What remains of man after the Singularity?

Authors:

Manu Srinath Halvagal (Microengineering), Aly Makhlouf (Bioengineering), Lucas Ramirez (Computer Science) and Hang Yuan (Neuroscience)

PROJECT SHS 1^{st} MASTER

Supervised by

Michael-Andreas Esfeld (Philosophy of Science) and Dustin Lazarovici (Philosophy of Science)

Report accepted on June 16, 2018

Lausanne, Academic Year 2017-2018

What remains of man after the Singularity?

June 16, 2018

Abstract In order to address the question, "What remains of man after the Singularity?", we must address both what is meant by the Singularity, as well as what it means to be human. In this essay, we first present an overview of the concept of the technological Singularity and outline existing ideas regarding human nature, as well as some potential pitfalls of these ideas. After demonstrating that these ideas are not satisfactory for our discussion, we introduce a novel framework in the second half of our essay. We propose this framework, inspired by both the Turing Test and contemporary views on moral philosophy, in order to address the limitations of current attempts at defining human nature. Finally, we discuss how our framework would view humans in a post-Singularity world, who may be so drastically altered, in some cases, that they no longer be considered human by our standards today.

Contents

1 Introduction

As early as the ancients Greeks, humans had started to contemplate what human nature entails. Today, some believe in biological determinism, envisioning human nature as simply an expression of our genome. Some conceptualize human nature as an aggregate of several defining traits, while others opt for more abstract notions. Although many existing ideas provide us with useful insights into human nature, so far no single idea entirely withstands scrutiny in the context of modern society. This is due to a sharp surge in computational capacity and the technological consequences that have followed, as well as the rapid development in natural sciences towards understanding the human body and the human mind.

The conceptualization of human nature is of paramount importance in modern society, as we begin to use scientific developments to radically modify ourselves, both for the purpose of fulfilling certain survival needs and for making functional enhancements. For instance, many individuals need cardiac pacemakers to regulate their heartbeat, without which survival would be difficult, while others use brain-computer interfaces to control machines for functional enhancement. There also exist numerous prosthetic devices with varying levels of functionality that help amputees regain limb function. These new developments in the sciences confront us with questions about various philosophical issues that might emerge as time goes by. In order to make our debate more comprehensive, we set our discussions in a hypothetical and currently somewhat ambiguous futuristic scenario - the technological Singularity. This will be a time of runaway technological growth, paving the way for drastic human enhancement that will deeply challenge our notions of human nature. By setting our discussion in this scenario, we can engender more fruitful results by truly stretching the philosophical scope of the concepts we discuss to the technological development that we envision.

More concretely, we will introduce the Singularity and its immediate consequences in section 2, to set up the overarching background of our central discussion. Then, we will inspect an array of different schools of thought on defining human nature, as well as their associated fallacies in section 3. After showing why current ideas are not entirely satisfactory, we introduce a more encapsulating, moral philosophy framework and its justifications in section 4, as an alternative for elucidating what it means to be human. Lastly, we will benchmark our framework against several extremely disruptive consequences brought forth by the Singularity in section 5, to understand how our view of humans may change in this context.

2 The Singularity

The concept of the Singularity has its roots in the ongoing process of technological progress. Although somewhat ambiguous, it is commonly thought of as being a major milestone in this technological progression, a point in time that clearly separates a certain pre-Singularity world from a post-Singularity one. It sets the context for unfathomable and persistent changes to human civilization which have the potential to disrupt its core features.

As we will see, the Singularity is intrinsically linked to the rise of information-systems that began during the last century. The definition we provide for this event will highlight this relation and allow us to argue the validity of our assumption concerning its occurrence. We may then open the Pandora's box of the consequences it can have on the human population.

2.1 Origins

Technological progress took a consequent turn halfway through the 20th century, as the mathematical models that stand behind today's computers were being formalized by John von Neumann and Alan Turing. In *Tribute to John Von Neumann* (Ulam, 1958), Polish-American scientist Stanislaw Ulam reported one conversation with Neumann who says "centered on the ever-accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity

in the history of the race beyond which human affairs, as we know them, could not continue". Around the same time, Alan Turing published *Computing Machinery and Intelligence* (Turing, 1950) in which he envisioned the existence of digital machines capable of fooling humans regarding their true nature: the Turing test was born as well as the idea that machines will one day be able to reach and eventually outclass human intelligence.

As technology has been having an increasing impact on individuals' daily lives, the Singularity has also been receiving more attention than ever before. We now take a closer look at the Singularity through the lens of Ray Kurzweil, one of the most prominent thinkers in this field, in the following section.

2.2 Definition

What exactly is the Singularity? For the sake of argument, we settle on one precise definition for this essay, noting that multiple definitions exist in the related literature. We define the Singularity as the creation of the first super-intelligent machine. The meaning attached to super-intelligence in the essay is the capacity to surpass any human in all spheres of intellectual activity. Formally speaking, it translates into the ability to solve any mathematical optimization problem more efficiently than any (non-augmented) human (Shanahan, 2015). Here, efficiency is measured by the accuracy of the solution, and the amount of time required. Scientific research is a good example of a maximization problem where the objective function is the amount of knowledge one can obtain about a particular field, whose axioms constitute the constraints as well as the rules that can be used to derive logical statements. A super-intelligent computer will be able to produce scientific knowledge for humanity at a rate no human being could ever achieve.

There is a specific case of optimization problem that is of particular interest to our discussion: the ability to design super-intelligent machines. Such a machine, by definition, will be able to improve on itself at super-human efficiency. Recursively, the resulting intelligent agent will itself have the ability to yield an even-better machine through self-improvement. The sole limits to this recursion are the number of materials needed to build these ever-improving computers and the physical limit of computational power in the universe.

One may wonder why such an event would be so significant that it warrant a name on its own. The reason is that this event represents the moment when technological development switches from being human-driven to machine-driven. As demonstrated, having super-human problem solvers will enable humanity to expand technologically at an unprecedented rate. This is reinforced by the observation that machines are by their very nature easily reproducible; once humanity succeeds in building the first super-intelligent computer, it is a matter of time before a group of these machines can be set-up to work in parallel on whatever problem they are given.

2.3 Computational Power & Brain Emulation

There are two essential components for the Singularity to materialize: sufficient raw computing power, and an efficient algorithm to solve all optimization problems.

For the former, it boils down to ensuring that the evolution of our computing power remains on the path it followed since the early days of computer science. This path was somewhat formalized by Gordon Moore (co-founder of Intel) in 1965 and is now known as Moore's law. It states that the size of a transistor (the most basic unit of computation in digital computers) in a dense integrated circuit is halved every two years. Accounting for this and other design improvements (e.g. faster-operating transistors, improved computer architecture), digital computers' performance-to-price ratio has effectively roughly doubled every 18 months from the 1960s until the 2000s.

One can view this evolution from a much wider scope. Ray Kurzweil linked the progression of information-technologies with the observation that the complexity of life forms on Earth has increased exponentially since the appearance of life. He uses this to argue in favor of his *Law of Accelerating Returns* (Kurzweil, 2001), which states that Moore's law is merely the continuation of this perpetual increase in complexity rather than an isolated phenomenon, and that this evolution will eventually result in a technological Singularity.

We now turn to the second problem: building an algorithm that will allow a machine with sufficient computing power to outperform humans in intellectual tasks. One of the preferred approaches in current research is to emulate the human brain, as we know it is its underlying biological mechanisms that yield the human intelligence we strive to surpass. If one could precisely define the 'brain's algorithm', then one could in theory emulate human intelligence, and overcome it simply by speeding up the machine's processor. Even though state-of-the-art artificial intelligence (AI) algorithms are still far from achieving this human-level intelligence, a few higher-level structures in artificial neural networks, similar to those of the human brain, are already being used to mimic real intelligence, notably the brain's learning mechanisms. An example of this is Google's 'Deep Mind', which is one of the most powerful AI ever designed and relies heavily on this approach to create 'intelligence'. Consequently, and perhaps surprisingly, our ability to create a super-intelligent machine in the future is likely to depend on our understanding of the human brain.

2.4 After the Singularity

In the previous section, we have clarified the path to the technological Singularity and now we are going to argue for its feasibility. The absence of a definitive argument against this event's embodiment, as well as the rate of progression of information-systems in the past decades, provides us with enough evidence to support its plausibility. Hence, we position ourselves in a post-Singularity world, and deal with the question: how does this milestone affect humanity at the individual level?

As highlighted earlier, technological progression will happen at a super-human rate, and a fast-paced flow of revolutions is to be expected. Ray Kurzweil argues that the development of three particular technologies will constitute the heart of these revolutions to come. He refers to them as the *GNR* (*Genetics*, *Nano-technologies*, *Robotics*) (Kurzweil, 2006). These are the technologies that he believes have the most disruptive potential to human nature. For example, one of the most discussed topics among futurists is transhumanism, which is the ability to introduce non-biological features into the human body. Possessing such augmentations, transhumans would have increased abilities compared to their non-augmented peers and would be able to transcend the limitations of their organic envelope. Achieving this level of engineering requires accurate models of how the various mechanisms of life interact with each other to create living beings, among which the field of genetics stands as the information process underlying all forms of life on Earth. The latter is already an incredibly active field of research whose application domain includes the possibility of eliminating diseases and of radical life extension by completely reprogramming our biology. Nano-technology, a domain in its early stages today, promises to overcome the limits of biological life by allowing the construction of structures at a molecular level. Swarms of nano-bots in our bloodstream could accomplish all sorts of tasks, from monitoring our vital signs to interfacing with our brain, the latter allowing us to live in virtual realities that we would be able to design at will. Combined with genetic engineering, nano-technologies would open the path to immortality and complete restructuring of the human body. On a different note, advancements in robotics (and computer science in general) will shorten the gap between already augmented humans and non-biological machines. Having the evolution of computers' complexity in sight, one may even envision the day where it is possible to entirely transfer a human mind into a non-biological substrate, marking the beginning of humanity's transfer to software.

The revolutions we just discussed are not exclusively within the domain of a post-Singularity world, as some of these technologies will probably see their first embodiments much before the Singularity. Nonetheless, the exacerbation of our technological capabilities, permitted by the Singularity, allows us to

consider them in their full potential, which we believe is of greater interest for the analysis we will conduct on human nature. Moreover, our discussion will focus on a period where the major revolutions described above have been completely integrated by societies to a point where machines are indistinguishable from humans. We will also discard issues related to the intermediate, 'transition' phase between now and the Singularity, focusing solely on the effects of the Singularity on human nature.

3 How do we define human nature?

Since the overarching theme of the paper is whether the Singularity will disrupt the very fabric of human nature, it is important to first understand the current state of the debate about exactly what human nature entails. Since a human is an elusive concept in this essay, we henceforth refer to it as 'human', indicating that it is a loosely-defined being. In broad terms, the definition is widely debated among those who see it entirely from the perspective of biological determinism, such as evolution and genetics, to more abstract notions of humanity (Savulescu, 2009). After studying these ideas in more depth, we can more critically address the question: Will we still be considered 'human' in the advent of the Singularity?

3.1 Existing Ideas

Over the centuries, there have been many ideas on how to define human nature. Aristotelian thought regarding the Four Causes, as laid out in *Physics* II 3 and *Metaphysics* V 2 (Falcon, 2015), attributes to each and every thing in existence a telos or a "final cause". As such, for Aristotle, there existed a final cause or higher purpose to human nature as well, and 'human' existence is intended to achieve this purpose. Apart from these metaphysical arguments, Aristotle does make a few clear statements about human nature:

- "Man is a conjugal animal.": This refers to the idea that the family unit or household is fundamental to human existence.
- "Man is a political animal.": Man has an innate tendency to form large communities and divide labor within these communities. To do so, man needs to use his unique capacity for reason, which is another defining trait of 'human' beings.
- "Man is a mimetic animal.": This represents classical Greek thought about art and representation, where the basis of humanity's penchant for creating and appreciating art is derived from an urge to imitate what already exists.

In these statements, Aristotle focuses on a few primary traits that distinguish us from other animals, identifying these traits as qualities of a higher order. He describes a 'best' 'human' as one who exhibits these finer qualities of human existence and one who behaves according to morals described in his works on virtue ethics (The *Nicomachean Ethics* II). Hence, for Aristotle, the qualities unique to 'humans' are not only what define us, but also what we are meant to achieve at our best. This idea of ascribing some higher purpose to human existence by no means belongs solely to Aristotle. Most religions, in fact, put forth this very idea in various forms that are built upon their own particular mythos.

With the discovery of the Theory of Evolution and the principles of genetic inheritance, a new dimension was added to the debate regarding human nature (Juengst and Moseley, 2016). These scientific ideas made it possible to describe, in purely objective terms, the exact differences between species, as well as what sets the 'human' species apart. This directly gave rise to a school of thought adopted by species preservationists, who emphasize the inviolability of species barriers that have naturally evolved over millenia. In his book 'Enough: Staying Human in an Engineered Age', McKibben (2004) argues that attempts to "direct evolution" through enhancements and genetic modification, which blur these barriers, are extremely irresponsible and could potentially be disastrous. Instead of trying to interfere with the natural order, he stresses the importance of preserving humanity's place within natural ecosystems.

Another prominent school of thought states that human limitations are part, and an important one at that, of our very nature. Human vulnerability and our struggle to confront our suffering are seen as a central aspect of what gives 'human' life meaning (Parens, 1995). There are sub-groups in this line of thinking that focus on different kinds of human vulnerabilities. "Life cycle traditionalists" condemn efforts to extend the human life span (Callahan, 1995). "Personalists" stress that human vulnerabilities are important in keeping 'humans' humble and modest (Fitzgerald, 2008). These views have been criticized primarily by "meliorists", who observe that 'humans' have endeavored to overcome their limitations throughout human history.

In his book 'Human Enhancement', Julian Savulescu challenges the basic notions of biological determinism and species conservationism through simple yet illustrative examples. By pure biological determinism, 'humans' have a total of 46 chromosomes and are defined, as a species, by their ability to reproduce with one another. This definition, he contends, is not sufficient, as it fails to encompass those with Down syndrome, who have 47 chromosomes, as well as those with Turner syndrome, who have 45 chromosomes and are unable to reproduce. By his reasoning, as long as we consider them to be 'human' with basic moral rights, then the defining traits of human nature cannot sufficiently be captured by this purely taxonomical classification. Motivated by this line of reasoning, Savulescu proposes a set of candidate traits, describing human nature as the "capacity to reason, act autonomously, engage in complex social relationships, display empathy and sympathy, have faith and act for moral reasons" (Savulescu, 2009, p. 243). Of the aforementioned traits, Savulescu believes that acting under a normative guide is the most important attribute to our nature, meaning that 'humans' are able to reason about and act on what they believe they should do, and not just on what they want to do. This normative behaviour is made manifest by displaying "practical rationality", which is the ability to reason about things out of self-interest and not strictly by a balance of evidence and argumentation, as well as having "judgement sensitive attitudes", which is the ability to exercise judgement in different situations and adjust behaviour accordingly Savulescu (2009).

Taking a more contrarian view, Arthur L. Caplan, a prominent bioethicist and philosopher of science, asserts that there is, in fact, no such thing as designated human nature (Caplan, 2009). Caplan establishes this stance on human nature in the context of a debate on the ethics of human enhancement. With regards to this issue, Caplan is an outspoken advocate of meliorism, a philosophical position maintaining that 'humans' can and should intervene in the natural processes of the world around them, ultimately for its betterment. This is in response to the anti-meliorist stance, which maintains that humans should not "play God" and that the technologies of the future, if left unchecked, promise to "destroy the very thing that makes us human - our nature." (Caplan, 2009, p. 201). It is here, Caplan interjects, that the anti-meliorists fall short, in their failure to precisely formulate a notion of human nature.

From Caplan's point of view, human evolution is a random walk; an aggregation of stochastic and external forces that incidentally intersected and drove our development (Caplan, 2009). Ultimately, in their basic evolutionary form, 'humans' were mere products of chance and not of design. However, 'humans' have always been dynamic beings, driving their own evolution in ways that often run counter to the forces of nature. Had it not been for these active efforts to develop technology and purposefully control the natural environment, 'humans' would not have survived until now. As such, this remnant of our evolution, our so-called human nature, is just an incidental state of being that holds no merit to how 'human' beings have come to be the way that they are. It is for this reason, he argues, that it cannot qualify as an essence that is significant to how we define ourselves, nor can it be seen as a designated, static concept, in light of the ever-evolving 'human' self. The way that 'human' beings are today is by no means a normative guide for how 'humans' used to be or should continue to be in the future, nor does it confine human nature to a strictly-defined set of qualities.

Taking this notion further, Caplan draws attention to the fact that the technologies we take for granted and use so freely today, as well as our sedentary lifestyle, would have drastically challenged our way of being as early humans; hunter-gatherers with only fire and the most basic tools at our disposal. Conse-

quently, Caplan argues that our conception of human nature entirely depends on the technology that is available to us. What it means to be 'human' today is starkly different to what it meant to be 'human' in our early evolutionary history and in the wake of drastic technological development, it is natural for there to be drastic challenges to the notions we have of being 'human'. Caplan aims to demonstrate that this intimate tie between human nature and technology has been a consistent hallmark of the human condition from time immemorial and that the technological progress of the 21st century is merely a difference of scale, not concept. If we insist on prescribing a singular trait to define human nature, it is our technological development that stands out throughout history as the focal point of being 'human', and "there is no reason to think that this creative manipulation of our environment, including our own bodies and minds, is any less worthy of inclusion as part of human 'nature'." (Caplan, 2009, p. 202).

3.2 Analysis

Without more clearly defining what we mean by being 'human', we cannot address the more pertinent question of how human nature will change in the face of extreme technological development. In the interest of moving forward with the discussion, it becomes important to narrow down our definition of 'human' to only the most seemingly salient features, as presented in the literature. This will allow a more focused and coherent framework for future analysis. Needless to say, the list of definitions presented above is by no means exhaustive. However, it serves to demonstrate the current state of the debate, as well as the breadth of opinions surrounding this controversial topic of defining human nature. By further analysis, what becomes apparent is that all of these definitions fail to capture a fully-defined notion of human nature, as can be demonstrated by simple rebuttals that will deconstruct the inherent inadequacies of each definition in isolation.

Proceeding with the Aristotelian definition, Aristotle asserts that being 'human' can be simply delineated by three broad traits, namely being "conjugal", "political" and "mimetic". Upon closer examination, this definition falls short in some aspects, namely in the case of 'humans' who do not marry, either out of choice or lack of opportunity, thereby violating the seemingly essential conjugality of being 'human'. Some 'humans' do not see conjugal relationships as being in any way essential to their existence, nor do they seek to live in groups or establish a household. Some 'humans' live in isolation from others and do not contribute to the ordained structure of labor division and large communities. Of course, one may say that Aristotle's primary thesis regarding human nature is not that the above three traits are a complete description of human nature, but rather it is the idea that humanity can be defined by a final cause ("telos") or purpose; 'humans' are defined by their higher purpose or what they seek to be, not by what they are. However, it cannot be said, with any degree of certainty, that 'humans' have reached any kind of consensus on a common 'higher purpose' as a species, nor do all 'humans' even subscribe to this train of thought. It is worth noting that the definition is not without its merits, and the set of traits proposed by Aristotle do attempt to dig deep into some of the inner workings of being 'human' that go beyond the often more obvious and categorical definitions. While it may sufficiently capture some seemingly prominent and, at times, more subtle features of human nature, it is most effective when aggregated across the entire human population and has clear shortcomings in select individual cases.

Turning to the more biologically-deterministic definition upheld by species conservationists, we refer back to Savulescus illustrative rebuttal of this definition, where he draws attention towards 'humans' with Down syndrome, who are characterized by an extra chromosome, and 'humans' with Turner syndrome, who are characterized by a missing chromosome and an inability to reproduce. In both cases, these 'humans' fail to subscribe to the strictly taxonomical definition that species conservationists propose, which calls to question the integrity of this definition altogether. Furthermore, genetic compositions vary with time and space and the very concept of an untampered genome that is a pure representation of our species does not exist in nature (Robert and Baylis, 2003). Rather than a rigid categorization based on exact features of the genome, it is perhaps a more valid argument to say that, so long as any being can still reproduce within the human species, that being can be said to human. However, even this statement is not without its own problems. There are many humans who are unable to have children, and we perceive them to

be just as human as others. In fact, the ability to have children is rarely a factor in such discussions at all.

However, the definition that Savulescu proposes in its stead is not without its own limitations. Savulescu had suggested that 'humans' are able to reason, act autonomously, engage in complex social relationships, display empathy and sympathy, have faith and act for moral reasons, as well as display "practical rationality" and "judgement sensitive attitudes". This elaborate list of traits still falls short in some regards, as we demonstrate by the following set of examples. For the first example, we examine the case of sociopaths who, with reference to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), act in accordance to their own set of norms and entirely out of self-interest. They are categorically unable to display empathy, having an impaired sensitivity to the well-being of others and lacking in any feeling of remorse. This puts sociopaths at odds with many of the traits that Savulescu deems essential to human nature, especially the ability to display empathy and sympathy. It can be argued that some of the other traits he proposes, namely the ability to exhibit practical rationality and judgement-sensitive attitudes, are in fact heightened in sociopaths, who are often fully able to manipulate others and navigate social interactions with a high level of tact and precision, usually motivated by an entirely self-interested goal. Perhaps they cannot be considered non-normative in their behaviour, although their behaviour does run counter to established norms, but rather they live by a different set of norms of their making. In this way, they are still fully able to display the sort of reasoning required to have the normative behaviour that is characteristic of 'humans'. Of course, we must recognize that any human is able to and often does act selfishly to the detriment of others. However, sociopaths are distinguishable in that this behaviour is not only made manifest in isolated events, but rather is an inherent part of their existence. The same can be said of people who display some form of autism which, although to a large extent manageable, is an ultimately incurable disorder that mainly impairs their ability to have complex social interactions and demonstrate judgement with a certain level of situational awareness. One might say that children, such as infants or toddlers, are driven almost entirely by their biophysical needs and are unable to display the normative behaviour and judgement-sensitive attitudes required for complex social interaction. However, most children have the capacity to develop these traits as they grow older, distinguishing them from sociopaths and those with autism. Furthermore, even those who do not have this capacity are in no way deemed any less 'human' for this reason.

Finally, placing Caplans definition under the same scrutiny, Caplan had said that 'humans' are characterized by the use of their intelligence to develop and use technology, specifically for the manipulation of themselves and the world around them. Here, we strictly define "technology" as any means of self-improvement that was not afforded to 'humans' naturally. This means that, in the context of our discussion, clothes and cutlery are equally valid forms of technology as computers and cars. Although not obvious, this definition appears impervious to the same criticisms of aforementioned definitions. In attempting to propose rebuttals to this definition, we quickly find that we cannot conceive of a 'human', living today, whose survival is not incumbent on the use of technology in one form or another. These technologies have a very strict particularity to 'humans' and appear to be truly essential to the way in which they exist. Even if we take the extreme case of uncontacted peoples, who live in almost total isolation from modern civilization and have had no exchanges with modern technology, we also find that they have no doubt developed their own technologies, as primitive as they may be, in a struggle to survive the environment in which they exist. Moreover, given the choice or the appropriate circumstances, all 'humans', including uncontacted peoples, have sufficient mental faculties to develop technology of the complexity that we see. Thus, the definition of human nature that Caplan proposes does not define the development and use of technology on the basis of choice or circumstance, but rather on the basis of ability.

In this way, Caplans definition quite selectively attributes this trait to 'humans', simply because he argues that 'humans' appear to be the only ones who are capable of it. Most oft-cited traits of human nature, such as intelligence and emotion, walk a very fine line between 'humans' and other animals, as we become increasingly aware of the ability of many animals to display both intelligence and emotion, similarly to 'humans'. In light of the fact that some animals have been observed to use very primitive tools, tools being a form of technology, we are cautious in acknowledging the fact that the use of technology by

'humans' may simply be attributed to a difference of scale, rather than concept. This is analogous to how we perceive human intelligence as merely a different level of intelligence to other animals, rather than denying the existence of non-human intelligence. However, the statement still holds that no other animals are able to observably leverage their intelligence to develop technology the same way that 'humans' do, nor can they actively drive their own self-improvement to become stronger, healthier or more intelligent. While comparing intelligence between 'humans' and 'non-humans' is a more abstract endeavor, technology is tangible and can be directly observed. Therefore, we believe it would not be constructive to simply attribute this technological disparity between 'humans' and other animals to a difference of scale, when the technological capabilities of even the most primitive human still far surpass and are qualitatively different to the best of any known animal. As such, the observable difference between the nature and complexity of 'human' technology compared to any other is so drastic so that we believe it to warrant its own classification. We therefore consider this trait of developing and using technology as one that is not immediately observed in all but 'humans'. Since we cannot conceive of 'non-humans' that fit Caplan's definition, nor 'humans' that do not fit the definition, it becomes very difficult to propose counter-examples that rebut this definition.

While our analysis shows that Caplan isolates, with impressive particularity, a trait that is highly specific to 'humans', this notion is called into question, upon closer examination, when we consider a more hypothetical case. We try to imagine a form of artificial super-intelligence or super-intelligent extra-terrestrial life, which has no resemblance to 'humans' today. We are suddenly faced with a new class of beings that, by virtue of their intelligence, are capable of both dramatic self-improvement and the development and use of technology for this self-improvement, in a way we thought only exclusive to 'humans' today. While, by Caplans definition, these beings would perfectly subscribe to a trait that he argues is definitively 'human' in nature, it takes no stretch of the imagination to conclude that we would most likely not consider these beings to be 'human'. This draws attention to a hidden but very much inherent limitation of Caplans definition, as it yet again appears unable to offer an entirely holistic definition of human nature.

It would appear, in fact, that none of the definitions presented above were perfectly able to capture the essence of what it means to be 'human'. While each definition has its own merits and provides insight into some of the defining features of being 'human', no single definition can be taken in isolation without insufficiency. In fact, the belief in a non-designated, non-specific human nature that is fluid and ever-changing, as originally proposed by Caplan, appears more in-line with the results of this discussion, as we consistently fail to place human nature under a static categorization. Perhaps this attempt at a bottom-up approach, whereby we lay down a set of candidate traits for human nature, build an increasingly more fine-grained definition and ultimately reach a master definition that fully encompasses everything that is quintessentially 'human', is a futile pursuit that cannot be practically realized. Ultimately, by this approach, it stands to reason that we will always be either too broad or too restrictive in our definition of human nature, which is one of the reasons why the concept remains so elusive until this day. Thus, we face a fundamental categorization problem, whereby it seems practically impossible to fully categorize human nature, either by a single definition or a combination of different definitions. For the purposes of progressing with the analysis, we believe it to be more productive to take an alternative approach, which we describe herein.

4 A Moral Philosophy Framework

Arthur Caplan's aforementioned ideas regarding the fluidity of human nature seem like a reasonable basis for developing our own thoughts, agreeing that the idea we propose must be extremely flexible. However, simply saying that there is no such thing as a designated human nature sidesteps the issue and does not provide any meaningful bounds for subsequent discussion. Drawing inspiration from one of Julian Savulescu's more prominent lines of reasoning, a moral philosophy approach provides a framework for addressing this categorization problem by placing general but meaningful constraints on what we consider

to be 'human'. By this approach, we need not dissect human nature into its constituent components; in fact, we need not define human nature at all.

We begin by imagining an individual, to whom we assign a select set of defining traits. If this imagined individual, by virtue of these assigned traits, possesses sufficient quality to make them worthy of basic moral rights, then we can conclude that this individual is 'human', without even examining the traits that underpin this individual. However, we acknowledge that assigning moral rights is a problematic approach, owing to the ambiguity of the concept and the fact that, even if it was more strictly formalized, the assignment of said moral rights is a highly controversial topic in and of itself. This may be demonstrated with a simple case, whereby many of us today consider other animals just as deserving of moral rights as 'humans', while others may not. In this way, we would not be able to use moral rights to differentiate 'humans' from other 'non-humans', if the assignment of these rights is not specific enough to 'humans'. In order to address this, we take a more narrow scope, where we speak only of the right to 'human' rights and the life enjoyed by these rights. To define a benchmark for 'human' on which our notion of 'human' life can be built, we simply look to ourselves, the conscious, intelligent beings who are reading this paper and pondering its questions. We give precedent to our current state as humans and take it as the base case for our analysis. Even though we cannot define being 'human', we know with certainty that our present selves are 'human' and we recognize other 'humans' when we see them, since we have not yet been confronted with the more ambiguous realities of the envisioned future. Accordingly, a 'human' life is simply the kind of life that we as 'humans' have a right to, decided by the only beings we know for sure to be 'human' at this moment in time, ourselves.

We precisely define what we mean by the right to 'human' life, so as not to appear as though we are stripping non-human beings of their own right to life. When we consider the case of a non-human being, such as a pesky or vile insect, we find that we often do not seem to value its life, nor do we see a need to preserve it. However, if we consider larger animals, such as cats, dogs and horses, to which we are far more sensitized, many believe in the need to protect the sanctity of their life. Indeed, many people claim to firmly believe this, but consume cows and chickens without hesitation. The seemingly contradictory notions we have about non-human life make clear that 'human' life has a particular quality to it that must be absent in other 'non-humans', as there is much stronger consensus in how we value 'human' life. We much more strongly condemn the loss of 'human' life than any other animal and by no means condone the sacrifice of other humans for nourishment. Moreover, 'human' life is protected not only by its preservation, but also by our moral pursuit of maintaining a minimum quality of 'human' life for all. This can be seen in the 'Universal Declaration of Human Rights' by its assertion of the right of all 'humans' to life, safety, justice, liberty, freedom of thought, freedom of expression, education, work, healthcare, social security, among others (UN General Assembly, 1948). In contrast, we often strongly protect 'non-human' life and speak of the need to preserve it, but do not demonstrate equal regard for the quality of its life; there is little to no discussion about protecting the liberties and freedoms of household pets as we do 'humans', nor about their right to education or social security.

Even with this clearer moral delineation between 'humans' and 'non-humans', others may argue that discrimination among humans is a timeless feature of the human condition and shows that not all humans are believed to be equal in their right to 'human life. Indeed, our moral standards two hundred years in the past deemed racial discrimination completely acceptable. Our moral standards then were starkly different than they are today and will likely be different two hundred years in the future. Since our moral framework is grounded on moral standards, it is important to note that it is an entirely pragmatic proposal that by no means claims moral absolutism in its conclusions. The metaethical question of whether absolute moral truths can be attained is of course an interesting question, but is not the focus of this paper. As such, we must point out that the conclusions drawn by using our framework would be entirely dependent on the time in which it is being applied, evolving as our notions of moral rights and the right to 'human' life evolve as well. However, the framework itself is transferable and can be applied at any point in time to reach conclusions that are particular to the time in which the framework is being applied. Most importantly, we propose it as an approach that we believe can give us useful insight into

this categorization problem as we tackle it as 'humans' today. It is also important to note that our moral framework is based on moral standards as they are in theory, not in practice. In practice, a person may be discriminated against for a job position, on the basis of race. In theory, however, no reasonable moral argument, by our current moral standards, can be made to condone this discrimination, nor assert that said person has no right to a job. Thus, by casting aside the prejudices of morals as they are exercised in practice, the set of rights that a 'human' individual has for a 'human' life, as imagined in theory, are more robust. With that, it is clear that while the notion of 'moral rights' may be ambiguous to us, the right to 'human' life is categorically different to other moral rights and is, to a large extent, more well-defined.

Now, with this framework, we are able to take a more binary approach to the categorization problem of what it means to be 'human', which is simple and intuitive. We simply consider an individual as either 'human' or 'non-human', on the sole basis of whether we believe they merit the right to a life that a 'human' should have. Simply feeling that a being is 'human', even if in ways that we have thus far failed to vocalize, is nonetheless sufficient and constructive for our discussion, if we feel the need to afford said being the right to 'human' life. This makes for a powerful analytical tool for delineating 'human' and 'non-human' beings, while avoiding the pitfall of having to formulate categorical, abstract ideas about human nature that often feel disconnected from the subject under question.

In order to move forward with this framework, we must address two questions that present themselves. The first and more immediate of these is the question of consensus. Who decides if something deserves the right to 'human' life and how much consensus is needed? It would appear, in this case, that consensus is a practical impossibility, since sub-populations of 'humans' will always differ in their opinions about any topic, let alone pressing philosophical issues. Here, it is up to each person who employs this approach to decide for themselves who deserves the right to 'human' life and who does not. There is no way of having consensus, nor is it possible to define a threshold for consensus. However, the belief that we as 'humans', whether consciously or unconsciously, are aware of the things most essential to our nature and would reject something if it felt 'non-human', means that we believe our moral framework is specific enough to garner appreciable consensus, in the face of the many hypotheticals that may be presented to it. However, we must recognize that it is by no means infallible and it does not claim to be. It does, however, seem general enough to encompass diverse human traits that were previously excluded by certain definitions, such as people with genetic abnormalities, psychological impairments or certain medical conditions. Another attractive aspect of our approach lies in the fact that our framework for defining human nature would be as general as the moral framework that we subscribe to, while still excluding evidently alien forms of life. If we encountered an extra-terrestrial population of beings that are just as sentient, intelligent and conscious as ourselves today, but who we do not qualify as worthy of the right to 'human' life, we would not consider them 'human' by our framework. They may most certainly be worthy of a right to life, a life of their own with whatever particularities that may come with it, but it would not be the right to the same life that we as humans have today. There is no suggestion that this life is better or worse than 'human' life, nor that denying the right to 'human' life is in anyway a relegation of their status. It simply illustrates their inherent difference from us, making them a being worthy of its own life but not a 'human' being. Why we do not qualify them as 'human' is inconsequential; perhaps it is because they are green and have a jelly-like form, or perhaps it is because they eat through their skin. Whatever the reason may be, if it seems sufficiently 'alien' as to be deemed 'non-human' by us with appreciable consensus, then it is not 'human'.

The second question we must address is the question of applicability. How do we apply this framework and under what circumstances is it valid? We know that we want to use this framework in the context of the Singularity, an imagined, highly technologically-advanced future. But exactly what point in time are we referring to? In order to better illustrate this problem, we may imagine a one-dimensional axis with left-most and right-most extrema. Current 'humans' occupy the left-most extreme of this axis, while purely electromechanical machines occupy the right-most extreme. At some time in the future, a so-called transitory phase from a pre- to post-Singularity world, 'humans' and machines are categorically distinct entities, but are becoming increasingly more intertwined as they converge towards a common point along

this imagined axis. When the Singularity is realized, we may then imagine that these two extrema have converged to a single, common point along this axis. During the transitory phase, we acknowledge that very large heterogeneity will likely exist among 'humans', as some 'humans' remain in their current form while others, with better resources and access to enhancement technologies, develop markedly improved physical and mental capabilities. In such a heterogeneous state of mankind, what then will become of the definition of 'human' and who then will be the one to decide? How will our framework apply in this situation? While this is an important philosophical question, it is not immediately pertinent to our discussion. We maintain that our discussion is only concerned with how 'humans' today would view 'humans' in a post-Singularity world, not on how different forms of 'human' will view one another as they transition from a pre- to post-Singularity world. The comparison we draw is between ourselves and the 'humans' that will exist when all transient dynamics have faded. The benchmark in our discussion remains grounded in the 'human' of today, while the 'human' against which this benchmark is measured is the 'human' of a post-Singularity world. As such, we explicitly make the assumption that at the time of the Singularity, man and machine will become indistinguishable entities and all 'humans' will exist in a homogeneous 'transhuman' form, of unknown but equal part 'human' and machine origin, unlike that which we know today. It is this 'transhuman' state that is the focal point of our analysis, as we analyze to what extent humans in a 'transhuman' form would still be considered 'human' by todays standards. The culmination of our entire methodology, therefore, can be summarized as follows: "If we imagine a homogeneous population of post-Singularity 'transhumans', subjected to the radical enhancement technologies we envision for the future, would we still afford these 'transhumans' the right to 'human' life as we know today? If so, they are no less 'human' than we are. If not, they are no longer 'human' by our current understanding." We then wish to examine how this framework holds up to the philosophical issues raised by the Singularity, as well as enhancement technologies in general, as we discuss in the next section.

5 How will humanity change?

5.1 Genetic Engineering

If we look to the timeless practice of selective breeding, it is evident that humans have long manipulated genomes for their benefit. Although this practice has long been embraced, the more precise, accelerated form of artificial selection offered by genetic engineering is considered highly suspect. While there are plenty of lines of reasoning opposing this practice, it is reasonable to conclude that none would argue that a genetically-modified tomato is no longer a tomato. In fact, in the absence of any indication to the contrary, many would consume genetically-modified tomatoes unknowingly.

In light of existing debates on genetic engineering, more interesting becomes the case when the organism undergoing selection is also the organism undertaking it. Does the prospect of changing the human genome at will or correcting deleterious genetic mutations disrupt our way of being? Would we deprive a being of 'human' status the moment the being decides to undertake some form of corrective gene therapy? It seems absurd to think that parents who choose to correct their child's embryonic predisposition for Huntingtons disease or give their child blue eyes, are in any way dehumanizing their child. When this child is born, how or why can it be viewed any differently, and is there any reasonable justification for viewing it as any less worthy of a right to human life than ourselves or yours. Is it the artificial nature of their insemination that would be called into question? Well, so too are children born of in-vitro fertilization today, whom we certainly do not deem unworthy of human life. Or is it the tampering of their genomes with exogenous agents? Exogenous retroviral genomes naturally exist as endogenous retroviruses in the human genome, the evolutionary artifacts of past infections. Genetic engineering today is none other than an approach drawing inspiration from these otherwise naturally-occurring processes.

More extreme is the case of genome engineering, the process of chemically assembling an entirely artificial genome, unbacked by the biological process of human fertilization. Why should this be viewed any differ-

ently, if the outcome of the process is still a human, indistinguishable from ourselves? Do we need prior knowledge of its genetic makeup or the nature of its insemination, to make the relevant judgement about conferring it 'human' status? In fact, its origin is inconsequential to us and should not be consequential in the future. While the processes by which we achieve genetic engineering are plenty, some more disruptive than others, we need only concern ourselves with the outcome. It follows, on the other hand, that if transhumans were genetically engineered to grow wings, they would less likely be considered 'human' by our current standards. No known human of the past or present has ever had wings and wings confer a drastically different functional lifestyle on the being that is given them. This illustrates that it is not the principle of enhancement, but rather the nature and magnitude of the enhancement that is important to the discussion. They may be just as intelligent or emotional as humans today, but we need only resort to our framework to elucidate the very alien nature of their being. As such, if we were to ask ourselves whether a being with wings is as worthy of the right to 'human' life as ourselves, the answer would likely be 'no' with appreciable consensus across our current population, in such an extreme case as this.

5.2 Artificial Bodies and Immortality

The influence of artificial bodies might first seem innocuous to our discussion of what it means to be 'human', since many people nowadays already employ some form of artificial body parts, and few have questioned their identity as 'humans'. Nonetheless, the issue gets more subtle when the artificial body parts no longer solely serve as medical apparatus, but as a means that can empower humans to a drastic degree, perhaps where even immortality is possible. Immortality can be achieved by replacing a failed organ with a new, synthetic one that could potentially be more functionally-reliable and energy-efficient. As long as one's brain remains intact, this person may become immune to many deadly diseases and physical destruction. Even under such extreme conditions, the majority of mankind would still probably consider those immortal beings as simply humans that can live longer. The extended lifespan will not change how those people interact with others, nor disturb their ability to exhibit 'human' behaviour and emotions. Why then would these immortals be considered not worthy of basic human rights?

While immortality does not make a 'human' non-human, based on our framework, some very puzzling issues arise if we extend our discussion to that of self-identification. In essence, when artificial bodies are possible, one person may perhaps be able to control several human-like robotic beings at the same time, using a single brain. The question then arises whether these robotic avatars are worthy of basic human rights as we have outlined. If one does not know of the true nature of these avatars, which appear to behave exactly like a human for all intents and purposes, their 'humanity' would not be called into question. However, when their origin and nature of being is understood, our perception, as we are today, may very likely change. In this case, we would likely look at the entire being (the controlling brain, along with the avatars) as a whole, and carry out our evaluation accordingly, treating each avatar more akin to a body part rather than something that merits rights in isolation. Thus, any rights we assign to the avatars would be derived from the rights we wish to assign to the whole being. A puzzling aspect here is whether it is correct for our perception to change, as it would, with our knowledge of the inner workings of any being, or whether the classification must be entirely pragmatic and functional. If the controlling brain was originally 'human', and has not been drastically altered, one could perhaps still consider such beings to be 'human'. However, the issue here is more contentious, since we are now essentially talking about a 'brain in a vat'. This leads us to the despondency regarding self-identification between the two possible answers to the question of whether the core of a man is just the mind or, as the embodiment argument goes, both the body and the mind are two inseparable components of an individual. This would then rely on the preconceptions of an observer, and answering this question of embodiment is outside the scope of our discussion. Admittedly, our moral philosophy framework breaks down when the consensus is hard to form. However, in the discussion of synthetic bodies and immortality, we would still consider post-Singularity beings as 'human', insofar as this kind of self-identification issue does not arise.

5.3 The Nature of Experience

Our senses continuously send streams of inputs by means of electrical impulses to the brain. The brain then integrates this data in real-time and creates a model of the physical world surrounding us. This is how we experience the reality that all humans are familiar with. What if we replaced these 'real' streams of inputs with artificial ones, carefully crafted as to simulate a reality that is not the one happening right now? If done correctly, this virtual (or simulated) reality would be indistinguishable to the subject from a 'true' representation of the world. This practice does not even require modifications on the subject's brain or body, one needs only to find a way to nullify the subject's senses and inject simulated ones instead (Shanahan, 2015).

This topic becomes more interesting when one considers a virtual reality shared by multiple human users. Some people might choose to spend most of their time inside a simulated world, while being surrounded with other remotely human-controlled avatars. Does this profoundly disrupt the intrinsic humanness of these individuals? Is there a threshold of real-world experience that is required for someone to be considered 'human'? There is a natural parallel with the practice of playing video games. Players often find themselves entirely immersed in universes designed by game developers. Although their sensory perceptions are not directly connected to the game, players often report or can be observed to lose track of their surroundings, including the feeling of time. The comparison with living in a virtual reality is even more valid when we touch upon multiplayer online games (so-called *Massively Multiplayer Online (MMO)* games) that let players interact and cooperate with each other to achieve common objectives.

We certainly consider players to be worthy of a right to human life, however addicted we find them to be to video games. Why would it make a difference to actually experience the game as an alternate reality then? The evolution of a person's avatar would still be guided by this person's human brain, since, as stated earlier, it is not necessary to artificially augment the brain to make someone experience a simulated reality. It would even be possible to construct real sensory-based relationships with other people through their respective avatars, something that video game players do to a lesser extent. In this regard, people choosing to spend their time in a virtual reality would perhaps display more of their human nature than online game players, and would thus still be considered as human beings according to our framework. A more interesting question is whether we consider the virtual avatars to be human. In this case, the answer is again the same as that for robotic avatars, in that any rights they have are only derived from the overall being, in this case, the human player and the avatar together. The intriguing aspect here is that even these virtual avatars will probably be afforded certain protections by merit of their being an important part of a human, whose loss could be potentially as distressful as physical injury or worse, due to a strong self-identification with the avatar.

5.4 Uploading Minds

Soon after sufficient computational power is available and we reach an understanding of the underlying processes of the brain and intelligence in general, one of the most radical technologies that can be expected to arise is the possibility to simulate a complete human brain on a computational platform. Once this technology becomes possible, a whole host of philosophical issues arise.

First of all, it becomes possible for a person to upload their entire personality and memories onto an inorganic substrate. The first issue that arises is whether the new entity can be considered 'human'. This is, in fact, simply a more extreme case of the issue we discussed in the case of artificial body parts. If we decide to recognize them, then would we treat the new being as an entirely new human person or would we need to have a new framework to account for such cases? Furthermore, once we consider the new being as human, we face issues of identity, such as which one is the original being, and whether this question is even a fair one to ask in the first place. In our opinion, this question would not truly matter since they start out exactly identical, otherwise, the procedure would be meaningless, and thus

both beings would have an equal claim to being the original. Going further, we are confronted with many more disturbing possibilities. For instance, once it is possible to make a digital copy of a person, there is nothing to stop any number of further replications, which further exacerbates the issue.

We also face a difficult ethical issue when we think further about the consequences of such a technology. Once a mind exists on a computational platform, it is much simpler to change or manipulate it compared to when it was a "wet", strictly biological brain. This kind of manipulation is dangerous, of course, when we think of bad actors. However, it is also possible, indeed easy, to make oneself more hardworking, or smarter, or any number of other enhancements that may be desirable and voluntarily undertaken. Another issue that is related to identity relates to the possibility of merging minds which could perhaps be possible soon after the above technology matures. At this point, individual identity may become extremely fluid and hard to describe in definite terms.

It is difficult to empathize with such beings, and it would be a stretch to see these extremely advanced and significantly transcendent beings as humans because they possess not only vastly enhanced abilities that humans possess as well, but also an array of abilities and possibilities that may not even be conceivable to us. Consider the framework we proposed. In the first place, giving such beings rights equal to humans seems completely inconsequential as they would be far more advanced than ourselves. Lastly, it would be fair to say that humans without this kind of computational intelligence would lack the capability to ever comprehend the thought processes of these beings even with mutual cooperation and efforts, and for this reason, they would seem to us to be completely alien or 'other'.

In an era when our bodies are full of nanobots acting as artificial cells and it is possible to upload our brains into a completely non-biological substrate, we have undertaken a discussion about what happens then to our ideas of what it means to be 'human'. Furthermore, when we shed what was given to us by evolution and our intelligence will explode to unprecedented levels, our ethics will need to be deeply questioned. For instance, if the enhancement is proved to make people more rational and empathetic, traits we now deem to be desirable in people, will it then be unethical to refuse to enhance oneself given such benefits to society at large? We end our discussion here, but answering such questions will become increasingly relevant in the coming decades.

6 Conclusions

We have presented arguments to demonstrate that categorical definitions of human nature fail to be comprehensive in a variety of scenarios. We then presented a moral framework for identifying humans, rather than defining them, avoiding this categorization problem. This kind of moral framework is more dynamic and accommodating to social changes, making this identification framework as fluid as we want it to be, or as fluid as our ever-changing morals demand it to be. This is more in tune with the reality of human existence; our standards always change, our views of the world around us always change, our view of ourselves always changes. Our morals are the best reflection of how we view ourselves, and this is why a moral framework is robust and makes the most sense.

Thereafter, we explore the question of what remains of man after the Singularity. The answer to this question depends on the extent of one's imagination. Within the scope of the cases that we chose to discuss in this paper, some technologies do not truly take away the essence of being human, while some other technologies would cause everlasting and irreversible changes that would completely transcend our notions of humanity.

The strength of our framework lies in how much consensus we can achieve with it. It is not in how inclusive or exclusive it is in identifying other beings. If, by our framework, we completely reject a 'transhuman' being as 'non-human', our conclusions are not called into question unless there would be significant disagreement among us today as to whether such beings are different enough from us to merit

their exclusion from the rights and advantages that humans enjoy. Our framework is more suited to such discussions of future radical technological changes because it agrees better with our intuitions, in the situations that arise with the Singularity, and holds up more robustly when we actually think about the moral issues that we would need to resolve.

As a concluding remark, we draw attention to one final thought. Our moral framework suggests that all 'human' rights are equally-weighted when assigning 'human' status to a being. Perhaps we must acknowledge, however, that the first and most fundamental of these rights, the right to life, is in fact more critical to the discussion, since all other rights must assume that the right to life has been given. If the right to life is given, this suggests that the life of this being is valued equally to the life of any 'human' today. Perhaps it is redundant to talk about a being's rights to education, healthcare and all the particularities of 'human' life, if we do not value its life equally to a human's. It seems natural, therefore, to attempt to pose our framework in a reduced form, by focusing soley on the right to life instead of general 'human' rights. In essence, if we were to pitch this being against an anonymous 'human' of today and were faced with the conundrum of having to choose who lives and who does not, would we value them both equally? Would this conundrum present a moral stale-mate where it would seem practically impossible to make a decision? If there is an inclination to value one life over another in the face of this conundrum, would this not automatically answer our question by exposing who we deem 'non-human'? Of course this presupposes two things about our current moral standards, namely that all 'humans' have equal right to life, in spite of whatever prejudices we may exercise, and that 'human' life is valued more than 'non-human' life. We imagine that in a future where moral standards have evolved such as to value 'human' and 'non-human' life equally, this reduced framework will cease to draw any meaningful conclusions.

We stress that this reduced framework does not condone the extermination of all evidently 'non-human' life. Rather, it illustrates that in the face of an extreme moral dilemma such as that described above, the decision we would make is highly telling of how we perceive ourselves against other 'non-human' beings, perhaps drawing on our most basic, primal instincts to preserve ourselves. The morality of such a proposition of course seems highly questionable, as it reduces the discussion to our lesser human impulses, which may end up being superficial and entirely appearance-based. It is for this reason that it still appears more reasonable to use our framework in the context of all 'human' rights and not solely the right to life. Nonetheless, there is inherent simplicity, almost triviality, in reducing our framework in this way, the applicability of which we leave open for discussion.

Word Count: 10976

References

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)*. American Psychiatric Association.

Callahan, D. (1995). Aging and the life cycle: A moral norm? In *A World Growing Old: The Coming Health Care Challenges*, pages 21–27. Washington: Georgetown University Press.

Caplan, A. L. (2009). Good, better, or best? In Savulescu, J. and Bostrom, N., editors, *Human Enhancement*, chapter 9, pages 199–209. Oxford University Press Inc., New York.

Falcon, A. (2015). Aristotle on causality. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2015 edition.

Fitzgerald, K. (2008). Medical enhancement: A destination of technological, not human, betterment. In Gordijn, B. and Chadwick, R., editors, *Medical Enhancement and Post-Modernity*, page 3955. Dordrecht: Springer.

Juengst, E. and Moseley, D. (2016). Human enhancement. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition.

Kurzweil, R. (2001). The law of accelerating returns.

Kurzweil, R. (2006). *The Singularity Is Near: When Humans Transcend Biology*. Penguin (Non-Classics).

McKibben, B. (2004). *Enough: Staying Human in an Engineered Age*. New York: St. Martins Griffin.

Parens, E. (1995). The goodness of fragility: On the prospect of genetic technologies aimed at the enhancement of human capacities. *Kennedy Institute of Ethics Journal*, 5(2):141–153.

Robert, J. S. and Baylis, F. (2003). Crossing species boundaries. *The American Journal of Bioethics*, 3(3):1–13.

Savulescu, J. (2009). The human prejudice and the moral status of enhanced beings: What do we owe the gods? In Savulescu, J. and Bostrom, N., editors, *Human Enhancement*, chapter 10, pages 211–247. Oxford University Press Inc., New York.

Shanahan, M. (2015). *The Technological Singularity*. MIT Press.

Turing, A. M. (1950). Computing machinery and intelligence.

Ulam, S. (1958). *Tribute to John von Neumann*, volume 64. American Mathematical Society.

UN General Assembly (1948). Universal declaration of human rights (udhr).