

Contradictions from the Enlightenment Roots of Transhumanism

JAMES HUGHES*

Trinity College, Hartford, Connecticut, USA

*Address correspondence to: James Hughes, PhD, Institute for Ethics and Emerging Technologies and Public Policy Studies, Trinity College, Williams 229B, 300 Summit Street, Hartford, CT 06106, USA. E-mail: director@ieet.org

Transhumanism, the belief that technology can transcend the limitations of the human body and brain, is part of the family of Enlightenment philosophies. As such, transhumanism has also inherited the internal tensions and contradictions of the broad Enlightenment tradition. First, the project of Reason is self-erosive and requires irrational validation. Second, although most transhumanists are atheist, their belief in the transcendent power of intelligence generates new theologies. Third, although most transhumanists are liberal democrats, their belief in human perfectibility and governance by reason can validate technocratic authoritarianism. Fourth, transhumanists are divided on the balance between democracy and the market. Fifth, teleological expectations of unstoppable progress are in tension with awareness of the indeterminacy of the future. Sixth, transhumanists are divided between advocates of ethical universalism and ethical relativism. Seventh, the rational materialist denial of discrete persistent selves calls into question the transhumanist project of individual longevity and enhancement.

Keywords: *democracy, Enlightenment, ethics, reason, transhumanism*

I. INTRODUCTION

Transhumanism, the belief that science can be used to transcend the limitations of the human body and brain, is an ideological descendent of the Enlightenment, a part of the family of Enlightenment philosophies (Burch, 2001; Bostrom, 2005). As such, transhumanism has also inherited the internal contradictions and tensions of the Enlightenment tradition.

By Enlightenment I refer to a wide variety of thinkers and movements beginning in the seventeenth century and continuing through the early nineteenth century. The Enlightenment was centered in Britain, France, and Germany, but as recent scholarship has increasingly documented (Outram, 2005), it had a global dimension with significant contributions from thinkers and movements across Europe, North America, and the Caribbean. These thinkers and movements broadly emphasized the capacity of individuals to achieve social and technological progress through the use of critical reason to investigate nature, establish new forms of governance, and transcend superstition and authoritarianism.

However, this framework of ideas was only understood as the core of the Enlightenment in hindsight. Specific thinkers and movements shared only part what is now thought of as Enlightenment values and clashed over radically different interpretations of these core ideas on questions of faith, the state, epistemology, and ethics.

Today the transhumanist movement has inherited and embodies all these tensions, divisions, and contradictions. Some divisions are due to transhumanists exploring different interpretations of core Enlightenment ideas. Others are due to following Enlightenment ideas to anti-Enlightenment conclusions.

II. THE UNSUSTAINABLE AUTONOMY OF REASON

“Reason is not self-legitimizing. Like all Enlightenment advocates for reason, transhumanists find that the project of Reason erodes all premises, including the superiority of reason over unreason. Consequently, transhumanists, like the Enlightenment as a whole, needs to defend its values with irrational *a priori*s.”

Reason was the central value of the Enlightenment. Some historians see the beginning of the Enlightenment in the early seventeenth century “Age of Reason,” associated with the Descartes, de Spinoza, Leibniz, Hobbes, Locke, and Berkeley. Historian Dorinda Outram defined the central claims of the Enlightenment around its appeal to reason:

Enlightenment was a desire for human affairs to be guided by rationality rather than by faith, superstition, or revelation; a belief in the power of human reason to change society and liberate the individual from the restraints of custom or arbitrary authority; all backed up by a world view increasingly validated by science rather than by religion or tradition. (Outram, 1995, 3)

When Kant (1784a) wrote his essay “Was ist Aufklärung” or “What is Enlightenment?” for the *Berlinische Monatschrift*, he summed up the slogan of the Enlightenment as “*sapere aude*” or “dare to know.”

Though divided by epistemology and theology, these thinkers attempted to ground philosophy on uncontested propositions such as “*cogito ergo sum*.” This thoroughgoing undermining of all irrational *a priori*s led to a number of philosophical dead ends, however, immediately generating a

score of post-rationalist movements. In the midst of the Enlightenment, Jean-Jacques Rousseau valorizes the primitive and decries the harmful effects of hyperrationalism on morality (Glendon, 1999). Eighteenth century Romanticism asserts the value of esthetic and emotional experience. Across Europe, post-rationalists of both Left and Right join the “Counter-Enlightenment” (Berlin, 1998). Nietzsche, for instance, was a thoroughgoing critic of the epistemological and political claims for reason made by the Enlightenment, from Kant, Descartes, and John Stuart Mill to Darwinism and contemporary politics.

After World War II, movements on both the right and left turned against Enlightenment rationalism. On the Left, the Frankfurt School writers criticized the Enlightenment’s instrumental rationality for its complicity in authoritarianism (Marcuse, 1964; Saul, 1992; Gray, 1995; Adorno and Horkheimer, 2002). Various strains of feminism and anti-imperialism attacked the patriarchal and Eurocentric construction of Enlightenment reason (Harding, 1982). Philosophers of the Right blamed communism on the totalizing logic of the Enlightenment’s assertion of utopian reason.

These post-rationalist movements rejected the autonomy and universality of reason because it came into conflict with other values of the Enlightenment such as respect for the rights of persons and for cultural diversity. But in the twentieth century Enlightenment, rationalism began to question its own first principles. One example is found in Wittgenstein’s turn from logical positivism. The logical positivists attempted to ban from philosophical discourse all terms and concepts without empirical referents. Ludwig Wittgenstein, although an early and influential advocate of this position, eventually changed his mind as he further investigated how language actually worked. Having turned empirical investigation on the process of reasoning itself, and attempting to purify language of all irrationality, Wittgenstein (1953/2001) concluded that the goal was chimerical. Language is a series of word games in which meanings are created only in reference to other words and not to empirical facts. The positivist project of building a rational philosophy from uncontested empirical observations is impossible.

Foucault, Derrida, and the postmodernists also represent the implosion of Enlightenment reason. All claims to reason are historically situated and biased by power and position. The Enlightenment is just one historical narrative among many, and there is no rational reason to choose the Enlightenment narrative over any other. Reason can only be argued for from metaphysical and ethical *a priori*s, even if those are only such basic assumptions as “it is good to be able to accomplish one’s intended goals.”

III. TRANSHUMANISTS AND REASON

Most transhumanists argue the Enlightenment case for Reason without awareness of its self-undermining nature. For instance, Max More’s Extropian Principles codified “rational thinking” as one of its seven precepts (More, 1998):

Like humanists, transhumanists favor reason, progress, and values centered on our well-being rather than on an external religious authority. (More, 1998)

The Transhumanist FAQ defines transhumanism as the consistent application of reason:

The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason . . . We might not be perfect, but we can make things better by promoting rational thinking, freedom, tolerance, democracy, and concern for our fellow human beings . . . Just as we use rational means to improve the human condition and the external world, we can also use such means to improve ourselves, the human organism. (Humanity+, 2003)

One of the central transhumanist blogs is Less Wrong, started at Oxford University under the aegis of transhumanist philosopher Nick Bostrom, dedicated to “the art of refining human rationality.” One of the frequent contributors there is Eliezer Yudkowsky, an autodidact writer on artificial intelligence and human cognitive biases who also is the founder of the Singularity Institute for Artificial Intelligence. Yudkowsky has said that one of his goals is to lead a “mass movement to train people to be black-belt rationalists.”

One of the central philosophical debates between bioconservatives and transhumanists, and “bioliberals” more generally, over the last two decades has been over the legitimacy of emotivist arguments such as Leon Kass’ (1997) “wisdom of repugnance” (Roache and Clarke, 2009). In 2003, the bioconservative Yuval Levin wrote in “The paradox of conservative bioethics” of the tragic dilemma faced by conservatives trying to devise rational arguments in defense of irrational taboos. Once liberal democracy forces the conservative to abandon appeals to tradition or intuition, democratic debate naturalizes the new.

The very fact that everything must be laid out in the open in the democratic age is destructive of the reverence that gives moral intuition its authority. A deep moral taboo cannot simply become another option among others, which argues its case in the market place. Entering the market and laying out its wares take away from its venerated stature, and its stature is the key to its authority. By the very fact that it becomes open to dispute—its pros and cons tallied up and counted—the taboo slowly ceases to exist . . . A conservative bioethics . . . is forced to proceed by pulling up its own roots, and to begin by violating some of the very principles it seeks to defend. (Levin, 2003)

Transhumanists and the Enlightenment face the opposite dilemma: how to advocate for *rationality* in a way that avoids its potential for self-erosion. Just as the bioconservatives cannot validate their taboos and ethical *a priori*s in the public square, there is likewise no rational reason why society should reject taboos and superstition in favor of a transhuman future; both value

judgments in favor of tradition, faith, and taboo or in favor of progress, reason, and liberty stem from pre-rational premises.

Furthermore, contemporary neuroscience, also a product of reason, now recognizes that reason severed from emotion is impotent (Damasio, 1994). Transhumanists need to acknowledge their own historical situatedness and defend their normative and epistemological first principles as existential choices instead of empirical absolutes.

One example of a transhumanist acknowledging the pre-rational roots of transhumanist values is antiaging activist Aubrey de Grey's (2008) essay "Reasons and methods for promoting our duty to extend healthy life indefinitely." de Grey directly addresses Leon Kass' emotivist argument and turns it on its head. What, de Grey asks, is more repugnant than sickness, aging, and death? Those arguing the antiaging cause, de Grey concludes, should start from these shared intuitions and prejudices instead of starting from reasoned arguments that presume the "objectivity of morality" and the "unreliability of gut feelings."

IV. ATHEISM VERSUS THE TRANSCENDENT POWERS OF INTELLIGENCE

"Most transhumanists are atheists. But many transhumanists, like many of the advocates for the Enlightenment over the last 300 years, are also religious believers. Some echo the accommodations of faith and reason of the Deists. Others believe that the transcendent potential of intelligence requires a new form of scientific theology."

Even though most of its advocates believed in some kind of divine being the Enlightenment is today often seen as a secularizing or even atheist movement. Certainly, the thrust of the Enlightenment thinkers was to encourage religious tolerance and skepticism about superstition and biblical literalism. But most also believed that the rational, scientific investigation of nature would affirm some form of deity.

One common theological stance of the Enlightenment thinkers was Deism. Deists rejected blind faith and organized religion and advocated the discovery of religious truth through reason and direct empirical observation. Deists believed that divine intervention in human affairs stopped with the creation of the world. They rejected miracles, the inerrancy of scriptures, and doctrines such as the triune nature of the Christian God. Deists like Thomas Jefferson, Thomas Paine (1794/1807), and Benjamin Franklin helped establish the separation of church and state in the new United States, arguing that doctrinal differences were irrelevant to good citizenship. Deism declined in the nineteenth century to be replaced by atheist materialism on the one hand and liberal strains of Christianity such as Unitarianism on the other.

Today self-identified transhumanists are mostly secular and atheist. In a survey conducted in 2007 of members of the World Transhumanist Association

(Humanity+, 2008), 93% answered yes to the statement “Do you expect human progress to result from human accomplishment rather than divine intervention, grace, or redemption?” Ninety percent denied “clear divinely-set limits on what humans should do,” and 90% affirmed that their “concept of ‘the meaning of life’ derived from human responsibility and opportunity, not than from divine revelation.”

When the transhumanists were asked for religious affiliations, two-thirds identified as atheist, agnostic, secular humanist, or nontheist. On the other hand, a third self-identified as religious of some sort, including Christian (8%), spiritual (5%), Buddhist (4%), religious humanist (2%), as well as pagans, Hindus, Jews, Muslims, and other faiths. Echoing [Goldberg’s \(2009\)](#) thesis that transhumanism is itself a religious point of view, about 1% of transhumanists listed transhumanism as their religion.

Intriguingly 1% of respondents offered “pantheist” or “scientific pantheist” as either a religious or a secular philosophy. Pantheism appears to have become popular because of the belief among some transhumanists in panpsychism, the idea that all matter in the universe partakes of consciousness ([Goertzel, 2004](#)). This conjecture emerges out of the hypothesis that matter is a form of computation and consciousness is an emergent property of matter.

Even if the universe is not currently suffused with consciousness, the transhumanist belief in the inevitable progress of intelligence, and the ability of science to ultimately control all matter, generates its own form of teleological theology similar to Pierre Teilhard de Chardin’s notion of humanity’s evolution into an Omega Point ([de Chardin, 1955, 1959](#); [Steinhart, 2008](#)). One example of such transhumanist theological teleology or “cosmotheism” is Frank [Tipler’s \(1995\)](#) argument for a resurrection of the dead at the universe’s end.

A more minimalist version of cosmotheology is found in Nick [Bostrom’s \(2003\)](#) “simulation hypothesis.” Bostrom proposes that if the universe generates vast superintelligences with billions of years to amuse themselves, one of their activities would be the creation of simulated civilizations. Given the vast numbers of simulations they will likely run, and the epistemological uncertainty about whether we are “real” or in a simulation, the likelihood is that we are living in a simulation.

Another version of transhumanist cosmotheism is found in the “Order of Cosmic Engineers” (OCE). The OCE describes itself as a transhumanist spiritual movement that foresees a future in which intelligence engineers the universe and becomes godlike. They distinguish between belief in a “supernatural” god and belief in inevitable natural superintelligent, superpowerful gods.

... (in the) very far future one or more natural entities—i.e. entities existing within our present universe—are highly likely to come into being—plausibly resulting from the agency of our and other species—which will to all intents and purposes be very much akin to “god” conceptions held by theist religions. We refer to conceptions of personal, omnipotent, omniscient and omnipresent super-beings, “deities” or “gods.” ([OCE, 2009](#))

These natural gods might in fact already exist, produced by prior civilizations, or might be able to reach back from our future to influence the past. Religious beliefs in gods, the OCE contends, might simply be a primitive apprehension of these superbeings.

The OCE, following Gardner (2007) and Lanza and Berman (2009), also suggests that these superbeings might have the power to shape our universe or create new universes specifically designed for life. They may then have dissolved themselves or diffused themselves into our universe at the moment of creation. The perfusing of intelligence into the universe will therefore lead to the (re-)creation of these godlike beings.

The OCE views as its ultimate, very long-term aspiration—its cosmic-scale mission if you like—the permeating of this universe—by means of cosmic engineering interventions such as so-called “computronium”—with benign intelligence. We see the perfusing of our universe with benign intelligence as a step towards the (re-)constitution or (re-)integration of (possibly hive-like) “societies of mind” or “global brains”. These in turn would ultimately evolve into—a possibly new and ever so slightly improved version of—these “original” god-like super-beings.

The Enlightenment attitudes toward religion ranged from atheism to advocates for new Enlightenment religions celebrating humanity and reason, to those who believed that science validated some minimal version of their faith in the Divine. Likewise transhumanism, although dominated by atheists, is developing its own theologies.

V. LIBERALISM VERSUS TECHNOCRATIC ELITISM

“Transhumanists, like Enlightenment partisans, believe that human nature can be improved. The liberal Enlightenment tradition has argued that individuals are best at finding their own interests and should be left to improve themselves in self-determined ways. But many people are mistaken about their own best interests. This creates a constant temptation to assert that those with a better understanding of the true interests of the masses should impose improvement.”

The Enlightenment rationale for liberalism, most powerfully articulated in Mill's (1869) *On Liberty*, was that if individuals were given complete liberty they would pursue self-perfection. Their various forms of self-improvement would be diverse, but they would attempt to maximize their talents in ways that would create a bountiful flourishing society. Since no one could truly know the interests of a citizen better than they themselves did, society should guarantee free debate and equal empowerment to ensure that policies would reflect the most general interests.

Unfortunately, free peoples often do not chose ends that advocates of Enlightenment philosophies believe natural and in their own best interests. As de Spinoza (1670/1989) said “the masses can no more be freed from their

superstition than from their fears . . . they are not guided by reason” (56). The benevolent rationale for authoritarianism has always been that rulers and their advisors understand the needs of the people better than the people do themselves. Before the Enlightenment, the source of this superior understanding was purported to be the rulers’ wisdom and spiritual guidance. After the Enlightenment, the idea that some people were more or less advanced on the path of reason and progress than others lent itself to justifications for colonialism and scientific socialist dictatorships of the advanced proletariat. More benignly, monarchs influenced by the Enlightenment, such as Frederick II of Russia, began to embark on programs of public betterment through education and social reform (Outram, 2005).

Transhumanists are overwhelmingly and staunchly civil libertarian, defenders of juridical equality and individual rights. But few believe liberal democracy to be the final and best form of government. Transhumanist Max More, for instance, looks toward a post-democratic minarchy:

Democratic arrangements have no intrinsic value; they have value only to the extent that they enable us to achieve shared goals while protecting our freedom. Surely, as we strive to transcend the biological limitations of human nature, we can also improve upon monkey politics? (More, 2004)

Like the Enlightenment tradition as a whole, transhumanists can also find rationales for technocratic authoritarianism. One such rationale was introduced by Nick Bostrom (2001) in his argument for a global “singleton” to mitigate “existential risks”. In order to avoid humanity’s evolution into something we might all individually and collectively choose, and which might be competitively more fit than the alternatives, we would need to impose a “singleton” to guide global civilization past these shoals.

A singleton does not need to be a monolith. It can contain within itself a highly diverse ecology of independent groups and individuals. A singleton could for example be a democratic world government or a friendly superintelligence. (Bostrom, 2001)

In his subsequent “What is a singleton?,” Bostrom (2006) defines the singleton as

. . . a world order in which there is a single decision-making agency at the highest level. Among its powers would be (1) the ability to prevent any threats (internal or external) to its own existence and supremacy, and (2) the ability to exert effective control over major features of its domain (including taxation and territorial allocation).

He again specifies that a singleton could be a democratic world republic, a dictatorship, or a superpowerful intelligent machine or posthuman. Bostrom proposes that the eventual formation of a global government will be necessitated by the need to solve numerous global problems. In addition to preventing the inadvertent evolution of humanity in undesirable directions,

such a global agency would be able to suppress wars and arms races, protect our common planetary and solar system resources from wasteful competition, relieve inequality, and establish a more rational economy. Technological innovations such as “improved surveillance, mind-control technologies, communication technologies, and artificial intelligence,” as well as the proliferation of apocalyptic technologies that require global invasive suppression, would all increase the likelihood of the emergence of a singleton.

Bostrom leaves open the possibility that the singleton could evolve from liberal democratic self-governance and be accountable to human beings in an equal and transparent way. But the prospect of a radical improvement in the cognitive powers and moral characters of posthumans and machine minds has led other transhumanists to advocate for humanity to abdicate self-governance to more enlightened successors. One example can be found in the work of the aforementioned artificial intelligence theorist and transhumanist Eliezer Yudkowsky.

Yudkowsky has focused a lot of his writing on the problem of human cognitive biases. Yudkowsky, like other believers in a coming artificial intelligence “Singularity,” believes that human cognitive limitations will be quickly superceded by the superrationality of a recursively self-improving artificial intelligence unconstrained by biology and evolutionary drives. Human brains will never, he argues, have the same capacity for self-improvement and perfect rationalization since machine minds will have “total read/write access to their own state,” the ability to “absorb new hardware,” “understandable code,” “modular design,” and a “clean internal environment” (Yudkowsky, 2008). In fact, argues Yudkowsky, human cognition is so irredeemably constrained by bias, and our motivations so driven by aggression and self-interest, that we should give up on the project of self-governance through rational debate and do our best to hasten the day when we can turn our affairs over to a super rational artificial intelligence programmed to act in our best interests.

In his 2004 essay *Coherent Extrapolated Volition* (CEV), Yudkowsky argues that a super-artificial intelligence (AI) would be able to intuit the desires and needs of all human beings and make the decisions necessary to satisfy them. In this, Yudkowsky (unconsciously) echoes the Marxist-Leninism. At the end of his essay, he asks “What if someone disagrees with the CEV?” to which he answers:

. . . Imagine the silliness of arguing with your own extrapolated volition. It’s not only silly, it’s dangerous and harmful; you’re setting yourself in opposition to the place you would have otherwise gone. . . . (Yudkowsky, 2004)

Ironically, Yudkowsky and many of his supporters are staunch libertarians, critical of all human governance. Only godlike AI programmed for human friendliness can be trusted as global totalitarian singleton.

As the transhumanist movement grows, there will undoubtedly be a growing conflict between transhumanist defenders of democratic self-governance and advocates of enlightened technocracy.

VI. MARKET VERSUS STATE

“Since the seventeenth century, the Enlightenment has included both advocates of both radical democratic and free market models of the ideal political economy. Today these are the two largest schools of political thought within transhumanism, the democratic left and the libertarian right.”

Both the radical democratic and free market wings of the Enlightenment believed in a society of equals managing their affairs through free association. For the radical democrats, free association meant political debate and collective decision making in a society in which the political equality of citizens balanced the social inequalities of wealth. For the free marketers, free association was free exchange with the state restricted to enforcing contracts and protecting the peace. The struggle between these two Enlightenment schools of thought has dominated political economy for the last century.

Likewise these are the two dominant political strains within contemporary transhumanism, frequently skirmishing but tenuously allied. In the 2007 global survey of transhumanists, almost half (47%) identified with politics of the Left, including US “liberal,” “social democrat,” “libertarian socialist,” and “technoprogressive.” The other major bloc within transhumanism, representing 20%–25% of members, was the free market-oriented “libertarians,” “Euro-Liberals,” and “anarcho-capitalists” (Hughes, 2008). The survey found very few conservatives (<2%) of any stripe.

VII. PROGRESS, CATASTROPHE, AND TELOS

“The Enlightenment transformed the Christian teleological eschatology into a conviction that humanity would be able to continually improve itself. But the scientific worldview does not support historical inevitability and suggests that there are absolute limits on the advance of progress. Today transhumanists are torn between their Enlightenment faith in inevitable progress toward Singularities and cosmological engineering and their rational awareness of the possibility of human stagnation or extinction.”

The Enlightenment secularized religious eschatology into a narrative of inevitable human social, scientific, and moral progress (Bury, 1920; Tuveson, 1949; Nisbet, 1979). Kant (1784b) argued the inevitable progress and moral perfection of man on religious grounds, and de Condorcet’s (1795) *Sketch for a Historical Picture of the Progress of the Human Mind* sketched a vision of humanity eventually conquering all oppression and ignorance and even death and the need to toil. This faith in progress continues down through Marxist historical determinism and neoconservative triumphalism about democratic capitalism to transhumanist and Singularitarian expectations about the inevitability of artificial intelligence, nanotechnology, and immortality.

But this belief in the historical inevitability of progress has always been in conflict with the rationalist scientific observation that humanity could regress or disappear altogether. In *D'Alembert's Dream*, for instance, Denis Diderot (1769) muses that humanity could regress to inertia, or into a Borg-anism, as easily as it could progress.

Who knows if everything isn't tending to reduce itself to a large, inert, and immobile sediment? Who knows how long this inertia will last? Who knows what new race could result some day from such a huge heap of sensitive and living points? Why not a single animal? . . . Be careful. Don't be in a rush to make judgments about the great work of nature . . . Draw justified conclusions from them and in an order where there is no large or small, no absolutely durable or temporary. Watch out for the logical fallacy of the ephemeral . . . when a transitory being believes in the immortality of things. (Diderot, 1769)

Later evolutionary thought would fuse with the Enlightenment faith in progress, aided in part by Darwin's own teleological interpretation:

As all the living forms of life are the lineal descendants of those which lived long before the Silurian epoch, we may feel certain that the ordinary succession by generation has never once been broken, and that no cataclysm has desolated the whole world. Hence we may look with some confidence to a secure future of equally inappreciable length. And as natural selection works solely by and for the good of each being, all corporeal and mental environments will tend to progress towards perfection. (Darwin, 1859)

But in fact the view more consistent with science and rationalism is that evolution is simply a random walk, human intelligence is an accident, and we could easily go extinct as many species have done. Today transhumanists have inherited the tension between Enlightenment optimism and realism about progress.

In the 1990s, transhumanists were characterized by exuberant Enlightenment optimism about inevitable progress. For instance, Max More's (1998) Extropian Principles defined "Perpetual Progress" as the first precept of their brand of transhumanism:

Seeking more intelligence, wisdom, and effectiveness, an indefinite lifespan, and the removal of political, cultural, biological, and psychological limits to self-actualization and self-realization. Perpetually overcoming constraints on our progress and possibilities. Expanding into the universe and advancing without end. (More, 1998)

Since the 2000 dot-com crash, however, transhumanists have tempered their expectations about progress. One influential example of this new realism is Nick Bostrom's (2001) essay "Existential Risks" that sketched out the "bangs," "crunches," "shrieks," and "whimpers" that could end human existence. Subsequently, Bostrom edited the 2008 *Global Catastrophic Risks* volume with the transhumanist astrophysicist Milan Circovic. Catastrophic risk is a programmatic focus for both the Future of Humanity Institute that Bostrom

directs and the Institute for Ethics and Emerging Technologies which Bostrom chairs, and for the transhumanist non-profit, the Lifeboat Foundation. Whereas some transhumanists still press for a full-court press for technological innovation on all fronts and oppose all regulation, others are focusing on reducing the civilization-ending potential of proliferating genetic engineering, artificial intelligence, and nanotechnology.

Another way that the tension between eschatological certainty and sober risk assessment has played out in transhumanism is in the debate over the Singularity. Ray Kurzweil (2005), for instance, defends the inevitability of his timeline to a utopian Singularity by pointing to data on the steady exponential march of technological progress through wars and depressions. He gives little weight to the dystopian and apocalyptic predictions of how humanity might fare under superintelligent machines, suggesting that we will merge with them into apotheosis. By contrast, the followers of Eliezer Yudkowsky gathered around his Singularity Institute for Artificial Intelligence believe that an apocalyptic outcome is much more likely than a utopian one and that any AI researcher not trying to ensure the “friendliness” of his creations is an agent of human extinction. These profoundly contradictory views seem certain to generate more violent conflicts.

VIII. ETHICAL UNIVERSALISM VERSUS RELATIVISM

“The Enlightenment thinkers proposed that all human beings should be accorded the Rights of Man. Transhumanists have asserted that a transhuman society could assure the legal equality of ur-human and posthuman citizens. But awareness that ethical views are historically situated, not absolute, and belief that future generations will achieve progress in ethics, convince some transhumanist thinkers that future posthumans will have different ethics and politics. This leaves many transhumanists ambivalent about any attempt to impose their values on humanity’s descendents or even about the possibility of human/posthuman coexistence.”

The Enlightenment advocated for moral universalism, that ethics and law should apply equally to all persons. The 1948 adoption of the United Nations’ Universal Declaration of Human Rights was a milestone in the institutionalization of Enlightenment values.

But the Enlightenment also generated its postmodern critique, that the rights of man are not self-evident and absolute, and that the “moral universals” are in fact deeply historically situated. Ethical relativism is therefore as much of an Enlightenment position as moral universalism.

Transhumanists are likewise caught between ethical universalism and relativism. Most transhumanists are certainly universalist in their assertion of the rights of all people to control their own bodies and brain and to take advantage of technological enablement. But they hesitate at the idea that

humanity today should attempt to constrain the moral choices of humanity's descendents. If our descendents are morally and intellectually our superiors, then our attempt to influence them would be as foolish as our Paleolithic ancestors attempting to ensure that we did not deviate from their values.

But what if posthumans decide to enslave ur-humans, treating us like we treat animals? Some transhumanists—such as myself (Hughes, 2004)—follow Enlightenment universalism in arguing for the enforcement of equal rights for both humans and posthumans. Other transhumanists accept that posthumans' vast superiority in power, cognition, and moral progress will make pet-like servitude the best of outcomes for humans. In a 2005 survey of transhumanists, a plurality (46%) agreed that “humans and posthumans will be able to coexist in one society and polity,” whereas 41% were unsure and 12% believed they could not coexist.

David Hume (1777) also came to a pessimistic conclusion about inequality in a society with vast differences in power in *An Enquiry Concerning the Principles of Morals*.

Were there a species of creatures, intermingled with men, which, though rational, were possessed of such inferior strength, both of body and mind, that they were incapable of all resistance, and could never, upon the highest provocation, make us feel the effects of their resentment; the necessary consequence, I think, is, that we should be bound, by the laws of humanity, to give gentle usage to these creatures, but should not, properly speaking, lie under any restraint of justice with regard to them, nor could they possess any right or property, exclusive of such arbitrary lords. Our intercourse with them could not be called society, which supposes a degree of equality; but absolute command on the one side, and servile obedience on the other.

Enlightenment thinkers such as Kant and Locke argued for psychological grounds for legal personhood and against mystical ideas like ensoulment. One implication of this theory today that has strongly influenced transhumanism is the idea that the moral standing of a person depends on their level of consciousness and not the biology or hardware they are instantiated on. But critics have pointed out that this system of valuation leaves open the possibility of valuing intelligent humans and posthumans more than less intelligent ones and even overturning legal equality (Fukuyama, 2003). Allen Buchanan (2009) has recently written, for instance, of the possibility of posthumans having a higher moral status and legal rights than unenhanced humans. An optimistic outcome might be that unenhanced humans would retain all extant rights, whereas posthumans might accrue new rights consistent with their enhancements. More pessimistically, powerful, superintelligent posthumans might value their own lives more highly than the lives of baseline humans.

Although, lacking posthumans to valorize, posthumanist supremacy is still very rare, it appears likely that conflicts will emerge between advocates of human/posthuman coexistence in a transhuman democracy.

IX. ILLUSIVENESS OF PERSONAL IDENTITY

“Enlightenment values are built around the presumption of an independent rational self, citizen, consumer, and pursuer of self-interest. Even the authoritarian and communitarian variants of the Enlightenment presumed the existence of autonomous individuals, simply arguing for greater weight to be given to their collective interests. Since Hume, however, Enlightenment empiricism has called into question the existence of a discrete, persistent self. Contemporary transhumanism has yet to grapple with the radical consequences of the erosion of liberal individualism on their projects of individually chosen enhancement and longevity.”

Seeking to build a society of freely associating rational citizens, the Enlightenment replaced the possession of a soul as the basis of moral standing with cognitive capacities such as the capacity to reason. Locke, for instance, proposed that memory connected one’s present self to one’s past and was therefore the basis of personal identity.

. . . to find wherein personal Identity consists, we must consider what Person stands for; which, I think, is a thinking intelligent Being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places . . . (Locke, 1689)

Continuity of memory would allow the resurrected dead at Judgment Day to be judged for their sins even if they inhabited different, discontinuous bodies; so long as they had the same memories, they would be the same people (Locke, 1689).

But the further investigation of the nature of memory and consciousness almost immediately began to erode this idea of personal identity. In Hume’s dissection, the self is merely

. . . a bundle or collection of different perceptions which succeed one another with an inconceivable rapidity and are in perpetual flux and movement . . . (Hume, 1739)

Hume’s skepticism about the self has remained a marginal point of view. But transhumanist technologies of radical personal modification have made newly relevant this unresolved contradiction between the Enlightenment’s liberal individualism and its erosion of the rational agent. Defining personal identity in the context of radical cognitive enhancement and body modification, including uploading the mind into a computer, was the focus of transhumanist philosopher Max More’s (1995) dissertation. More argued after Locke that so long as the posthuman remembered its former life, and the chain of events that led to its “vastening,” then personal identity had been maintained.

Conversely, transhumanist Susan Schneider (2009) in her essay “Future minds: Transhumanism, cognitive enhancement and the nature of persons” suggests that transhumanist theories of psychological continuity like More’s,

or Kurzweil's (2005) "patternism," were inadequate to establish the continuity of personal identity after radical cognitive enhancements or uploading.

Bostrom acknowledged the problem of personal identity for transhumanism in the 2003 Transhumanist FAQ:

Many philosophers who have studied the problem think that at least under some conditions, an upload of your brain would be you. A widely accepted position is that you survive so long as certain information patterns are conserved, such as your memories, values, attitudes, and emotional dispositions, and so long as there is causal continuity so that earlier stages of yourself help determine later stages of yourself These problems are being intensely studied by contemporary analytic philosophers, and although some progress has been made, e.g. in Derek Parfit's work on personal identity, they have still not been resolved to general satisfaction. (Humanity+, 2003)

I addressed this issue in my 2001 essay on "The Future of Death," adopting a Buddhist and Parfitian point of view, that personal identity is a convenient illusion that is now eroding under the impact of neurotechnology. In particular, that essay looked at the challenge to personal identity of the remediation of severe brain damage that has caused personality erasure. Future cognitive enhancement would, I argued, push liberal democratic society to adopt a post-personhood standard of personal identity and moral standing. It is hard to discern, however, what meaning "liberty, equality, and fraternity" would have without the convenient fiction of autonomous individuals as citizens.

X. CONCLUSIONS

Transhumanism reflects within itself the myriad tensions and contradictions of the Enlightenment tradition of which it is a product.

- Transhumanists are staunch advocates of the supremacy of reason, but like all Enlightenment partisans they need to develop nonrational grounds on which argue for reason and Enlightenment values.
- Most transhumanists are atheists, but some see their religious faiths as consistent with transhumanism, as many Enlightenment thinkers believed their faiths consistent with reason. The transhumanist belief in the eventual apotheosis of omnipotent intelligence, again inherited from the Enlightenment, leads some transhumanists toward materialist theologies.
- Most transhumanists support a liberal democratic polity. But belief in the irrationality of democratic self-governance and the possibility of benevolent technocratic rule legitimate transhumanist authoritarianism.
- Transhumanists, like Enlightenment thinkers generally, are deeply divided between advocates of radical democracy and free markets.

- Transhumanists are divided between those who believe in the Enlightenment narrative of the inevitability of moral and social progress and those who more soberly acknowledge the possibility of human stagnation or even extinction.
- Many transhumanists are advocates of ethical universalism and believe that humans and posthumans can coexist in a transhuman society. But some are ethical relativists, believing that posthumans will create their own superior moral code that humans cannot judge.
- The neurotechnologies that transhumanists believe will be available in the future will call into question the discrete, continuous self on which Enlightenment political theory is based.

Just as every Enlightenment movement has schismed over these radically different interpretations of Enlightenment values, so the young transhumanist movement appears on the cusp of developing many radically divergent ideological variants.

REFERENCES

- Adorno, T. W., and M. Horkheimer. 2002. *Dialectic of Enlightenment*. Trans. Edmund Jephcott. Stanford, CA: Stanford University Press.
- Berlin, I. 1998. *The proper study of mankind: An anthology of essays*. New York: Farrar Straus and Giroux.
- Bostrom, N. 2001. Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology* 9(1). Available: <http://www.nickbostrom.com/existential/risks.html>. (Accessed December 10, 2009).
- . 2003. Are you living in a computer simulation? *Philosophical Quarterly* 53:243–55. Available: <http://www.simulation-argument.com/simulation.html>. (Accessed December 11, 2009).
- . 2005. A history of transhumanist thought. *Journal of Evolution and Technology* 14(1):1–25. Available: <http://jetpress.org/volume14/bostrom.html>. (Accessed December 5, 2009).
- . 2006. What is a singleton? *Linguistic and Philosophical Investigations* 5(2): 48–54. Available: <http://www.nickbostrom.com/fut/singleton.html>. (Accessed December 10, 2009).
- Buchanan, A. 2009. Moral status and human enhancement. *Philosophy & Public Affairs* 37:346–81.
- Burch, G. 2001. Progress, counter-progress and counter-counter-progress [On-line]. Available: <http://www.gregburch.net/progress.html>. (Accessed December 4, 2009).
- Bury, J. B. 1920. *The idea of progress: An inquiry into its origin and growth*. London: Macmillan.
- Damasio, A. 1994. *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.
- Darwin, C. 1859. *On the origin of species* [On-line]. Available: <http://www.literature.org/authors/darwin-charles/the-origin-of-species/chapter-14.html>. (Accessed December 14, 2009).

- de Chardin, T. 1955. *Le Phénomène Humain (The human phenomenon)*. trans. B. Wall. New York: Harper Collins.
- . 1959. *The future of man*. trans. N. Denny. New York: Doubleday.
- de Condorcet, M.-J.-A.-N. C. M. 1795. *Sketch for a historical picture of the progress of the human mind* [On-line]. Available: <http://oll.libertyfund.org>. (Accessed December 13, 2009).
- de Grey, A. D. N. J. 2008. Reasons and methods for promoting our duty to extend healthy life indefinitely. *Journal of Evolution and Technology* 18(1):50–5. Available: <http://jetpress.org/v18/degrey.htm>. (Accessed December 1, 2009).
- de Spinoza, B. (1670) 1989. *Tractatus Theologico-Politicus*, translated by S. Shirley with an introduction by B S Gregory. Leiden. Available: http://www.philosophyarchive.com/index.php?title=A_Theologico-Political_Treatise_-_Spinoza. (Accessed December 1, 2009).
- Diderot, D. 1769. *D'Alembert's dream* [On-line]. Available: http://records.viu.ca/~Johnstoi/diderot/revedalembert_tofc.htm. (Accessed December 1, 2009).
- Fukuyama, F. 2003. *Our posthuman future*. New York: Picador.
- Gardner, J. 2007. *The intelligent universe: AI, ET, and the emerging mind of the cosmos*. Franklin Lakes, NJ: New Page Books.
- Glendon, M. A. 1999. Rousseau & the revolt against reason. *First Things* 96(October 1999): 42–7.
- Goertzel, B. 2004. Patterns of awareness: A pattern-theoretic, panpsychist solution to the hard problem of consciousness [On-line]. Available: <http://www.goertzel.org/dynapsyc/2004/HardProblem.htm>. (Accessed November 11, 2009).
- Goldberg, S. 2009. *Does the wall still stand? The implications of transhumanism for the separation of church and state*. Workshop on transhumanism and the future of democracy, templeton research lectures at the Arizona State University Center for the Study of Religion and Conflict, April 24, 2009 [On-line]. Available: <http://lsr.nellco.org/georgetown/fwps/papers/107/>. (Accessed November 1, 2009).
- Gray, J. 1995. *Enlightenment's wake: Politics and culture at the close of the modern age*. New York: Routledge.
- Harding, S. 1982. Is gender a variable in conceptions of rationality? A survey of issues. *Dialectica* 36:226–41.
- Hughes, J. 2001. The future of death: Cryonics and the telos of liberal individualism. *Journal of Evolution and Technology* 6(1). Available: <http://jetpress.org/volume6/death.html>. (Accessed December 9, 2009).
- Hughes, J. 2004. *Citizen cyborg: Why democratic societies must respond to the redesigned human of the future*. Cambridge, MA: Basic Books.
- . 2008. Report on the 2007 interests and beliefs survey of the members of the World Transhumanist Association [On-line]. Available: <http://www.transhumanism.org/resources/WTASurvey2007.pdf>. (Accessed December 1, 2009).
- Humanity+. 2003. Transhumanist FAQ [On-line]. Available: <http://humanityplus.org/learn/transhumanist-faq/>. (Accessed December 1, 2009).
- Hume, D. 1739. *A treatise of human nature: Book I: Of the understanding. Part IV: Of the sceptical and other systems of philosophy. Section VI: Of personal identity* [On-line]. Available: <http://www.mnstate.edu/gracyk/courses/web%20publishing/TreatiseI.iv.vi.htm>. (Accessed December 5, 2009).
- . 1777. *An enquiry concerning the principles of morals* [On-line]. Available: <http://www.gutenberg.org/dirs/etext03/nqpmr10.txt>. (Accessed December 13, 2009).

- Kant, I. 1784a. Was ist Aufklärung. *Berlinische Monatschrift* (Dezember-Heft): 481–94. Available: <http://www.uni-potsdam.de/u/philosophie/texte/kant/aufklaer.htm>. (Accessed December 10, 2009).
- . 1784b. Idea of a universal history from a cosmopolitical point of view [On-line]. Available: <http://www.marxists.org/reference/subject/ethics/kant/universal-history.htm>. (Accessed December 14, 2009).
- Kass, L. R. 1997. The wisdom of repugnance. *The New Republic* 216(22):17–26.
- Kurzweil, R. 2005. *The singularity is near*. New York: Viking.
- Lanza, R., and B. Berman. 2009. *Biocentrism: How life and consciousness are the keys to understanding the true nature of the universe*. Dallas, TX: Benbella Books.
- Levin, Y. 2003. The paradox of conservative bioethics. *New Atlantis* 1:53–65.
- Locke, J. 1689. Essay concerning human understanding [On-line]. Available: http://oregonstate.edu/instruct/phl302/texts/locke/locke1/Essay_contents.html. (Accessed December 15, 2009).
- Marcuse, H. 1964. *One-dimensional man: Studies in the ideology of advanced industrial society*. Boston, MA: Beacon Press.
- Mill, J. S. 1869. *On liberty* [On-line]. Available: <http://www.utilitarianism.com/ol/one.html>. (Accessed December 13, 2009).
- More, M. 1995. The diachronic self. Identity, continuity, transformation. PhD dissertation, University of Southern California (published under his former name, Max T. O'Connor). Available: <http://www.maxmore.com/disscont.htm>. (Accessed December 1, 2009).
- . 1998. *The extropian principles v3*. Extropy Institute. Available: <http://www.maxmore.com/extprn3.htm>. (Accessed December 1, 2009).
- . 2004. *Democracy and transhumanism*. Extropy Institute. Available: <http://www.extropy.org/politicaltheory.htm>. (Accessed December 1, 2009).
- Nisbet, R. 1979. The idea of progress: A bibliographical essay. *Literature of liberty: A review of contemporary liberal thought* 2(1). Available: <http://oll.libertyfund.org/index.php?Itemid=259&id=165>. (Accessed December 13, 2009).
- Order of Cosmic Engineers. 2009. Prospectus [On-line] Available: <http://cosmeng.org/index.php/Prospectus>. (Accessed December 3, 2009).
- Outram, D. 1995. *The Enlightenment*. Cambridge, UK: Cambridge University Press.
- . 2005. *The Enlightenment*, 2nd ed. Cambridge, UK: Cambridge University Press.
- Paine, T. 1794/1807. *The age of reason; being an investigation of true and fabulous theology* [On-line]. Available: <http://www.archive.org/details/agereasonbeinga01paingooq>. (Accessed December 2, 2009).
- Roache, R., and S. Clarke. 2009. Bioconservatism, bioliberalism and repugnance. *Monash Bioethics Review* 28(1):4.1–4.21.
- Saul, J. R. 1992. *The dictatorship of reason in the west*. New York: The Free Press.
- Schneider, S. 2009. Future minds: Transhumanism, cognitive enhancement and the nature of persons. In *The Penn Center guide to bioethics* (pp. 844–56), ed. V. Ravitsky, A. Fiester, and A. L. Caplan. New York: Springer. Available: <http://ieet.org/index.php/IEET/more/schneider20090204/>. (Accessed December 5, 2009).
- Steinhart, E. 2008. Teilhard de Chardin and transhumanism. *Journal of Evolution and Technology* 20(1):1–22.
- Tipler, F. 1995. *The physics of immortality: Modern cosmology, god and the resurrection of the dead*. New York: Anchor.
- Tuveson, E. L. 1949. *Millenium and Utopia: A study in the background of the idea of progress*. Berkeley, CA: University of California Press.

- Wittgenstein, L. 1953/2001. *Philosophical investigations*. Oxford: Blackwell Publishing.
- Yudkowsky, E. 2004. *Coherent extrapolated volition*. San Francisco, CA: Singularity Institute for Artificial Intelligence.
- . 2008. *Artificial intelligence and society*. *Transcript of a speech given at the Artificial Intelligence and Society event at the University of Santa Clara* [On-line]. Available: <http://www.acceleratingfuture.com/people-blog?p=1611>. (Accessed December 10, 2009).