Can transhumanism lead to happiness?

-

A thought experiment

Bachelor thesis
submitted in partial fulfilment of the requirements
to obtain the academic degree of
“Bachelor of Science”
in Liberal Arts and Sciences
at the University College Freiburg
of the University of Freiburg

submitted by

First and Last Name:	Benjamin Janis Schult
Matriculation Number:	4745437
Majoring in:	Life Sciences
First examiner (supervisor):	Prof. Dr. Regine Kather
Second examiner:	Dr. Simon Büchner
Word count:	11,597

Date of submission: 2nd April 2024

Declaration

I hereby declare that I am the sole author and composer of my bachelor thesis and that no other sources or learning aids, other than those listed, have been used. Furthermore, I declare that I have acknowledged the work of others by providing detailed references of said work. I hereby also declare that my bachelor thesis has not been prepared for another examination or assignment, either wholly or excerpts thereof. I declare that the submitted electronic version of the bachelor thesis corresponds in content and formatting to the hard copy.

_____ _____

Place, date Signature

Acknowledgements

I want to thank my supervisors Prof. Dr. Regine Kather and Dr. Simon Büchner for their openness of mind and for giving me the freedom and opportunity to choose a topic corresponding to my own interests. I want to thank Prof. Dr. Michael Müller for inspiring and encouraging conversations that sparked the idea to focus on transhumanism as a topic combining the life sciences and philosophy. Moreover, I want to thank all my proofreaders for their help and invaluable feedback over the course of this work's composition. Finally, I want to thank Dr. Jürgen Kurth for his unapologizing authenticity, uncompromising sincerity and his loving, encouraging support over the last years of my therapeutic process that helped me to find the courage to follow my own path and to become more and more who I am.

Abstract

Transhumanism is one of the most influential ideologies of the 21st century. It principally directs the development of technologies in the field of artificial intelligence, synthetic biology or nano technology. The key focus of this ideology is the transcendence of all human limitations by means of technology. Following David Pearce's ideas, this "abolitionist project" shall lead to eternal happiness, not only for mankind but for nature in general. Transhumanist beliefs are deeply rooted in the dualistic, secular, and anthropocentric notions of Renaissance-humanism, in the externalist-scientific optimism of the Enlightenment as well as in the relativistic understanding of meaning and self in postmodernism. Its dualistic concept of hedonism, of pain and pleasure as being mutually exclusive, leads to its focus on abolishing any kind of pain or painful limitation (thus "abolitionism"). Thus, whether transhumanism will lead to happiness depends on its ability to eradicate suffering as such. Due to the important information function of pain, it cannot be eliminated directly. Rather, all its sources such as emotional pain or death must be remediated. The elimination of death is improbable due to the finite nature of our physical universe to which any substrate of consciousness is bound. The elimination of emotional pain is improbable as the necessary control to ensure it is a limitation itself causing the pain of unfreedom. Therefore, transhumanism is bound to fail. Yet, through the despair of failure it might lead to a surrender of its beliefs, finally opening to the experience of happiness beyond pain. The radicality of the transhumanist personality becomes its greatest strength and centrally discerns it from the moderate externalism of contemporary culture. Transhumanism's opposite, internalism, might lead to happiness in the same way, by failing to achieve it directly. Its interrelations to externalism and mysticism pose to be a promising topic for further investigation.

Contents

Disclaimer on gender specific language

For convenience and readability purposes, I do use masculine pronouns in this paper (which are my own). Naturally, all genders are addressed.

1 Introduction

Does our increasing use of technology lead to more happiness? If we would ask this question to a proponent of transhumanism, then the answer would most likely be an unshaken "Yes!". As it will become evident in this paper, the actual answer is rather a "Yes, it might, but for other reasons than suspected". Correspondingly, my thesis is that transhumanism might lead to happiness only because its attempts to achieve it will ultimately fail. Understanding why transhumanism does or does not lead to happiness might have drastic consequences for our modern way of life, as it probably is the most-influential ideology of the 21st century. The development of many of our cutting-edge technologies, especially in the field of life sciences, such as artificial intelligence, brain prosthetics or gene therapy are fueled by the belief in this world view[1]. A change of attitude towards transhumanism might deeply change our handling of these new technologies. In general, transhumanists want to improve the human condition with the power of technology. They belief that the self-determined use of technology will ultimately free us from the vices of life such as ignorance, sickness, and even death[2]. They belief that, thanks to all this technological power, mankind will finally become like God. This new human, or "homo deus"[3], as Yuval N. Harari calls this vision, will transcend all limitations, all-powerful, all-knowing, always happy. However enticing this grandiloquent story might be, whether all of this will become an actual reality, or even whether it is desirable, is far from being self-evident. Therefore, I will use this paper to investigate the question whether transhumanism can lead to eternal happiness, its assumed end goal for humanity, when its basic assumptions are taken seriously. I will subject transhumanism to the test against its own premises. My aim is not to discuss whether it will

[1] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 4-10.
[2] Bostrom, *The Transhumanist FAQ*., p. 1.
[3] Harari, *Homo Deus: A Brief History of Tomorrow*.

be sensible to follow transhumanist ideas, as Francis Fukuyama for instance did at length[4]. Instead, I will investigate whether we could reach the goals of transhumanism if we decided to attempt it. To do that I will, firstly, introduce transhumanism as an ideology and movement with its various definitions, history, and key characteristics. This will set the testing ground on which I will conduct a thorough examination of the assumptions which transhumanism builds on and in how far they allow for happiness to be achieved. This examination will take the form of a thought experiment. Finally, I will discuss the consequences of my findings in terms of the societal and individual evaluation of transhumanism.

2 Theoretical background

With the aim to test transhumanism concerning its ability to lead to happiness, one naturally must wonder why happiness was chosen as the reference point and what concrete understanding of happiness is applied? This will be our first focus before outlining transhumanism as an ideology afterwards.

2.1 Happiness

Transhumanism aims to transcend all human limitations. This can be framed as the goal of achieving superlongevity, superintelligence, and super wellbeing[5]. Some theorists, such as David Pearce, would argue that super wellbeing is the ultimate goal of transhumanism (or life) because immortality and superintelligence can be seen as mere means to get rid of the obstacles of death and ignorance that stand in the way of eternal happiness[6]. I go with Pearce's view and therefore focus on the transhumanist claim to reach eternal happiness in

[4] Fukuyama, "Transhumanism"; Fukuyama, *Our Posthuman Future*.
[5] Ross, *The Philosophy of Transhumanism*., p. 168.
[6] Pearce, "The Hedonistic Imperative."

this paper. Moreover, because of this, David Pearce can be seen as the key reference for this journey.

The understanding of happiness that is applied is crucial to how we answer the central question of this paper. As we will see over the course of this work, transhumanism is not able to lead to happiness in its narrow understanding as an absence of pain. Nevertheless, it might if we understand happiness and pain not as mutually exclusive. I will therefore use the term "happiness" when I simply mean the narrow understanding by transhumanism and the term "meta happiness" when I mean an understanding beyond the exclusive dualism of pain and pleasure (/happiness).

The ethic that transhumanism's understanding of happiness is mainly based on is called negative utilitarianism. It is a view in which happiness is a state only possible to experience in the absence of pain. Happiness and suffering are diametrically opposed. Due to that, the experience of happiness is dependent on the suffering-free constitution of the human being (not experiencing pain e.g. because of a disease such as cancer). The elimination of suffering (thus "negative") becomes the main goal of this view. [7]

2.2 Transhumanism

2.2.1 Etymology

The term *transhumanism* was in its modern form first used by Julian Huxley (1887-1975) in his book *New Bottles for New Wine*[8] from 1957. He coined this term to discuss a new ideology which proposes a radical change of human nature. An ideology that aims to transcend humanity and was thus called *trans*humanism. The "trans" in transhumanism already foreshadows its key characteristic of aiming to transcend all human limitations.

[7] Pearce.
[8] Huxley, *New Bottles for New Wine*.

2.2.2 Definitions

There does not exist one uniform definition of transhumanism. Various authors use various versions to highlight particular aspects such as the technological focus or its historical roots. The following version is used by *Humanity+* (formerly *World Transhumanist Association*), one of the leading organizations to support transhumanist interests, featuring many of the leading heads of the movement, such as David Pearce, Nick Bostrom, Max More, or Anders Sandberg[9]:

> [Transhumanism is…]
>
> *(1) The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to eliminate aging and to greatly enhance human intellectual, physical, and psychological capacities.*
>
> *(2) The study of the ramifications, promises, and potential dangers of technologies that will enable us to overcome fundamental human limitations, and the related study of the ethical matters involved in developing and using such technologies.* [10]

Philosopher Benjamin Ross puts it this way:

> *Transhumanism is a cultural movement which advocates a philosophy predicated on the argument that humans ought to transcend the limits imposed by our biological heritage.*[11]

[9] "Leadership." https://www.humanityplus.org/leadership
[10] Bostrom, *The Transhumanist FAQ*., p. 4.
[11] Ross, *The Philosophy of Transhumanism*., p. 1.

> *Transhumanists want to abolish the suffering which is a consequence of human finitude by making the boundaries imposed by death, ignorance, and psychological pain obsolete.*[12]

So why is happiness as a reference point ideal for discussing transhumanism? Following Ross' definition, we see that its aim for happiness includes all its other aims as well, it captures its roots.

Transhumanists, such as Nick Bostrom, often claim historical roots of the transhumanist movement in the thought of the great minds of secular humanism and the Enlightenment[13]. Thus, the following investigation of its historical inspirations will help to answer the question: What is transhumanism in its essence and how did it come about? Subsequently, we will be ready to start investigating its capabilities to achieve happiness.

2.2.3 Historical precursors

Ancient roots of the theme

Central themes of transhumanism, such as the overcoming of death, can be found in scriptures from various cultures in ancient times already. Probably the oldest example is the Babylonian *Epic of Gilgamesh* (approx. 2000 BCE)[14].

Humanism

One of the main proposed precursors of transhumanist thought is the movement of (secular) humanism, especially the humanism of the Renaissance. In that time, focus of life shifted towards the non-religious life before death resulting in a strong anthropocentric secularism. Renaissance-humanists believed in human agency and concentrated on the development of

[12] Ross., p. 11.
[13] Bostrom, *The Transhumanist FAQ*., p. 38-40.
[14] Bostrom., p. 39.

humanitas, the human qualities, such as compassion, benevolence, righteousness, or dignity[15]. Human nature was seen as not perfect by birth, but its humanity had to be developed by means of education and cultural refinement, i.e. by a development of character[16].

Enlightenment

Out of the secularity and anthropocentrism of Renaissance-humanism grew the movement of Enlightenment with its focus on science and rationality, another strong influence for modern transhumanism[17]. The Enlightenment is characterized by a purely optimistic view of the human as a rational being[18]. Its focus laid on the scientific method to study and conquer nature for human betterment as the renown Francis Bacon (1561-1626) proposed[19]. Scientific progress was a matter of indefinite advancement[20]. The movement put an increasing emphasis on the individual, e.g. by claiming individual rights[21]. Rules for living and the limits of knowledge were not given by a divine power anymore but based on universal principles accessible to human beings by virtue of their own rationality as Immanuel Kant (1724-1804) affirmed[22]. This individualism intensified later in postmodernism where universal principles are denied, and the individual becomes the creator of all meaning, purpose, and identity. It is also the basis for the transhumanist idea that the human being should decide upon its own fate.

---

[15] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 99.
[16] Ross, *The Philosophy of Transhumanism*., p. 49.
[17] Baumann, "Humanism and Transhumanism.", p. 70.
[18] Tiefensee, "Was ist der Mensch - was soll er sein? Humanismus - Antihumanismus - Transhumanismus - Posthumanismus.", p. 67.
[19] Bostrom, *The Transhumanist FAQ*., p. 39.
[20] Ross, *The Philosophy of Transhumanism*., p. 51f.
[21] Ross., p. 49.
[22] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 100.

Ideographical for the growing influence of scientific thinking was the expanding notion of mechanization in the conception of the world and of humanity. From Galileo Galilei (1564-1642) over René Descartes (1596-1650) and Issac Newton (1642-1726) to Julien Offray della Mettrie (1709-1751), nature and later the human being itself were seen as governed by natural laws resembling a machine[23]. The human being became a comprehensible, reproduceable, mechanical mechanism[24], the foundation of modern reductionism, a characteristic feature of transhumanist's concept of man.

(Post-) Modernity

From Charles Darwin's (1809-1882) theory of evolution transhumanism lends the idea of life (and humanity) as an ongoing, accessible process without a fixed nature that must be retained. Its comprehensibility allows for its manipulation. The biological process of evolution can be directed by (trans)humanist intentions[25].

In postmodernism this idea is expanded beyond biology to all meaning and purpose of life. It influenced transhumanist thinking in that it dispensed with the idea that there are a universal nature or universal principles that human life is governed by. The world is seen as a subjective construct and thereby only what we want it to be. Based on this relativistic subjectivism all external limitations to the realization of a self-created identity are seen as cruel and their abolishment thus demanded[26]. This relativism gives transhumanists the freedom to create a 'new human being' (the trans-/posthuman) in whatever way they please, as there is no innate nature nor principle that they must surrender to. The human being is what we make it to be[27].

---

[23] Kather., p. 26.
[24] Ross, *The Philosophy of Transhumanism*., p. 49f.
[25] Ross., p. 52.
[26] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 100.
[27] Gubrium and Holstein, "Grounding the Postmodern Self."

Another philosopher that is often discussed as proto-transhumanist is Friedrich Nietzsche (1844-1900). Some transhumanists, such as Stefan Sorgner (*1973), see a similarity between Nietzsche's concept of the "overhuman" and the transhumanist ideal of the "posthuman"[28]. Both concepts portrait an overcoming of the current human being. Other philosophers, such as Nick Bostrom or Benjamin Ross disagree with this interpretation by referring to Nietzsche's concepts of the "last human", "amor fati" and Nietzsche's' idea of personal growth[29]. For them, both concepts show Nietzsche's focus on the split between the present and a transcendent ideal as the key problem that should be overcome and not humanity itself. "Amor fati" in particular demands an acceptance and love of one's fate, contradictory to a technological change of oneself. The controversy stresses the common notion of transhumanist thinking to seek the salvation of humanity in a transcended, technology-enabled ideal (last/trans-/posthuman), superior to the current form of human beings.

Furthermore, the school of behaviorism has influenced transhumanist views decisively. The behavioristic conception of the human being completely ignores the existence of inwardness[30]. Humans are solely understood in terms of their behavior caused by environmental stimuli[31]. This view is a continuation of the scientific tradition and its objective approach to knowledge. What cannot be observed externally, whose existence is denied.

Lastly, with the advent of computer science in the second half of the 20th century this behavioristic understanding of the human being changed over in a focus on cognitive functions and later into a computational view of consciousness. Finally, transhumanism conceives of the human being as a mere pattern of information or data

[28] Sorgner, "Nietzsche, the Overhuman, and Transhumanism."
[29] Bostrom, "A History of Transhumanist Thought"; Ross, *The Philosophy of Transhumanism*., p. 145.
[30] Baumann, "Humanism and Transhumanism.", p. 69.
[31] Watson, *Behaviorism*.

(dataism/patternism)[32]. One's 'personality' is understood as a particular configuration of data that is not bound to the human body but could be realized in any desired substrate, e.g. silicon chips[33].

To sum it up, transhumanism is built on the secular anthropocentrism of humanism and the scientific-reductionist, techno-optimist thinking of the Enlightenment. This led to a relativistic, self-determined and externalist notion of the individual in behaviorism and postmodernism, which greatly influenced the movement, getting rid of the idea of any fixed constant that would bound human identity.

Having studied the precursors of transhumanism, one might wonder now how these translate into the key assumptions that transhumanism is based on. What are the convictions that define transhumanism concretely?

2.2.5 Key characteristics & assumptions

My aim is to test transhumanism's potential to realize eternal happiness against its own premises. The following overview of key assumptions will therefore also function as a reference for the subsequent thought experiment. This fact also leads to the question of how the exclusively dualistic notion of happiness, which Pearce's abolitionist project is based upon, comes about.

Dualism (the separation of I and the world)

Transhumanism is deeply rooted in humanistic dualism[34], i.e. the position that there is a separation between inside and outside, the self and the world, good and bad.

[32] Ross, *The Philosophy of Transhumanism*., p. 83-86.
[33] Kurzweil, *The Singularity Is Near*.
[34] Crinson, "The Dialectics of the Biosocial: Addressing Ontological Dualism and Contemplating Transhumanism."

Secular Anthropocentrism

Based on this dualistic understanding, transhumanists see the human being as separated from nature. Nature is simply the objective environment in which we are embedded. Because of this separation, humanity can disregard any form of external power such as God and can make itself the center, purpose, and measure of all things (as Protagoras might have said). With the renunciation of any external power, humanity must focus on the life before death. Transhumanists understand mankind as the most sophisticated product, the tip of evolution. A process which can now be taken over and directed by humans themselves. [35]

Progressivism / pessimism

Along with the separation from nature, humans are able to see that nature was not benevolent to them and created the human as a defective being with many limitations (such as pain and death; pessimism). Yet, as the human being is the purpose of all things, the improvement of the human condition towards less suffering is imperative (progressivism). Whereas Bacon advocated for a change of our natural environment, in transhumanism mankind itself becomes the object of change and of its self-directed evolution. [36]

Liberal Egalitarianism

Building on this humanist notion of the imperative for improvement, transhumanism borrows the ideal of liberal egalitarianism, meaning that not only that the human being ought to be free from all external constraints but that humans should also have the freedom to determine their own destiny (liberty as key value). As they are all the center of purpose, this "right" applies to all humans in the same way (egalitarian) and lays the ground for its utilitarianism[37]. This liberalism becomes a central part of the transhumanist concept of man.

[35] Bostrom, "A History of Transhumanist Thought.", p. 2.
[36] Baumann, "Humanism and Transhumanism.", p. 71.
[37] Hughes, "Contradictions from the Enlightenment Roots of Transhumanism.", p. 629.

The human essence is understood by its constant strive for transcendence of all limitations, or as Kurzweil puts it:

> *The essence of being human is not our limitations - although we do have many - it's our ability to reach beyond our limitations.*[38]

Externalism/Behaviorism

The separation between inside and outside as well as the pessimistic view of one's nature (inwardness) leads transhumanism to adopt a strict externalism. This means that not only can the world and the human being only be studied, described, and explained from an 'objective' outside perspective in an explicit way (in terms of its behavior), but it goes even further than that in assuming ontologically this is all there is (no inwardness). [39]

Epistemological certainty / scientism / techno-optimism

Rooted in this externalism, transhumanism is convinced of the merit of science and technology. In this view, science is the only real source of knowledge. The conviction of epistemological certainty means that everything can eventually be known in scientific terms. There are no mysteries and no problems that could not be solved by applied reason. [40] Scientific knowledge and technology have and will improve the human condition for the better (less suffering). Technological advancement is the main mode to fulfill our human potential. With their techno-optimism transhumanists plea for the proactionary principle, i.e. focusing on the opportunities instead of the risks, or put simply 'do whatever is possible'. [41]

Physical Informationalism

Related to the beliefs in dualism, externalism and scientism, physical informationalism is of particular importance to the problem of consciousness. The premise is that human

---

[38] Kurzweil, *The Singularity Is Near*., p. 230.
[39] Watson, *Behaviorism*.
[40] Ross, *The Philosophy of Transhumanism*., p. 13f.
[41] Ross., p. 21.

consciousness is an informational phenomenon, i.e. arising from a complex pattern of contrast, e.g. binary activation states, states of electrical conductance or binary code (0/1). In this view, consciousness, similar to a software program, is only an epiphenomenon of the physical body (the hardware). The structure and interaction (the pattern) of the matter give rise to a perception of consciousness. [42]

Hedonism / Negative Utilitarianism / Abolitionism

In accord with the pessimistic understanding of humanity, the human condition is seen as dominated and characterized by pain and suffering. The goal of life is happiness, which is understood dualistically as pure pleasure (hedonism[43]) and absence of pain (negative utilitarianism). Followingly, the experiences which someone can have in life are seen as either good (pleasure) or bad (pain, death). Their meaning is not a matter of subjective interpretation. The possibility to experience pleasure despite pain is disregarded. Pain and pleasure are mutually exclusive. To achieve happiness, pain must be eliminated. David Pearce formulated this intention in his terms as "hedonistic imperative" or "abolitionism" (abolishing suffering). The aim is to abolish suffering for the greatest number of people and by that to create the most benefit for humanity (utilitarianism). Theorists, such as Pearce, expand this claim even to non-human animals and nature in total[44]. The existence of death and pain is seen as objective evil inherently limiting happiness. Its overcoming is thereby a subproject of transhumanism[45]. Building on externalism, this is attempted with the use of externalist-scientific-technological methods.

In short, the transhumanist hedonistic understanding of the world is deeply rooted in its foundational assumptions of dualism, pessimism, and externalism. As it builds on a combination of various more deep-rooted beliefs this hedonistic vision becomes

[42] Searle, "How to Study Consciousness Scientifically.", p. 1938.
[43] Veenhoven, "Hedonism and Happiness."
[44] Pearce, "The Hedonistic Imperative."
[45] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 121.

representative of the transhumanist worldview. The test of its ability to achieve the desired happiness thereby becomes the acid test of transhumanism as such. The following thought experiment will therefore tackle exactly this question. Can transhumanism (based on its own premises) lead to happiness?

3 Transhumanism thought through

3.1 The main goal

As we have seen, happiness is the ultimate goal of all transhumanistic efforts, be it in the eradication of sickness, the attempts to achieve immortality or the limitless enhancement of intelligence. Following the transhumanist understanding of happiness, suffering and pain are the main characteristics of the current human condition and the main roadblocks to achieving happiness. I will use the term ‘suffering’ to describe the general problematic condition of humanity and ‘pain’ for a more specific sensation. Yet, for the transhumanist there is no distinction between pain and its evaluation that some might call suffering. It ignores the possibility of interpretation. In the end, suffering is characterized by the perception of something that is at the very least emotionally painful, and which is opposing our yearning for happiness (e.g. also fear, shame, etc.). Therefore, suffering means pain, and pain means suffering. Over the course of this experiment, it will become clear that the exclusive understanding of happiness is transhumanism’s own defeat. Corresponding to the term meta happiness I will therefore use the term “meta suffering” when talking about the suffering caused by one’s own condemnation of pain to make this point clearer.

The eradication of suffering and pain in all its forms is the starting point and will remain the overarching goal on our journey. Say we were to put ourselves in the shoes of a convinced transhumanist, how could we achieve this goal? This is what this thought experiment is about.

3.2 A direct attempt

If pain is the problem we want to be free from, the most straightforward way to our success would be to alter the human being in a way that we cannot perceive pain anymore. If we disable the feeling for physical and emotional pain, suffering would be eradicated too. This is the idea behind David Pearce's "abolitionist project"[46]. To achieve this, we might use advanced neuro-informational technology such as a futuristic brain-computer-interface, psychoactive drugs or change our biologic constitution overall using methods such as gene therapy. A problem of course is that even though pain is an unpleasant perception it has a necessary information function for our physical survival (e.g. warning us of deadly heat, when touching fire). If we were simply to deactivate pain perception, we would soon die from a variety of causes that we could not prevent as we would not understand them in their dangerous potential. We would die for reasons we could have prevented when we would have noticed the warning pain.

But what if we were to substitute the painful, biological warning mechanism with a non-painful information channel? This technology could inform us about all kinds of dangerous situations such as a dangerously low body temperature or a sharp object in our immediate proximity without causing the direct physical perception of pain, without the "nastiness of pain" as Pearce would say. Would this work? Probably not. Even though we would have changed our channel of information, the painful situation itself would remain the same. Similar to the (emotional) pain we experience when we hear that a war has started in an area where beloved people live or when we hear about the upcoming economic recession and accompanying tax increases, in the same way we would still suffer from the (now non-painfully transmitted) information itself. This is because of the dualistic, fixed evaluation of experiences, meaning that there are objectively bad incidents in life and that their evaluation as "bad" is not a matter of subjective interpretation. Death is bad and as long as we can die

[46] Pearce, "The Hedonistic Imperative."

and know about it, we will suffer under the limitations of freedom it puts on us. Even with the described technology, we would still know that there is danger and death out there waiting for us.

Would this still be the case if we changed ourselves so that we could only perceive gradients of wellbeing as Pearce proposes? Would we still associate death with pain if we do not know the experience of pain? This question is tricky because it is hard to ultimately decide from our perspective upon the perception of reality of a being that would be fundamentally different from us. In the end, probably only the artificially created pain-free being itself could answer this question authentically, whereby for such a being the question probably would make no sense and thereby it would not be able to answer it. The problem is that this question would only be relevant for the kind of being that is not predestined to answer it, namely Homo sapiens. Nevertheless, there are some relevant aspects that make it questionable whether this attempt to create pain-free (and thus ever-happy) beings will be successful. The creation of beings with only gradients of well-being but no perception of pain is like the idea of creating beings that can see only gradients of white or light shades of grey but do not perceive black or dark shades of grey. It misses the point that every shade of grey, even the lightest grey imaginable, is only possible due to the influence of black pigment. Without black pigment there are no shades of white or light grey, there is only (absolute) white. Similarly, gradients of well-being are only sensible in a field of polarity between well-being and non-well-being. Even though one might find oneself always in the 'light half of experience' these gradients of well-being depend on the potential for the opposite. Now, one could argue that we might aim for eternal, total bliss without gradients of well-being, but this faces the problem of the necessity of differences for information processing as even Pearce acknowledges[47]. To orient ourselves, to avoid danger, to pursue anything at all in life, to process any kind of information and meaning we depend on polarity.

[47] Pearce.

To use an informational metaphor, one cannot code anything with only one symbol (e.g. only 1), and in the same way, life and experience would be meaningless. This wish for unipolarity is also not consistent with the transhumanists' informational understanding of consciousness that presupposes duality (0 and 1) as the ground of our algorithmic consciousness. This aspect will be discussed in more detail in the section about consciousness below.

To sum it up, the direct attempt to abolish the perception of pain in general, even though not completely impossible, does not seem sensible from our current point of view. Consequently, our goal must be to eradicate all dangers causing us to experience such suffering, i.e. we must *eradicate all sources of pain*. To eradicate the sources of pain, we must gain control over them. The main sources of suffering as I will discuss them here are sickness, death and (emotional) pain. In the following I will discuss each of these sources and how they could be eradicated little by little.

3.3 Sickness

3.3.1 Mental sickness

If we are not able to eradicate pain directly through altering the pain information pathways, could we instead optimize the brain towards an ongoing state of happiness? Could we create an ongoing state of pleasure and positivity similar to a MDMA trip? Or could we not invent even more sophisticated techniques that would make it possible to optimize the brain in a way that we would not experience depression, schizophrenia, or phobias anymore? To answer this question properly, we would first have to elaborate the root causes of conditions such as depression or schizophrenia, which already touches upon a huge debate in psychology[48]. It is strongly debated and therefore quite questionable whether mental illnesses are reducible to pure physical processes such as errors in the biological

[48] Rosenberg, "Contested Boundaries."

management of neurotransmitter release for example. But as we are testing transhumanism against its own premises, let us suppose it were that easy. Let us suppose we were able to optimize our brains to be always happy and mentally stable. Would that not solve the issue? Sadly, for our endeavor it would not. The reason is that even though we might always be in a state of pleasure, as we have seen before, the potential for experiencing pain and thereby suffering cannot be simply 'optimized' in this way. As long as we are able to experience pain, and may it only be because we know that it is 'out there', our chance for eternal happiness is diminished. The quest remains to focus on the sources of painful experiences.

3.3.2 Physical sickness

If our attempts to optimize our mind for happiness directly fail, could we at least optimize our body so that we do not have to experience the tormenting pains of physical sickness anymore? This would be the goal of a transhumanist medicine (and in many ways is the goal of our modern medicine as well)[49]. Using technologies such as genetic engineering or nano robotics we would try to fix all weak points of our bodies and thereby prevent getting sick. As it is impossible to tell whether or not this might be successful, let us suppose it were to be successful. Firstly, it might be that with the ongoing mechanization and optimization of the human body we would get rid of all which is alive in us, meaning we might die of mechanical substitution. As death, in whatever form it may be, is a painful vision, this would lead to a failure of our attempt to reach happiness. Secondly, even if we did not use mechanical technology but only an advanced alteration of our biological systems, and the quasi death through mechanical substitution would not occur, again our attempt would fail as there are still external factors that could cause a destruction of the body and thereby pain. As an example, in a storm a branch could break off from a tree and hit us hurtfully or a jealous, fellow human being could hurt or kill us. Death and external sources of pain (by nature or society) would still be a possibility. Can we control death, can we control nature

[49] Marcum, *An Introductory Philosophy of Medicine*., p. 77f.

and the whole of society? Transhumanists believe we might be able to in the future. This is what we will test next.

3.4 Death (or non-aging)

Transhumanists have come up with various ideas to tackle the problem of mortality. Some prefer the approach to stop aging and make the human body immortal (as Aubrey de Grey)[50], others directly dispensed with the human body as substrate of the future and focus on immortality of the person in a new substrate (as Ray Kurzweil)[51]. I will follow this and discuss both general approaches separately.

3.4.1 Physical immortality

Closely related to our attempt to cure sickness is the goal to reach physical immortality. Some authors go as far as to argue that sickness is indeed only the prolonged and stretched-out process of ageing and thereby a sub phenomenon of death[52]. We cannot eliminate death without at the same time eliminate ageing and sickness. Thus, many of the transhumanist efforts circulate around the topics of anti-ageing methods and advanced medicine. If we were to invent technologies that make us immune against illnesses and would stop our ageing process at a desired age, to prevent our death these technologies themselves would need to be immortal, i.e. infinitely self-preserving or self-repairing, themselves. Otherwise, they would stop functioning and again, we would age and die. Could such a technology exist? Following the laws of physics, it is rather improbable. Particularly the laws of thermodynamics, namely conservation of energy and the ongoing increase of entropy would make it impossible to prevent the entropic deterioration of our life-saving technology forever[53]. We would still age and be mortal after all. Transhumanism builds on the

[50] De Grey and Rae, *Ending Aging*.
[51] Kurzweil, *The Singularity Is Near*., p. 112.
[52] De Grey et al., "Is Human Aging Still Mysterious Enough to Be Left Only to Scientists?", p. 669.
[53] Penrose, *The Road to Reality*., p. 690.

presumption that the scientific understanding of the world is the only correct one. If now this understanding, including the laws of nature described above, if these laws are indeed true and immutable then physical immortality is impossible. Even if our current understanding of the universe is wrong so that we could generate new energy for infinitely long times, and thereby secure the self-preservation of our technology and of ourselves forever, even then the technology might fail due to external interferences of nature or humanity. These two factors still must be controlled.

Even among convinced transhumanists the hope for physical immortality is not very strong. As an alternative, the possibility of reaching mental immortality, i.e. immortality of our consciousness, is seen as a much more promising endeavor and is therefore more popular among them as well[54].

3.4.2 Mental immortality

Consciousness uploading or WBE ('artificial base substrate approach')

The mental approach to immortality is not only filled with more hopes than the physical one but it is also more commonly known in the general public due to the widespread discussions of AI and brain-computer interfaces such as Neuralink these days[55]. This vision is usually termed "consciousness uploading" or "Whole Brain Emulation" (WBE)[56]. What is meant by that term is the idea that similar to the software data on a computer we might be able to copy or transfer our consciousness to a base substrate (or hardware) other than our biological body, e.g. to an advanced neural network. In that way, we could live on even when our biological bodies would have expired their lifetime. On the road to realize this vision we are facing a number of obstacles, the first one being the "mind-body problem" in general. The mind-body problem describes the difficulty for philosophers and scientists to grasp what

[54] Ross, *The Philosophy of Transhumanism*., p. 39.
[55] Fiani et al., "An Examination of Prospective Uses and Future Directions of Neuralink."
[56] Ross, *The Philosophy of Transhumanism*., p.40.

consciousness means, what it is constituted of and how it relates to our physical body[57]. Is consciousness an epiphenomenon of the brain? Or does it have an independent reality and it only coincides with the activation pattern in the brain? Even if it is only an epiphenomenon of the brain, when do the physical processes of atoms and molecules turn into the experience of subjectivity? And how? These are some of the questions that philosophers, trying to solve the psycho-physical problem, deal with and some argue that we are far away from finding a definite solution to it[58].

As it is an important issue in philosophy and also becoming more and more relevant in the discussion of AI, I want to shortly elaborate on why the problem of consciousness is so difficult to deal with and why this makes consciousness uploading improbable.

When we start to think about consciousness, we have to acknowledge that we are not thinking about human consciousness in general but rather about our personal one. Based on our experience, we intuitively infer that other human beings experience a similar form of consciousness, and we might even infer that animals have some kind of consciousness or subjectivity, too. Nevertheless, the only form of consciousness we really know for certain is our own. This is the first predicament that makes it hard to study consciousness objectively. We do not have external access to it. This problem is also called the qualia problem of consciousness and was much debated in the second half of the 20th century (see e.g. John Searle's Chinese Room thought experiment)[59]. David Chalmers coined the term "the hard problem of consciousness" to describe this inability to explain subjective experience[60].

The second predicament is that even inside our subjective experience, our consciousness, similar to our self, is something that cannot be observed directly as such. It rather functions as some kind of 'background canvas' on which our subjective experience

[57] Fodor, "The Mind-Body Problem."
[58] Clarke, "Forget the Hype, We've No Idea How to Reach Human-like AI."
[59] Searle, "Minds, Brains, and Programs."
[60] Chalmers, *The Conscious Mind*.

(thoughts, images, etc.) arises. We do not see the canvas itself but only the colorful forms that are painted onto it. We can only infer to the canvas by the means of the paint, but we cannot observe it directly (figure-background problem).

The question arises, how shall we study something scientifically if it is only accessible from the (scientifically unreliable) subjective perspective and even then, still impossible to observe directly? And how shall we understand and replicate something that we cannot study scientifically? This is the problem that philosophy of mind, and with it all the other disciplines interested in consciousness, face. Creating technology for consciousness uploading would presuppose an explicit understanding of consciousness and how it comes into being. Only then would we be able to manipulate, control and transfer it precisely. Thought about it in this way, the problem of consciousness, and thereby of consciousness uploading, seems to be impossible to solve. The only possible way I can see by which a technology of consciousness uploading could really be created, would be by chance, without an explicit understanding of what it is. This might seem improbable due to the complexity of such a technology but is not necessarily impossible after all.

Does this solve our problem of mental immortality, at least hypothetically? No, it does not. Even if a technology for consciousness uploading would be invented by chance, the base substrate on which our consciousness would exist (the hardware) would, similarly to our biological body, be bound to the laws of nature discussed above in the section of physical immortality. It might have a much longer lifetime, but it will still be finite. It is no solution to *eternal* happiness.

If the physical constraints of the base substrate are the hard limit of mental immortality, could we then not create a consciousness disentanglement technology which frees consciousness from its physical constraints entirely?

Consciousness disentanglement (the 'no base substrate approach')

Transhumanism builds on a scientific worldview rooted in physical reductionism which basically says that everything is a form of physical matter and can be reduced to physical

processes[61]. In line with that is the basic tenet of computer science that information necessarily depends on a (physical) carrier medium[62] . If we take this to be true and bring this together with the transhumanist conviction that consciousness is a process of information[63], then consequently it would be impossible to create a technology that disentangles consciousness from all physical media. As we cannot detach consciousness from a material base substrate in general and as material base substrates are finite in their lifetime and underly the deterioration dynamics already discussed, the attempt for (mental) immortality through consciousness disengagement fails. Here again, the premises of transhumanism (physical informationalism) preclude its own chance of success. So, if we cannot disentangle consciousness from matter, is the quest for mental immortality finally lost? It is not, at least not necessarily, as there is one option left.

Consciousness conflation (all matter as base substrate')

The attempt for physical (and mental) immortality failed because of the limited lifetime of the physical base substrate which we cannot simply detach consciousness from. But what if, instead of using a particular technological device or no matter at all as base substrate, we were to use *all* the matter as substrate, what if we were to use the entire universe as medium? Could we reach immortality if our substrate were not limited in extent and not separated from the entirety of the matter of the cosmos? Even while artificial products of all kind decay, the universe as such seems to exist further on. So, if we were to use not a part of it but its totality as medium, our substrate of consciousness could, in one way or the other, persist. There are a few caveats and uncertainties to this attempt for mental immortality.

Firstly, it must be acknowledged that in our former attempts we could still imagine that our consciousness would somehow stay the same, even though it is transferred to another

[61] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 118.
[62] Tegmark, "Consciousness as a State of Matter."
[63] Bostrom, *Superintelligence*., p. 30.

medium. With the attempt of consciousness conflation, the idea of separated consciousnesses is difficult to sustain. For this approach to succeed, every single consciousness of mankind would have to be calculated through the entire process of existence. There could not be distinguishable parts (or partitions, to use the computer metaphor) that are responsible for the emergence of single consciousnesses. Every consciousness would have to be a function of the whole. Following the physicalist logic, a conscious system is bound to its material substrate and therefore an epiphenomenon of the informational processes occurring in that substrate (such as software represents electrical currents and states in a modern computer)[64]. This means that all consciousnesses we aim to use for conflation must be realized in and thereby represent the same base substrate processes. We would end up having innumerable identical consciousnesses. They would all be the same as they represent the same informational processes. We would end up with only one consciousness after all, thus conflation.

This new, unified consciousness, or "God consciousness" as we might call it, would arguably be very different from the consciousnesses we were aiming at making immortal. This new consciousness might be immortal, but it would not represent the individual human hoping to become immortal. In some sense, for the God consciousness to emerge the individual consciousness would have to vanish, i.e. to die. A unified consciousness, even though it might be infinite in lifespan, does not represent a solution to the wish for immortality of the individual.

There is yet another caveat to the idea of consciousness conflation. Even if we would accept reaching a unified consciousness as mental immortality of the individual, we have to test whether the universe is indeed infinite in lifespan. If not, even a unified consciousness based on the entire matter of the cosmos would not offer immortality. We therefore have to ask what will happen to the universe over time, i.e. with time going towards infinity? In

[64] Duncan, "Ontological Distinctions between Hardware and Software."

physical cosmology there are many different scenarios proposed of how the universe will develop in the far future, with three of them being especially popular. Yet, all these scenarios (Big Freeze, Big Rip, Big Crunch/Bounce) postulate the occurrence of a singularity that would mean the end of all information and physical structure[65]. The universe is considered to be finite in lifetime by state-of-the-art science. A unified consciousness would end with these singularities and therefore be finite as well.

We have seen now that the attempts to reach physical immortality, as well as mental immortality in all its different forms, may it be consciousness uploading, disentanglement or conflation, are bound to be futile based on our current understanding of physics, i.e. based on the transhumanist premise that the scientific perspective is the only one to be taken seriously. What remains then? If we cannot become immortal ourselves then at least we might become immortal in significance due to the immortality of an artificial product created by us. We would become immortally significant by being the creator of an indeed immortal species or object. This is roughly the position posthumanists take on the purpose of human life. We shall examine it as our last hope for immortality.

3.4.3 Immortality through a product (posthumanism)

Living on in a new, immortal, living species

Can we become immortal (in significance) by designing a conscious species, which not only outshines us in capabilities but will also outlive us through being immortal? The first problem of this idea is that our experience of consciousness is deeply embedded in our experience of being human and by that of being mortal. It will be very difficult if not impossible to imagine, let alone design, a species which is conscious (maybe even more so than we are) but does not know death. But suppose we were able to design such a species, would it be helpful for our quest? Unfortunately, it would not. Even if we were to design a new species, this species would be facing the same problems again as non-aging humans

[65] Adams and Laughlin, "A Dying Universe."

would do. If it underlies our laws of physics its lifespan would be limited by the finite nature of our physical universe as well. Some trans-/posthumanists argue that such a superior species, powered by the so called "intelligence explosion"[66], would be able to find a way to reach immortality by inventing an advanced technology and changing the very laws of nature[67] (or even that we humans might reach that power). Even though this possibility cannot be ruled out completely, unfortunately for our quest, it seems rather like a desperate "out of the blue" idea without a realistic base, meant to save the hope for immortality. This possibility does not adhere to the principles of transhumanism, such as scientism, as the conception of natural laws in science means particularly that such laws are independent from our influence (objective reality).

What might be possible, even though it seems unlikely from a contemporary physics perspective, is that our conception of the natural laws is not fundamentally wrong (it is clearly not perfect as we do not have a theory of everything yet) and that a superior species might be able to improve upon our understanding of nature and our ability to manipulate it greatly. Whether this species might be able to change the course of the universe altogether seems impossible to predict, but as of now we have not much reason to believe it will be. Furthermore, even if it would become powerful enough to change the course of nature, it would still have to navigate the dangers from inside its own species. If immortality of significance is at best unlikely by creating a new, living species, could we achieve this goal at least by creating an indestructible artificial object?

Living on in a non-living product

Due to the phenomenon of entropy, there cannot be any individual object or physical structure that remains the same for eternity, artificial objects included. This approach faces the same difficulties as the attempt for physical immortality. Moreover, this object would

[66] Bostrom, *Superintelligence*.
[67] Kurzweil, *The Singularity Is Near*.

not only have to be protected against destruction due to natural causes, but also from within our own society (e.g. by opponents of the transhumanist project).

Before we proceed, I want to take a short break and reflect upon where we are standing in our thought experiment. Up until now, we have seen that for us transhumanists to be able to achieve happiness, we have to eradicate suffering by eradicating pain. As we cannot eradicate pain directly, we have to eradicate all sources of it, such as sickness, death or misfortune. While the eradication of sickness might be possible with advanced technology, the attempt to avoid death and become immortal, may it be in terms of physical or mental immortality, or by immortality of significance poses to be unsuccessful due to the finite nature of the physical universe. At the very best, attaining immortality is improbable. If, for the sake of our experiment, we would suppose that we were able to attain immortality and therefore eradicate sickness and death as sources of suffering, we would still have to eliminate all sources of suffering that hurt us during our lifetime. We would have to prevent sources of pain from outside humanity as well as from inside. Therefore, this will be our next step of investigation: The control of nature as a source of pain from outside humanity (following the dualistic understanding of transhumanism).

3.5 Pain from outside humanity (control of nature)

Even if immortal, we still must deal with other sources of pain such as natural catastrophes as pain reception were still possible. For that we would have to be able to completely control nature around us. In some sense, human life has always been about controlling parts of nature to create our own habitat[68]. With the increasing power of technology this nature controlling capability arguably has been increased so far as we are nowadays practicing what is called “geoengineering”, i.e. the voluntary change of the earth by changing big scale phenomena such as producing artificial clouds, removing huge amounts of carbon dioxide

[68] Marsh, *Man and Nature*.

from the atmosphere or potentially block off solar radiation[69]. The history of increasing human influence on natural events could be extrapolated and interpreted in that a complete control and prevention of natural catastrophes might be possible in the future.

Besides catastrophes originating on earth there are further dangers that would have to be handled. The control of nature would have to be expanded outwards and include protection against events originating in the proximity of the earth such as asteroid impacts or solar storms. Projecting the advancement of our technology into the future, it might be possible to protect earth and humanity against such dangers from our solar system.

As the sun will eventually have used up all its energy[70], humanity will probably have to leave the solar system in the far future. We therefore would have to be able to control phenomena outside our solar system as well. These would include the strongest cosmic events known such as supernovae and gamma ray bursts[71]. Whether or not we will be able to create technology safeguarding us against such powerful events is hard to foretell. Maybe we will simply be able to predict and therefore avoid such events.

There is yet another more incalculable threat. We would have to protect humanity against a foreign intelligent species (aliens) with a hostile attitude towards us. This one is tricky because the existence of a much more advanced and therefore more powerful alien species is possible, if not even probable concerning what we currently know about the universe[72]. Consequently, this threat is impossible to rule out. The only hope that remains is that we will never meet with an alien species or that it will be either less powerful than us or at least not hostile minded.

If we were able to mitigate all those potential dangers for humanity in the future, we would still have to face the problem of the mentioned ultimate fate theories of the universe

---

[69] Caldeira, Bala, and Cao, "The Science of Geoengineering."
[70] Sackmann, Boothroyd, and Kraemer, "Our Sun. III. Present and Future."
[71] Livio, Panagia, and Sahu, *Supernovae and Gamma-Ray Bursts*.
[72] Armstrong and Sandberg, "Eternity in Six Hours."

which pose a huge potential for pain one way or the other. Similar to our attempt to avoid death of age (immortality), our attempt to avoid pain would lead to the necessity for complete control over the universe similar to an omnipotent God-entity. Followingly, the only hope for our quest is that our current conception of the universe is wrong or that humanity will indeed become all-powerful. But, even if mankind became an all-powerful species in relation to nature, we would still have to deal with the potential dangers originating from inside our species.

3.6 Pain from inside humanity (control of society)

For our quest for happiness to be successful we would have to eliminate all sources of pain. Even if we were able to mitigate the suffering caused by sickness, ageing, or natural dangers we would still have to deal with the danger we humans pose to each other and ourselves. To avoid these dangers, such as murder, torture, or emotional hurt, we have to control every single human being, we have to control our society, i.e. we need a complete social control. Based on transhumanist principles, our first attempt to achieve this would be by means of external control, e.g. technology (externalism).

3.6.1 External control

Control by fear of punishment (subsequent)

At first attempt, we could try to steer human behavior by creating a fear of punishment for doing the wrong things (murder, torture, etc.). We would steer by subsequent measures in the case a rule violation has happened. For example, we could create complex and powerful policing systems that prosecute every rule violation in the most encompassing manner. Similarly, we could reward positive (or non-negative) behavior to further increase the inclination to follow the rules. In the behaviorist tradition we would use operant conditioning[73] to influence the behavior of our subjects, i.e. all human beings. If the reward

[73] Wheeler, *Beyond the Punitive Society*.

is appealing and the punishment fearsome enough, this method might work to eradicate negative behavior, such as murder, in most cases. Nevertheless, it also confronts us with several problems. First off, however extreme our measures would be, it is impossible to preclude the possibility that somebody, for whatever reason, might not care about (or not think of) the consequences of his actions and still hurt somebody else. Therefore, the retroactive control by operant conditioning does not yield a guarantee for pain to be obviated. It would create a rule-based dictatorship of morals. Moreover, the punishment itself, as well as the fear created to bring the members of society 'in line', are a form of pain and suffering which we aimed to prevent. In the end, the attempt to eradicate suffering through external control using operant conditioning would cause the particular problem it was intended to solve in the first place.

To circumvent this problem, the members of society would have to be controlled in such a way that they want to follow the rules of moral behavior themselves without any desire for disobedience.

Control by mind manipulation (preventive)

In this type of external control, we would influence human beings into doing the right actions by manipulating their minds using all available methods such as omnipresent propaganda, cultural norms, genetic alteration, technological implants, drugs, pre-emptive policing, etc., yet without trying to scare or reward them. The result would be a society similar to the ones depicted in Aldous Huxley's "Brave New World"[74] or in the movie "Minority Report" from 2002[75]. In this scenario, we could avoid the problems of menacing punishment for any rule violation because no member of society would even think about breaking a rule, they would neither question their morals nor wish for anything else. They would live exactly the way

[74] Huxley, *Brave New World*.
[75] *Minority Report.*

they want (or were made to want). This approach may initially seem like the perfect solution, but it still has some caveats.

The approach of mind control would change us human beings into predetermined, unfree puppets, similar to unconscious marionettes controlled by their puppet masters. We would give up our free will, our power for self-determination and thereby violate a central premise of transhumanism itself (liberalism). We would have to become less conscious, less free and thereby more limited than we are now. We would create our own prison in our minds. This vision of an unfree future is deeply painful for transhumanists as freedom from all constraints (e.g. pain) is the very road that shall lead us to happiness. We see then that the attempt for external control by means of alteration of the individual's mind is futile. It does not prevent human suffering (to achieve happiness), but instead produces it.

When external control of society neither by subsequent, nor by preventive measures work to prevent pain and suffering caused by humans themselves, what options to achieve happiness do remain?

The problem of mind control, and external control in general, is the unfreedom it causes to the individual. Thus, the only solution to this problem is that the individuals must control themselves willingly to act in the right, non-harmful way without external interference.

3.6.2 Internal control

As we have seen, external control is problematic, so we must learn to control ourselves if we want to achieve happiness (following transhumanist premises). Here, we face another obstacle. The attempt to control ourselves by means of our own consciousness does not align with the transhumanist principles of techno-centricity based on externalism. The attempt to control oneself internally does rather resemble the aim for a development of character in Renaissance-humanism[76] or other ancient self-improvement practices common to the Stoa

[76] Kather, *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*., p. 99.

or to various East-Asian religions (e.g. Buddhist meditation)[77]. The internal approach lies outside of the transhumanist line of thinking and would neither be aimed for by transhumanists, nor could it be accounted for as a success for transhumanism if happiness would be achieved that way.

One might wonder now whether transhumanists are really that strict about the externalism issue? Would they not simply do whatever is possible to achieve happiness, even if it would be an internal approach not using technology? After all they are deeply influenced by utilitarianism, so if something proves to be useful, they will vouch for it, won't they?

When we go back to the premises of transhumanism, we can see that externalism is a central characteristic of this ideology. The sympathy for utilitarianism, the scientific method, behaviorism and even its relation to postmodernism are all based on this foundation. The complete lack of inwardness in behaviorism[78] or the idea that there is no narrative-independent self in postmodernism[79] are telling evidence for that. Accordingly, the transhumanist personality is one that seeks for the explanation of life and the solution for happiness always outside of oneself. The avoidance of any confrontation with one's inwardness and the mistrust against anything internal and subjective are the core motivations for advocating transhumanism. This fear of oneself lies at the heart of it. Their externalism can also be understood as anti-internalism. Thus, the internal approach is disregarded. It would not be attempted because the reality of its very basis (inwardness) is denied.

[77] Goerger, "Moral Practice in Late Stoicism and Buddhist Meditation."
[78] Baumann, "Humanism and Transhumanism."
[79] Gubrium and Holstein, "Grounding the Postmodern Self."

3.7 Results

What was our last hope to guarantee the non-harmful behavior of mankind, i.e. internal control of oneself, has proven to be non-transhumanist by virtue. Internal control is no solution to eliminate the danger of pain caused by society. Thus, we have come to a point of dilemma. Happiness, that is our highest goal, could only be achieved if the attempt for control is completely abandoned (due to the pain of unfreedom), which in turn would make this happiness uncertain (due to flawed nature of humanity). We have reached a double bind, the dead-end of our journey. And now? What is the result of our thorough thought experiment?

In the end, we have reached the following insight: Under the given assumptions of transhumanism, e.g. that human beings are deficient beings, that the scientific-externalist approach is the only valid one, or that happiness is only possible in the absence of limitations such as pain, death, or unfreedom, *under these assumptions there is no solution to the misery of the human condition.* Either pain and suffering are indeed inevitable in human life and eternal happiness is impossible to achieve, or the given assumptions of transhumanism are wrong.

If transhumanism fails to achieve happiness based on its own premises, why did I purport that it might be successful in the introduction of this paper? The answer is simple yet surprising. Transhumanists will probably attempt to achieve happiness following their own ideals despite all counterarguments and improbability of its success, and they will fail. Their failure will pose an opportunity for the acceptance of pain, and of their own limits in general. By failing in every conceivable way, surrender will remain their only option left. Their acceptance and surrender will be the liberation from the double bind of being bound to suffer by fighting against the inevitable pain of life. By accepting pain, they will become free from their urge to abolish it. They might realize that their condemnation of pain was the source of their (meta) suffering and by letting it go, they will have freed themselves from it. They become able to experience the bliss that lies beyond the categories of pain and pleasure (meta happiness). So paradoxically transhumanism might lead to happiness particularly because its methods fail to achieve it.

One might ask now: How can we be sure that all transhumanist efforts will fail? Only because it seems to be improbable now, does not mean that our understanding of physics or

even of logical argumentation in general might not change in the future. After all, how can we say something is impossible if we have not tried it out? These questions are legitimate and hard to put aside. I will answer them indirectly in the following.

In response to the defense of transhumanism we might ask why cling so strongly to the technological solution at all? Why not accept its defeat? Moreover, one might wonder why so many people advocate for the transhumanist attempt to create happiness by use of technology if it seems so futile? What is the psychological root of supporting this ideology?

To answer this question, we will examine the ongoing debate between transhumanists and their critics, especially the debate around David Pearce's hedonistic imperative. The way Pearce reacts to criticism is a telling example of the character traits of the transhumanist personality, will answer the question and therefore be the focus of the following chapter.

3.8 David Pearce and the transhumanist personality

Pearce's vision for humanity builds on gradients of well-being without the nasty perception of pain, provoking the question of the interdependency of meaning and polarity as discussed before. Pearce responds to such objections that they are merely expressions of a psychological habituation and defense mechanism (his critics got habituated to pain and try to protect this part of their identity). Alternatively, he believes this "fatalism" (belief that suffering is inevitable) functions as a coping-mechanism[80] Furthermore, he argues that unenhanced humans (or "Darwinian humans" as he calls us)[81] are not able to understand the phenomenological experience of a being that, for biological reasons, cannot conceive of suffering. Both arguments show that there is no rational way to argue against his convictions. First, because every argument against the realizability of his paradise is interpreted as a problematic psychological condition that simply hinders one from being free enough to imagine a world without pain. Second, because he uses the opaqueness of subjective

[80] Pearce, "The Hedonistic Imperative.", p. 51.
[81] Pearce., p. 20.

experience to disqualify the credibility of us Darwinian humans to understand and discuss the phenomenal experience of the post-Darwinian, enhanced human being. If every argument against the hedonistic imperative is only due to a psychological condition, and in any case we as humans are not able to judge the posthuman anyways, then by that he makes his ideas non-falsifiable. In this way he never has to abandon his transhumanist belief.

Pearce responds to criticism of his ideas by questioning the psychological integrity of his critics. The same can be done with him as well and will be revealing for understanding the transhumanist personality. Indeed, Pearce states in interviews that his life since his childhood was characterized by suffering and by the suffering about the suffering in the world[82]. Consequently, he says that if he could, he would turn off existence as a whole, he would eradicate the world in the present form (because it is so painful)[83] . Therefore, one could wonder under which circumstances Pearce was raised, how painful his socialization and family life must have been to develop such a negative outlook upon existence. The same question can be applied to the transhumanist personality in general. Furthermore, Pearce naturalizes psychological suffering by believing it is genetically given (a given hedonic "set-point")[84], and by ignoring the possibility to find happiness beyond pain. These beliefs highlight his own psychological resistance against the fact that one has a responsibility for one's own outlook on life, that a state of unhappiness does not have to be chiseled in stone but can be overcome also during the lifetime of us "Darwinian" humans. By ignoring this possibility and by projecting the only solution to happiness into their technology-ruled, imagined utopian future transhumanists do not have to examine their own prejudices about life and themselves. They can stay the suffering persons they learned to be and blame biology for it. The conviction that, rather than their beliefs, the whole nature of humanity

[82] Sosis, "What Is It like to Be a Philosopher: David Pearce"; Despres, "Interview with David Pearce: Utopian Biology?"
[83] Sosis, "What Is It like to Be a Philosopher: David Pearce."
[84] Pearce, "The Hedonistic Imperative.", p. 72.

and biological life on earth must be changed[85] is telling of the severeness of this fear to face oneself, to enter this process of self-discovery with unknown outcome. In Pearce's case this might be an explanation why he sees a coping mechanism behind the convictions of people such as Buddha, or Jesus, or other people who allegedly achieved happiness despite their hardships in life[86]. He sees it as a coping mechanism because his own beliefs are a coping mechanism in the first place, helping to ignore the painful present reality. This might also be the reason why transhumanist convictions are getting increasingly popular and why our society is following many of its premises, at least in weakened form. These premises take the responsibility for the suffering away from the individual, blame evil nature and promise the salvation in a technological future, i.e. by external means. By doing that, transhumanism actually achieves the opposite of what it intends. It intends the empowerment of the human being, the becoming of God, but by externalizing responsibility to technology, it disempowers the human being into beings condemned for impotence. All their hope focusses on the mercy of external, technological manipulation. Transhumanists entertain the helplessness syndrome that they critique their disbelievers for[87].

Pearce beliefs that his utopian future of happiness can only be achieved by the abolishment of pain with technology and not by accepting it[88] (probably because he cannot accept this own pain). The self-conception of being defective becomes the transhumanist conception of man that I have discussed before. The sad truth is that by identifying happiness with the absence of pain, and by fighting against this pain, transhumanists fail to see the metalevel of their suffering. They fail to see that the source of their (meta) suffering is their own conviction. By remaining blind to the roots of their suffering they fail to see the possibility to experience happiness in the present, not only in a utopian future.

---

[86] Pearce, "The Hedonistic Imperative.", p. 67.
[87] Pearce., p. 51.
[88] Pearce.

Concludingly, the transhumanist personality can be characterized by an intense aversion of pain in combination with a lack of inward self-awareness resulting in the focus on external methods and a readiness for their radical implementation. This radicality is what discerns a transhumanist from a normal advocate of modern medicine for example. It differentiates them from the moderate hedonistic, externalist approach common to modern, western society[89]. This radicality also becomes transhumanists' protruding strength as it leads them into the dead-end with accompanying surrender and experience of the meta happiness beyond pain. A process that the moderate externalist might never experience because he sticks to his approach without going it the whole way and thereby might never experience the despair leading to surrender and freedom of his own fallacy.

3.9 Societal considerations

When the radicality of transhumanists is their key strength and when their convictions often include a fundamental change of humanity and nature as a whole, how shall one deal with them? Should their ideas be realized on a societal level? Or only on an individual basis? How will we deal with the impact of enhanced humans on unenhanced ones, e.g. in the competition for jobs?

The question of societal implementation poses a problem for transhumanist success. As we saw, transhumanism can lead to happiness only if it is allowed to attempt everything in its power and consequently fails. If a transhumanist alteration is restricted only to the individual level and not allowed on a general scale it might be that this denied change of environment might keep a transhumanist from realizing the dead-end of his approach. One might claim, similar to Pearce, that one is still unhappy because one still lives in an unenhanced environment creating pain, not in paradise. Like a moderate externalist, he might be trapped in his unfulfilled convictions. As modern society in general is moderately

[89] Veenhoven, "Hedonism and Happiness."

externalist and arguably unhappy with it, there is a growing push towards more radical versions of externalism. The growing influence of transhumanism is therefore a logical consequence of the scientific-externalist mindset of western society and its ongoing failure to achieve happiness. Suppose there will not be a miracle that will lead to the drop of the externalist ideology altogether, the only logical continuation of this situation seems to be an intensification of externalist efforts until they fail ultimately, leading to an acceptance of and surrender to the limitations of reality. The common urge for ongoing 'progress' in the sense of "more is better" is a manifestation of this state of mind (seeking salvation by more of the same). It seems that modern humanity is doomed to defy the structures of existence until reaching the abyss of self-destruction. Transhumanism is sometimes called the "world's most dangerous idea"[90], yet it is simply a logical continuation of convictions, such as externalism or exclusive hedonism, common to western societies. It thus constitutes a powerful object of reflection for one's own convictions upon oneself, the world, and the path to happiness.

3.10 A note on internalism

The externalist path to happiness might bring us to the edge of destruction before enabling happiness. One might wonder therefore whether there is an alternative not threatening our survival? As we have excluded the option of internal control from our considerations due to transhumanist premises, the question arises: What if we were not transhumanists, would internalism be a solution to the externalist dilemma?

In my point of view, internalism can lead to happiness in the same way that externalism does, by virtue of its own failure and subsequent surrender of one's beliefs. This might be surprising as the internal approach is advocated in many religions and schools especially for the purpose of reaching "enlightenment" or liberation from suffering, thus leading to

[90] Grassie et al., *H+/-*, p. 13.

happiness[91]. Nevertheless, methods such as meditation might lead to happiness only in the moment of accepting that they do not, that the limitations cannot be transcended. I understand "enlightenment" as the enlightening realization that all efforts to reach happiness (or unlimitedness) are in vain. In practice, it might seem from the outside that the method of meditation has led to this realization, where in reality, it was its failure.

To understand possible reasons for this, one might remember the limitations of observing and thereby understanding one's own consciousness (figure-background problem) touched upon in my thought experiment above. The nature of oneself, of one's own consciousness, remains an unsolvable mystery and thereby poses a limitation that cannot be transcended. It demands acceptance and surrender. Followingly, I see both internalism and externalism as possible paths towards mysticism, the surrender to the mystery of life[92].

Due to the restricted space of this paper, I will not further elaborate on the internal approach here. I end my considerations by emphasizing that this similarity of externalism and internalism, their relationship to mysticism and the dynamic of becoming free from limitations (such as suffering) by accepting them, might be a topic worth of further investigation by the philosophical community.

[91] Watts, *Psychotherapy East & West*.
[92] Rohr, *The Naked Now*.

4 Conclusion

In summary, transhumanism is an influential ideology in the modern world impacting much of the cutting-edge developments in technology. It shares many assumptions with our modern, western society, such as scientism, externalism, or hedonism. One of its key concerns is to achieve happiness by eradicating the sources of pain using technology. The presented thought experiment showed us that under the assumptions of transhumanism happiness is impossible to achieve. This is because transhumanism presupposes the elimination of pain and suffering for happiness to be experienced. At the same time, it is impossible to eradicate all sources of suffering following the transhumanist approach. By virtue of its radicality, transhumanism can result in the surrender of one's beliefs when coming to the final dead-end of its efforts. Paradoxically, it might lead to happiness because it is unable to achieve it. Its radicality is what discerns the transhumanist personality type from the mediocre externalist one common to western society. Apart from that, transhumanism is a logical continuation of the externalist mainstream ideology. It can serve as a surface for reflection, yet humanity might need to go the route of externalism all the way, until threatened by extinction, to become ready to surrender to the limitations of life and thereby experience a meta happiness beyond pain. The internal approach to happiness might be successful in a corresponding way, too, by failing to achieve its goal directly. The relation between internalism, externalism, and mysticism poses to be a worthwhile topic for further investigation.

Bibliography (Chicago-style)

Adams, Fred C., and Gregory Laughlin. “A Dying Universe: The Long-Term Fate and Evolution of Astrophysical Objects.” *Reviews of Modern Physics* 69, no. 2 (1997): 337–72.

Armstrong, Stuart, and Anders Sandberg. “Eternity in Six Hours: Intergalactic Spreading of Intelligent Life and Sharpening the Fermi Paradox.” *Acta Astronautica* 89 (2013): 1–13.

Baumann, Fred. “Humanism and Transhumanism.” *The New Atlantis*, no. 29 (2010): 68–84.

Bostrom, Nick. “A History of Transhumanist Thought,” 2005.

———. *Superintelligence: Paths, Dangers, Strategies*. First Edition. Oxford: Oxford University Press, 2014.

———. *The Transhumanist FAQ*. World Transhumanist Association, 2003.

Caldeira, Ken, Govindasamy Bala, and Long Cao. “The Science of Geoengineering.” *Annual Review of Earth and Planetary Sciences* 41 (2013): 231–56.

Chalmers, David John. *The Conscious Mind: In Search of a Fundamental Theory*. Philosophy of Mind Series. New York: Oxford University Press, 1996.

Clarke, Laurie. “Forget the Hype, We’ve No Idea How to Reach Human-like AI.” *Tech Monitor* (blog), 2021. https://techmonitor.ai/technology/we-have-no-idea-how-to-reach-human-like-artificial-intelligence.

Crinson, Iain. “The Dialectics of the Biosocial: Addressing Ontological Dualism and Contemplating Transhumanism.” *Journal of Marxism and Interdisciplinary Inquiry* 12, no. 1 (2021): 76–88.

De Grey, Aubrey, John W. Baynes, David Berd, Christopher B. Heward, Graham Pawelec, and Gregory Stock. “Is Human Aging Still Mysterious Enough to Be Left Only to Scientists?” *BioEssays* 24, no. 7 (2002): 667–76.

De Grey, Aubrey, and Michael Rae. *Ending Aging: The Rejuvenation Breakthroughs That Could Reverse Human Aging in Our Lifetime*. New York: St. Martin’s Publishing Group, 2007.

Despres, Jonathan. “Interview with David Pearce: Utopian Biology?” HEDWEB, 2005. https://www.hedweb.com/hedethic/interview.htm.

Duncan, William D. “Ontological Distinctions between Hardware and Software.” *Applied Ontology* 12, no. 1 (2017): 5–32.

Fiani, Brian, Taylor Reardon, Benjamin Ayres, David Cline, and Sarah R Sitto. “An Examination of Prospective Uses and Future Directions of Neuralink: The Brain-Machine Interface.” *Cureus*, 2021.

Fodor, Jerry A. “The Mind-Body Problem.” *Scientific American* 244, no. 1 (1981): 114–23.

Fukuyama, Francis. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. New York: Farrar, Straus and Giroux, 2003.

———. “Transhumanism.” *Foreign Policy*, no. 144 (2004): 42–43.

Goerger, Michael. “Moral Practice in Late Stoicism and Buddhist Meditation.” *Comparative Philosophy* 8, no. 1 (2017).

Grassie, William, Gregory Hansell, Hava Samuelson, Nick Bostrom, Aubrey de Grey, Natasha Vita-More, Sky Marsen, et al. *H+/-: Transhumanism and Its Critics*. Philadelphia: Metanexus Institute, 2011.

Gubrium, Jaber F., and James A. Holstein. “Grounding the Postmodern Self.” *The Sociological Quarterly* 35, no. 4 (1994): 685–703.

Harari, Yuval Noah. *Homo Deus: A Brief History of Tomorrow*. New York: Random House, 2016.

Hughes, James. “Contradictions from the Enlightenment Roots of Transhumanism.” *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 35, no. 6 (2010): 622–40.

Humanity+. “Leadership.” Accessed March 20, 2024. https://www.humanityplus.org/leadership.

Huxley, Aldous. *Brave New World*. London: Chatto and Windus, 1952.

Huxley, Julian. *New Bottles for New Wine*. London: Chatto and Windus, 1957.

Kather, Regine. *Die Verheißung Gesteigerter Lebensqualität: Philosophische Hintergründe von Künstlicher Intelligenz Und Transhumanismus*. Ostfildern: Matthias-Grünewald Verlag, 2022.

Kurzweil, Ray. *The Singularity Is Near*. New York: Viking, 2005.

Livio, Mario, Nino Panagia, and Kailash Sahu. *Supernovae and Gamma-Ray Bursts: The Greatest Explosions Since the Big Bang*. Cambridge: Cambridge University Press, 2001.

Marcum, James A. *An Introductory Philosophy of Medicine: Humanizing Modern Medicine*. Berlin, Heidelberg: Springer Science & Business Media, 2008.

Marsh, George Perkins. *Man and Nature*. Seattle: University of Washington Press, 2003.

Minority Report. Twentieth Century Fox, Dreamworks Pictures, Cruise/Wagner Productions, 2002.

Pearce, David. “The Hedonistic Imperative,” 1995.

Penrose, Roger. *The Road to Reality: A Complete Guide to the Laws of the Universe*. 1. publ. London: Cape, 2004.

Rohr, Richard 1943-. *The Naked Now: Learning to See as the Mystics See*. New York: Crossroad Publ. Company, 2009.

Rosenberg, Charles E. “Contested Boundaries: Psychiatry, Disease, and Diagnosis.” *Perspectives in Biology and Medicine* 49, no. 3 (2006): 407–24.

Ross, Benjamin. *The Philosophy of Transhumanism: A Critical Analysis*. First edition. Emerald Points. Bingley, United Kingdom North America Japan India Malaysia Chine: Emerald Publishing, 2020.

Sackmann, I.-Juliana, Arnold I. Boothroyd, and Kathleen E. Kraemer. “Our Sun. III. Present and Future.” *Astrophysical Journal* 418 (1993): 457–68.

Searle, J. R. “How to Study Consciousness Scientifically.” *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 353, no. 1377 (1998): 1935–42.

Searle, John R. “Minds, Brains, and Programs.” *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24.

Sorgner, Stefan Lorenz. “Nietzsche, the Overhuman, and Transhumanism.” *Journal of Evolution and Technology* 20 (2009): 29–42.

Sosis, Cliff. “What Is It like to Be a Philosopher: David Pearce.” What Is It Like to Be a Philosopher?, 2022. http://www.whatisitliketobeaphilosopher.com/david-pearce.

Tegmark, Max. “Consciousness as a State of Matter.” *Chaos, Solitons & Fractals* 76 (2015): 238–70.

Tiefensee, Eberhard. “Was ist der Mensch - was soll er sein? Humanismus - Antihumanismus - Transhumanismus - Posthumanismus.” In *Die Perfektionierung des Menschen? Religiöse und ethische Perspektiven*, 65–80. Münster, Westfalen: T. Bahne und K. Waldner, 2018.

Veenhoven, Ruut. “Hedonism and Happiness.” *Journal of Happiness Studies* 4, no. 4 (2003): 437–57.

Watson, John B. *Behaviorism*. Milton Park, Oxfordshire: Routledge, 2017.

Watts, Alan. *Psychotherapy East & West*. Novato, CA: New World Library, 2017.

Wheeler, Harvey. *Beyond the Punitive Society: Operant Conditioning: Social and Political Aspects*. San Francisco: W. H. Freeman, 1973.