cognitive biases, transhumanists' epistemological views range widely (where these are explicitly stated – not all transhumanists delve into such recondite areas of philosophy). Some form of Piercean pragmatism seems to be quite popular while others may accept some form of externalism, explicitly or implicitly.

Another approach with substantial support (which I presented at the first conference organized by Extropy Institute) is pancritical rationalism (PCR), also known as comprehensively critical rationalism (Bartley 1962; More 1994). This epistemology, based on the work of philosopher of science Karl Popper, differs radically from much Enlightenment epistemology. Many Enlightenment thinkers defended some form of foundationalism, starting with Descartes' quest for utterly certain foundations for knowledge which he located in his supposedly clear and distinct idea of God. Empiricists and idealists (other than Bishop Berkeley) usually left God out of the picture yet still sought certain foundations in the form of unquestionable sense impressions or self-evident concepts or intellectual intuition.

Critical rationalism, by contrast, rejects this "justificationism" – the view that beliefs must be justified by appeal to an authority of some kind – in favor of the view that nothing is justified or beyond question. There are no foundations to knowledge. Acquiring and improving knowledge is based essentially on conjecture and criticism. According to critical rationalism, we can give up justification while retaining a respect for objectivity, argumentation, and the systematic use of reason. Despite the popularity of critical rationalism and its obviously close fit with the transhumanist drive toward continual improvement and challenging of limits, it must be acknowledged that there exists a small contingent of Ayn Rand-inspired transhumanists who remain committed to a foundationalist epistemology. Rand's foundationalist view explicitly claims that knowledge is hierarchical in nature, being based on undeniable axioms (Rand 1979).

Any discussion of transhumanist concerns quickly raises multiple issues in the area of metaphysics. Several of these revolve around the nature and identity of the self. With few exceptions, transhumanists describe themselves as materialists, physicalists, or functionalists. As such, they believe that our thinking, feeling selves are essentially physical processes. While a few transhumanists believe that the self is tied to the current, human physical form, most accept some form of functionalism, meaning that the self has to be instantiated in *some* physical medium but not necessarily one that is biologically human – or biological at all. If one's biological neurons were gradually replaced, for example, with synthetic parts that supported the same level of cognitive function, the same mind and personality might persist despite being "in" a non-biological substrate (Koene 2012; Merkle 2012).

Some critics who read discussions of "uploading" minds to non-biological substrates claim that transhumanists are dualists. Those critics are confusing dualism with functionalism. A functionalist holds that a particular mental state or cognitive system is independent of any specific physical instantiation, but must always be physically instantiated at any time in *some* physical form. Functionalism is a form of physicalism that differs from both identity theory (a mental state is identical to a specific brain state) and behaviorism (mental terms can be reduced to behavioral descriptions). According to functionalism, mental states such as beliefs and desires consist of their causal role. That is, mental states are causal relations to other mental states, sensory inputs, and behavioral outputs. Because mental states are constituted by their functional role, they can be realized on multiple levels and manifested in many systems, including non-biological systems, so long as that system performs the appropriate functions.

Functionalism comes in several versions, and transhumanists vary in their stands. One extension of functionalism is known as revisionary materialism or, at its extreme, eliminative materialism. Eliminativism (Churchland 1992) holds that the common-sense view of the mind ("folk psychology") is false, and that some kinds of mental states do not exist. The idea is that some of our common psychological concepts such as belief, desire, or intention are so poorly defined that they will be found to lack any coherent neurological basis – just as the concept of "caloric" was thrown out completely in favor of a thermodynamic conception of heat. Revisionary materialism (or what could equally well be called revisionary functionalism) takes the intermediate position that mental states may be reducible to physical phenomena, but only after some significant changes and refinements are made to the folk psychological concept. Given transhumanists' interest in using scientific knowledge to reconceptualize and revise human cognitive architecture, transhumanists may be uniquely open to this position in the philosophy of mind.

Transhumanists' commitment to technologically mediated transformation naturally generates great interest in the nature and limits of the self. The high level of interest in philosophy and neuroscience among transhumanists has led to a wide acknowledgment that the simple Cartesian view of the mind or self as a unitary, indivisible, and transparent entity is unsupportable. As we store more of our memories externally and create avatars, it is also becoming increasingly apparent that the boundaries of the self are unclear and may not be limited to the location of a single body. Complementing these questions about the nature and identity of the self at any one time are questions about the identity of the self over time, especially for a self that undergoes major cognitive and somatic changes over an extended lifespan. Discussions of theories of personal identity have long been a mainstay in transhumanist forums and publications (Parfit 1984; More 1995; Hughes 2012).

Another particularly relevant area of metaphysics concerns the idea of the world as simulation. As computers have become ever more powerful, simulations for both scientific and ludic purposes have proliferated and rapidly grown in sophistication. Although humans have always lived their

lives entirely in the physical world as revealed by the unmediated senses, we may come to spend much of our time in simulated environments, or in "real" environments with virtual overlays. Simulated worlds raise questions about what we value. For instance, we do value the experience of achieving something or *actually* achieving it, and how clear is the distinction (Nozick 1974)? Taking this line of thinking further, transhumanists from Hans Moravec to Nick Bostrom have asked how likely it is that we are *already* living in a simulation (Moravec 1989; Bostrom 2003).

An obvious metaphysical question to raise here is the compatibility or otherwise of religion and transhumanism. In my 1990 essay that first set forth modern transhumanism as a distinct philosophy under that name, I explained how transhumanism (like humanism) can act as a philosophy of life that fulfills some of the same functions as a religion without any appeal to a higher power, a supernatural entity, to faith, and without the other core features of religions (More 1990). The central place accorded to rationalism suggests a tension between transhumanism and religion. But are they actually incompatible?

Since rationalism is an approach to acquiring knowledge and says nothing about the content of knowledge, it is possible in principle for a transhumanist to hold some religious beliefs. And some do. The content of some religious beliefs is easier to reconcile with transhumanism than the content of others. Christian transhumanists, while not completely unknown, are very rare (and I know of none who are fundamentalists, and such a combination would surely indicate deep confusion). There are more Mormon transhumanists (although some of these are cultural rather than religious Mormons), perhaps because that religion allows for humans to ascend to a higher, more godlike level, rather than sharply dividing God from man. Several transhumanists describe themselves as Buddhists (presumably of the secular, philosophical type), and there seem to be few obstacles to combining transhumanism with liberal Judaism. However, the vast majority of transhumanists do not identify with any religion. A pilot study published in 2005 found that religious attitudes were negatively correlated with acceptance of transhumanist ideas. Those with strong religious views tended to regard transhumanism as competing with their beliefs (Bainbridge 2005).

II.   History

Before outlining the precursors, roots, and formation of transhumanism, a brief note is in order on the origin of the term itself. Many terms have been independently coined multiple times – although not necessarily with precisely the same meaning, and this is true of "transhumanism." In 1312, in his *Divine Comedy*, Dante Alighieri uses the term *transumanare*, meaning to pass beyond the human, but his usage was religious or spiritual in nature. T.S. Eliot's used of "transhumanized" in his 1935 *The Cocktail Party* is about "illumination" rather than technologically mediated transformation.

A closer fit is Julian Huxley's brief chapter "Transhumanism" in his 1957 book, *New Bottles for New Wine*. He used it to mean "man remaining man, but transcending himself, by realizing new possibilities of and for his human nature." He did not, however, develop this evolutionary view into a philosophical position, and his usage came to light years after the term was independently coined as part of the contemporary transhumanist movement.

F.M. Esfandiary had written a chapter using the term "transhuman" in a 1972 book, and went on to develop a set of transhumanist ideas in which transhuman was a transition from human to posthuman, yet he never referred to them as "transhumanism." The term was introduced

explicitly to label a deliberately transhumanist philosophy in the 1990 essay, "Transhumanism: Toward a Futurist Philosophy" (More 1990).

The current section is able to provide only a brief outline of the roots and history of transhumanism. Different scholars will emphasize differing aspects of the history, so it's important to consider multiple sources (Jones 1995; Bostrom 2005). We should be careful to distinguish precursors and proto-transhumanists from early transhumanists proper. For instance, we can easily regard the European alchemists of the thirteenth to eighteenth centuries as proto-transhumanists. Their search for the Philosopher's Stone or the Elixir of Life looks like the search for a magical form of technology capable of transmuting elements, curing all disease, and granting immortality.

In 1995, at Extro-1 – perhaps the first explicitly and exclusively transhumanist conference – as part of a survey of the roots of the transhumanist concept of extropy, Reilly Jones brought our attention to the renowned Renaissance philosopher Pico della Mirandola. The Judeo-Christian tradition has standardly pictured a gulf between human and God that is absolute and unbreachable. That led Ludwig Feuerbach to accuse that tradition of debasing humanity by transferring all the creative potential of our species into an external, perfect being who we can only worship. But Pico della Mirandola saw a far more mutable distinction between the human and the divine. In his 1486 piece, *Oration on the Dignity of Man*, he portrays God as the Craftsman explaining to humanity its nature in a way that sounds much closer to transhumanism than to the religious worldview it emerged from:

> Neither a fixed abode nor a form that is thine alone nor any function peculiar to thyself have we given thee, Adam, to the end that according to thy longing and according to thy judgment thou mayest have and possess what abode, what form, and what functions thou thyself shalt desire. The nature of all other beings is limited and constrained within the bounds of laws prescribed by Us. Thou, constrained by no limits, in accordance with thine own free will, in whose hand We have placed thee, shalt ordain for thyself the limits of thy nature. We have set thee at the world's center that thou mayest from thence more easily observe whatever is in the world. We have made thee neither of heaven nor of earth, neither mortal nor immortal, so that with freedom of choice and with honor, as though the maker and molder of thyself, thou mayest fashion thyself in whatever shape thou shalt prefer. Thou shalt have the power to degenerate into the lower forms of life, which are brutish. Thou shalt have the power, out of thy soul's judgment, to be reborn into the higher forms, which are divine.

The realization of transhumanist goals – or perhaps even the full articulation of the philosophy – would not be possible before the development and use of scientific method. In that light, we can see Francis Bacon as a precursor. In *The Advancement of Learning* (1605) and *Novum Organum* (1620), Bacon advocated inductive reasoning and helped Western thought turn away from Scholastic and Platonic approaches and toward empirical methods. As science flowered, some Enlightenment thinkers began to think along proto-transhumanist lines, as in this passage by the Marquis de Condorcet (1743–94) from *Sketch for a Historical Picture of the Progress of the Human Mind* (1795):

> In fine, may it not be expected that the human race will be meliorated by new discoveries in the sciences and the arts, and, as an unavoidable consequence, in the means of individual and general prosperity; by farther progress in the principles of conduct, and in moral practice; and lastly, by the real improvement of our faculties, moral, intellectual and physical, which may be the result either of the improvement of the instruments which increase the power and direct the exercise of those faculties, or of the improvement of our natural organization itself? (1795: 319)

Condorcet's immensely influential formulation of the idea of progress placed it at the center of Enlightenment though. With no reference to supernatural forces, he contended that growing knowledge in the natural and social sciences would enable us to create a world of growing material abundance, individual freedom, and moral compassion. He went so far as to make the following transhumanistic statement:

> Would it even be absurd to suppose this quality of melioration in the human species as susceptible of an indefinite advancement; to suppose that a period must one day arrive when death will be nothing more than the effect either of extraordinary accidents, or of the flow and gradual decay of the vital powers; and that the duration of the middle space, of the interval between the birth of man and this decay, will itself have no assignable limit? (Condorcet 1795: 368)

Enlightenment thought contained a range of views about the nature of progress, ranging from a sense of its inevitability to the view that humanity had to work hard and persistently to maintain it. To this day, some transhumanists seem attracted to the iron logic of progress, often expressed in graphs showing accelerating technological progress. However, no one goes so far as to believe in genuine inevitability in the sense often attributed to Hegel and Marx. Some formulations of transhumanism explicitly address the issue. The Principles of Extropy include the concept of "practical optimism" or "dynamic optimism" which tempers an optimistic sense of radical possibility with an insistence that we actively create the future we desire.

For several decades, it has been fashionable in some circles (especially the postmodernists and poststructuralists) to sneer at Enlightenment ideas, to declare that they are outdated, human-centric, or naive. Transhumanism continues to champion the core of the Enlightenment ideas and ideals – rationality and scientific method, individual rights, the possibility and desirability of progress, the overcoming of superstition and authoritarianism, and the search for new forms of governance – while revising and refining them in the light of new knowledge. The search for absolute foundations for reason, for instance, has given way to a more sophisticated, uncertain, and self-critical form of critical rationalism. The simple, unified self has been replaced by the far more complex and puzzling self revealed by the neurosciences. The utterly unique status of human beings has been superseded by an understanding that we are part of a spectrum of biological organisms and possible non-biological species of the future.

Before Enlightenment ideas fed into transhumanism, they were filtered through an evolutionary perspective. With the 1859 publication of Darwin's *Origin of Species*, the traditional view of humans as unique and fixed in nature gave way to the idea that humanity as it currently exists is one step along an evolutionary path of development. Combined with the realization that humans are physical beings whose nature can be progressively better understood through science, the evolutionary perspective made it easy to see that human nature itself might be deliberately changed. In a philosophical rather than scientific form, Friedrich Nietzsche picked up this idea and declared that humans are something to be overcome, and asked "What have you done to overcome him?" (Nietzsche 1896). Although Nietzsche seemed not to see a role for technology in this transformation, his bold language inspired some modern transhumanists (More 2010).

One of the more interesting precursors to transhumanism was Nikolai Fedorovich Fedorov (1829–1903), a Russian Orthodox Christian philosopher and participant in the Russian cosmism movement, who advocated using scientific methods to achieve radical life extension, physical immortality, resurrection of the dead, and space and ocean colonization. According to Fedorov, the evolutionary process led to increased intelligence culminating, so far, in human

beings. Humans must use reason and morality to shape further evolution. Especially crucial was to overcome mortality and even to restore everyone who had ever died to life. They would be restored not in their former physical forms, but in a self-creating, immortal form. Shortly after Fedorov another, less known, thinker took up the immortalist, transhumanizing cause. Jean Finot (1856–1922) advocated using science to engineer life and fabricate living matter in his book, *The Philosophy of Long Life*.

The first part of the twentieth century saw other proto-transhumanists who may have had an indirect influence on later transhumanism. These include British geneticist and evolutionary biologist J.B.S. Haldane who, in his 1924 book, *Daedalus; or, Science and the Future*, envisioned scientific advances including *in vitro* fertilization ("ectogenesis") and foresaw a world where humans direct their own evolution to their benefit. J.D. Bernal's 1929 book, *The World, the Flesh and the Devil*, developed visionary ideas about space colonization (including the Bernal sphere), and the enhancement of human intelligence and lifespan. The turn of the century and early twentieth century saw other proto-transhumanists, such as Charles Stephens and Alexander Bogdanov (Stambler 2010) as well as the beginnings of modern science fiction, which has helped expand our sense of the possible.

Transhumanism as we know it today finally began to take form in the latter part of the twentieth century. Champions of life extension played a central and persistent part in this development. Not all advocates of extending the maximum human lifespan had well-developed ideas beyond that single goal, but many had at least some sense that the same technological advances that could deliver longer, healthier lives could also enable us to change ourselves in other ways. The "father of cryonics," Robert Ettinger, was one of the latter. After explaining in his first book, *The Prospect of Immortality* (1964), that we could have another chance at life by preserving ourselves at ultra-low temperatures at the point of clinical death, his 1972 *Man into Superman* explored other transformative possibilities, and explicitly used the term "transhuman." Another enduring supporter of life extension and cryonics, Saul Kent, not only wrote practically and speculatively about extending the human lifespan, but also about other possibilities in his 1974 book, *Future Sex*.

One of the most comprehensively (if sometimes idiosyncratic) transhumanist thinkers of this period was F.M. Esfandiary (later known as FM-2030). Esfandiary's approach was more literary than academic, even though he taught at the New School for Social Research in New York in the 1960s. Starting in 1966, while teaching classes in "New Concepts of the Human," he outlined a vision of an evolutionary transhuman future. He also brought together optimistic futurists in a loosely organized group known as UpWingers. In his 1989 book, *Are You a Transhuman?*, he defined a transhuman as a "transitional human," whose use of technology, way of living, and values marked them as a step toward posthumanity. FM-2030's writing and social activity importantly underscored the practical elements of the philosophy. The idiosyncratic and personal nature of FM-2030's transhumanism was displayed in his book, which contained extensive questionnaires, then rated the reader as more or less transhuman. Some of his measures included how much someone traveled, what alterations they had made to their body (even though the existing technology remained primitive), the degree to which they rejected traditional family structures and exclusive relationships, and so on.

The dependence of transhumanist goals on major technical and scientific advances means that technically oriented visionaries have played an influential role. One central and persistent concern has been with the development of greater than human intelligence and the possibility of an "intelligence explosion." In 1970, Marvin Minsky made what turned out to be highly optimistic forecasts of the advent of super-intelligent artificial intelligence (AI), then in a 1994 *Scientific*

American article explained why vastly extended lives will require replacing our biological brains with superior computational devices.

The idea of accelerating technological progress driven by machine super-intelligence dates back several decades. This idea, now frequently referred to as "the singularity," was explicitly pondered in a 1958 conversation between Stanislaw Ulam and John von Neumann during which they discussed "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" (Ulam 1958). In 1965, I.J. Good argued that AI development would lead to an intelligence explosion. These ideas were taken up and elaborated and extended by several other influential writers (Bostrom 1998; Broderick 2001; Kurzweil 1990, 1999; Moravec 1989; Vinge 1993). Although numerous technological pathways have informed transhumanist projections, nanotechnology as envisioned by Eric Drexler has been especially influential (Drexler 1986).

Philosophy, science, and technology are not the only influences on the development of transhumanist thinking. The arts have also enriched our appreciation of what is possible and what is important. The role of science fiction in elaborating on possible transhumanist futures is clear, but a topic too large to even touch on here. Other art forms and art theory have played a role since around 1982, when Natasha Vita-More wrote the Transhuman Manifesto (Vita-More 1983), followed by the Transhuman Arts Statement (Vita-More 1992, revised 2002). In 1997, a later version of the manifesto was released first onto the Internet and signed by hundreds of creative thinkers and then placed aboard the Cassini Huygens spacecraft.

The activities of the first fully, explicitly, and exclusively transhumanist organization, Extropy Institute (ExI), shaped the intellectual and cultural movement of transhumanism starting in the late 1980s. *Extropy* magazine (subtitled "Vaccine for Future Shock" and then "The Journal of Transhumanist Thought") was first published in 1988 by Max More and Tom W. Bell (the latter coined the term "extropy"). That publication, in its paper and online versions, presented ideas from numerous leading transhumanists, and included the 1990 essay "Transhumanism: Toward a Futurist Philosophy." A vast amount of discussion focused on transhumanism started taking place when ExI created an early (and still existing) email list in 1991. This was followed by a series of influential conferences starting in 1994, a website in 1996, and an online Vital Progress Summit in 2004, which led to the development of the Proactionary Principle (More 2004). The history of ExI has been detailed further elsewhere (Extropy Institute 2005).

In 1998, the Transhumanist Declaration was crafted by an international group of authors. Other groups sprang up in the 1990s, including Aleph in Sweden, De:Trans in Germany, and Transcedo in the Netherlands. In the same year, another general-purpose transhumanist organization – the World Transhumanist Association – was founded. After Extropy Institute was closed, the WTA – renamed Humanity+ – became and remains the central organization of the movement in general, although organizations such as the Institute for Ethics and Emerging Technologies (IEET) and the Future of Humanity Institute play a strong role in the academic arena.

III.   Currents

At this point in its development, it is probably impossible to list necessary and sufficient conditions for transhumanism that would be accepted by all self-described transhumanists. At the start of this essay, some of the core elements were identified, including the view that it is both possible

and desirable to overcome biological limitations on human cognition, emotion, and physical and sensory capabilities, and that we should use science, technology, and experimentation guided by critical and creative thinking to do so. Beyond these shared and rather general views, transhumanists vary widely in their assumptions, values, expectations, strategies, and attitudes.

In addressing moral and ethical concerns, transhumanists typically adopt a universal standard based not on membership in the human species but on the qualities of each being. Creatures with similar levels of sapience, sentience, and personhood are accorded similar status no matter whether they are humans, animals, cyborgs, machine intelligences, or aliens. Yet the meta-ethical basis for making moral decisions and according rights can be consequentialist, deontological, or virtue-based. A genuinely pure deontological ethics appears to be uncommon. At least since the advent of the extropian transhumanism, many transhumanists have established their morality on a virtue foundation. In recent years, some prominent transhumanists have assumed a consequentialist foundation, in the form of various kinds of utilitarian – most radically in David Pearce's "hedonistic imperative."

Transhumanism supports a rich diversity of political perspectives. Given shared goals, it is unsurprising that transhumanists all support personal choice in the use of self-directed technological transformations, including anti-aging treatments and cryonics, cognition-enhancers, and mood-modifiers. From the late 1980s and through the 1990s, many and perhaps most transhumanists evinced a broadly libertarian politics. This perspective continues to receive far more support among transhumanists than in the general population, but over the previous decade or so, liberal democrats have become just as well recognized, some adopting the term "technoprogressive." Only the latter are likely to favor some form of world government which, in one scenario, might be controlled by a machine super-intelligence. Interestingly, even some libertarian transhumanists share this scenario, although typically seeing the super-intelligent mind acting as a guide or local governor rather than a planetary authority.

Transhumanists of multiple varieties share the view that we can make radical changes to the human condition. Major scientific and technological progress is both possible and desirable. Beyond that, consensus immediately dissolves, with differences over the expected rate, shape, and risks of progress. Those who expect a technological singularity anticipate a drastic acceleration in the rate of change, either as a one-time jump caused by the advent of super-intelligence, or as a continuous acceleration driven by exponential trends in computing power. Others expect technology to advance at different rates in various sectors and to go through faster and slower periods, depending not only on technologies themselves but on economic and social conditions.

An optimistic flavor necessarily permeates transhumanism. Someone cannot believe that radical transformations of the human condition are both possible *and* desirable while also believing that we are doomed to failure or disaster. In the early days as a self-aware movement, transhumanist discussions tended to emphasize the positive – arguing, against the cultural consensus, that technology-enabled transformations were plausible and that we should pursue them vigorously and look forward to the resulting future. (Although consideration of risk has never been entirely absent. An early book of note here is Jonathan Glover's 1984 *What Sort of People Should There Be?*) As the movement has grown, much more of the population has come to accept the plausibility of many of the technological pathways described (AI, advanced biotechnology, nanotechnology, and so on). That facilitated a shift within transhumanism to more fully consider potential risks and downsides.

While transhumanists all continue to resist "bioconservatives" and other opponents of the transhumanist vision, some emphasize the enormous uncertainties of the future. At the extreme,

this has become a focus on the risks that AI or runaway self-replication or other technologies might lead to the extinction of the human race. Bill Joy brought these kinds of dangers to the attention of a wide audience with his 2000 *Wired* article, "Why the Future Doesn't Need Us." The philosophy is enriched by a clear-eyed view of potential risks. At the same time, a heavy and persistent focus on catastrophes can make them appear inevitable and even, in some cases, desirable (the "sweet lemons" phenomenon). So far, unlike Joy, the more risk-focused transhumanists remain supportive of strong yet cautious technological advancement, believing that attempts to block advance will only increase risks.

IV.   Misconceptions

Philosophies can be clarifying not only by explaining what they *are*, but also by clarifying what they are *not*. As with any controversial view, critics often take the easy path by constructing straw men then proceed gleefully to burn them down and dance around the ashes. I have addressed several misconceptions elsewhere (More 2010b) but here will very briefly comment on some of the more common, involving notions of perfection, prediction, the body, and the attitudes toward what is "natural" but unchosen.

The frequency with which critics talk of transhumanists as wanting to "perfect" human beings or to achieve a state of perfection or to bring about a utopian society suggests that they haven't actually read much transhumanist literature. More likely, they read it with cognitive blinders on, distorting what they read to fit their preconceptions. For instance, Don Ihde (in Hansell and Grassie 2011) characterizes transhumanists as looking forward to a future posthuman world that would be a utopia. (He labels this purported goal "The Idol of Paradise.") This criticism, and the others like it, confuse the goal of continual improvement or enhancement with the longing for a state of final perfection. These are actually radically different. The former is essentially a process of perpetual change whereas the latter is a state of stasis.

Transhumanism reflects the Enlightenment commitment to meliorism and rejects all forms of apologism – the view that it is wrong for humans to attempt to alter the conditions of life for the better. Nothing about this implies that the goal is to reach a final, perfect state. The contrary view is made explicit in the transhumanist concept of extropy – a process of perpetual progress, not a static state. Further, one of the Principles of Extropy is Perpetual Progress. This states that transhumanists "seek continual improvement in ourselves, our cultures, and our environments. We seek to improve ourselves physically, intellectually, and psychologically. We value the perpetual pursuit of knowledge and understanding."

In my own formulations of transhumanism, I found the Idol of Paradise and the idea of a Platonically perfect, static utopia, is so antithetical to true transhumanism that I coined the term "extropia" to label a conceptual alternative. Transhumanists seek not utopia, but perpetual progress – a never-ending movement toward the ever-distant goal of extropia. If the transhumanist project is successful, we may no longer suffer some of the miseries that have always plagued human existence. But that is not reason to expect life to be free of risks, dangers, conflicts, and struggle. Outside, perhaps, of David Pearce's goal of eliminating all suffering, you will have to search far and wide to find any suggestion of utopia or perfection in transhumanist writing. This is why such a mischaracterization typically is not supported by quotations from relevant sources.

Another misconception – one which is somewhat more justifiable than the last – is that transhumanism essentially makes predictions about the future. Certainly some transhumanists

(whether or not they apply that term to themselves) do make predictions. In recent years the most well-known example is Ray Kurzweil, who has provided numerous detailed predictions based on his analysis of exponential trends in computer-based technologies. Beyond the forecasts made by particular individuals, it's reasonable to say that transhumanism depends on very general expectations about continued technological advance.

No specific predictions, however, are essential to transhumanism. Transhumanism is defined by its commitment to shaping fundamentally better futures as defined by values, goals, and general direction, not specific goals. Even to the extent that a goal is somewhat specific – say, abolishing aging, becoming post-biological, or enhancing cognitive abilities to some arbitrary degree – the means and time frame in which these might be achieved are open to differing views. Transhumanism per se says much about goals but nothing about specific means or schedules.

A third common misconception is that transhumanists loathe their biological bodies. The origin of this mistaken view is hard to fathom. Transhumanists do seek to improve the human body, by making it resistant to aging, damage, and disease, and by enhancing its senses and sharpening the cognition of our biological brains. Perhaps critics have made a flying leap from the idea of being dissatisfied with the body to hating it, despising it, or loathing it. In reality, transhumanism doesn't find the biological human body disgusting or frightening. It does find it to be a marvelous yet flawed piece of engineering. It could hardly be otherwise, given that it was designed by a blind watchmaker, as Richard Dawkins put it. True transhumanism *does* seek to enable each of us to alter and improve (by our own standards) the human body and champions morphological freedom. Rather than denying the body, transhumanists typically want to choose its form and be able to inhabit different bodies, including virtual bodies.

A related misconception is the reflexive assumption that, because we seek to overcome biological aging and the inevitability of death that we are terrified of death. While some transhumanists – like anyone else – may fear a painful, prolonged death, we understand that death is not something to be feared. It is nothing. It is simply the end of experience. What makes death extremely undesirable is not that it is a bad condition to be in, it is that it means the end of our ability to experience, to create, to explore, to improve, to live.

I have attempted to provide as accurate a view of the philosophy of transhumanism as is possible at present. The philosophy will continue to grow and develop, and various individuals will emphasize certain aspects of it. But these elaborations and emphases should not be confused with the core elements.

References

Bacon, Francis (1620) *Novum Organum*, trans. R.L. Ellis and J. Spedding, in J. Robertson, ed., *The Philosophical Works of Francis Bacon*. London: Routledge (1905).

Bainbridge, William S. (2005) "The Transhuman Heresy." *Journal of Evolution and Technology* 14/2 (August), pp. 1–10.

Bartley, William Warren (1962) *The Retreat to Commitment*. New York: Knopf.

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

Bostrom, Nick (2003) "Are You Living in a Computer Simulation?" *Philosophical Quarterly* 53/211, pp. 243–255.

Bostrom, Nick (2005) "A History of Transhumanist Thought." *Journal of Evolution and Technology* 14 (April).

Churchland, Paul M. (1992) *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press.

Condorcet, Jean Antoine Nicolas de Caritat (1795) *Sketch for a Historical Picture of the Progress of the Human Mind*. Google digitized edition, pp. 319, 368.

Drexler, K. Eric (1986) *Engines of Creation: The Coming Era of Nanotechnology*. London: Fourth Estate.

Ettinger, Robert (1964) *The Prospect of Immortality*. New York: Doubleday.

Ettinger, Robert (1972) *Man into Superman*. New York: St. Martin's Press.

Extropy Institute (2005) "History and Achievements of ExI from 1988 – 2005." http://www.extropy.org/history.htm (accessed January 15, 2012).

FM-2030 (1989) *Are You a Transhuman?* New York: Warner Books.

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

Hansell, Gregory R. and Grassie, William, eds. (2011) *H+/–: Transhumanism and Its Critics*. Bloomington, IN: Xlibris Corporation.

Hughes, James (2012) "Transhumanism and Personal Identity." In M. More and N. Vita-More, eds., *The Transhumanist Reader*. Oxford: Wiley-Blackwell.

Huxley, Julian (1957) *New Bottles for New Wine*. London: Chatto & Windus.

Jones, Reilly (1995) "A History of Extropic Thought: Parallel Conceptual Development of Technicism and Humanism," *EXTRO-2 Proceedings*, Extropy Institute.

Koene, Randal A. (2012) "Uploading to Substrate-Independent Minds." in M. More and N. Vita-More, eds., *The Transhumanist Reader*. Oxford: Wiley-Blackwell.

Kurzweil, Ray (1990) *The Age of Intelligent Machines*. Cambridge, MA: MIT Press.

Kurzweil, Ray (1999) *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Viking.

Merkle, Ralph (2012) "Uploading." in M. More and N. Vita-More, eds., *The Transhumanist Reader*. Oxford: Wiley-Blackwell.

Minsky, Marvin (1994) "Will Robots Inherit the Earth?" *Scientific American* (October).

Moravec, Hans (1989) *Mind Children*. Cambridge, MA: Harvard University Press.

More, Max (1990) "Transhumanism: Toward a Futurist Philosophy." *Extropy* 6 (Summer), pp. 6–12. Revised June 1994 and 1996.

More, Max (1993) "Technological Self-Transformation: Expanding Personal Extropy." *Extropy* 10, 4/2 (Winter/Spring), pp. 15–24.

More, Max (1994) "Pancritical Rationalism: An Extropic Metacontext for Memetic Progress." *EXTRO-1 Proceedings*, Extropy Institute.

More, Max (1995) *The Diachronic Self: Identity, Continuity, Transformation*. Ann Arbor, MI: A. Bell & Howell. http://www.maxmore.com/disscont.htm.

More, Max (1997) "Mind Morph: Technologically Refined Emotion and Personality." *EXTRO-3 Proceedings*, Extropy Institute.

More, Max (2003) "Principles of Extropy, Version 3.11.2003." http://www.extropy.org/principles.htm. Original version: "The Extropian Principles." *Extropy* 5/5 (May 1990).

More, Max (2004) "The Proactionary Principle" (May). http://www.maxmore.com/proactionary.htm, http://www.extropy.org/proactionaryprinciple.htm.

More, Max (2009) "Hyperagency and Hope: Critiquing Human Limitationism." Metanexus Institute, July 19. vimeo.com/9860229.

More, Max (2010a) "The Overhuman in the Transhuman." *Journal of Evolution and Technology* 21/1 (January), pp. 1–4.

More, Max (2010b) "True Transhumanism." In Gregory R. Hansell and William Grassie, eds., *H+/–: Transhumanism and Its Critics*. Bloomington, IN: XLibris Corporation.

Nietzsche, Friedrich W. (1896) *Thus Spake Zarathustra*, trans. Alexander Tille. New York: Macmillan.

Nozick, Robert (1974) *Anarchy, State, and Utopia*. New York: Basic Books.

Parfit, Derek (1984) *Reasons and Persons*. Oxford: Clarendon Press.

Pico della Mirandola, Giovanni (1486) *Oration on the Dignity of Man*.

Rand, Ayn (1979) *Introduction to Objectivist Epistemology*. Denver: Mentor.

Sandberg, Anders (2001) "Morphological Freedom – Why We Not Just Want It, but Need It." http://www.nada.kth. se/~asa/Texts/MorphologicalFreedom.htm. Retrieved November 21, 2011.

Stambler, Ilia (2010) "Life Extension: A Conservative Enterprise? Some Fin-de-Siècle and Early Twentieth-Century Precursors of Transhumanism." *Journal of Evolution and Technology* 21/1 (March), pp. 13–26.

Ulam, Stanislaw (1958) "John von Neumann 1903–1957." *Bulletin of the American Mathematical Society* (May), part 2, pp. 1–49.

Various (2002) "The Transhumanist Declaration." http://humanityplus.org/philosophy/transhumanist-declaration/.

Various (2003) "The Transhumanist FAQ: v 2.1." World Transhumanist Association. http://humanityplus. org/philosophy/transhumanist-faq/.

Verdoux, Philippe (2009) "Transhumanism, Progress and the Future." *Journal of Evolution and Technology* 20/2 (December), pp. 49–69.

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

Vita-More, Natasha (1983) "Transhuman Manifesto." http://www.transhumanist.biz/transhumanmanifesto.htm.

Vita-More, Natasha (1992) "Transhumanist Arts Statement." Revised 2002. http://www.transhumanist.biz/transhumanistartsmanifesto.htm.

Further Reading

Bell, T. W. and Murashige, K.H. (1999) "Who Owns Your Genes?" EXTRO-4: Biotech Futures Conference, Berkeley, CA.

Moravec, Hans (1999) *Robot: Mere Machine to Transcendent Mind*. New York: Oxford University Press.

Sandberg, Anders (1997) "Amplifying Cognition: Extending Human Memory and Intelligence." *EXTRO-3 Proceedings*, Extropy Institute.

Sandel, Michael (2004) "The Case Against Perfection." *The Atlantic* (April).

Stock, Gregory (2002) *Redesigning Humans: Our Inevitable Genetic Future*. New York: Houghton Mifflin.