Volume 6 · Issue 1 · 2021

Journal *of* **Philosophy**

contents

articles

With or Without the Self? Arguments in Favor of the Hindu Concept of the Atman over the Buddhist Understanding of the Anatman

Rocco Angelo Astore

CUNY: Borough of Manhattan Community College, USA
E-mail address: rastore@bmcc.cuny.edu
ORCID iD: https://orcid.org/0000-0002-5126-0211

Abstract

As found in the Upanishads, readers come to encounter many ideas regarding the "Self" as opposed to the "self," or the Atman versus the atman. Now, complicating matters further readers encounter the antithetical concept, of the Atman and atman, or the Buddhists understanding of the anatman as found in the Dhammapada. First, this piece will unpack the idea of both the "Self" and the "self," or the Atman and the atman, as understood in Hindu philosophy and theology. Next, this essay will then describe the Buddhists understanding of anatman, or the absence, or emptiness of the self, or the nullification of the self, or the atman as well as even of the Self, or the Atman. Finally, this piece will resolve by arguing in favor of the Self, or the Atman as well as the self, or atman over the Buddha's notion of the anatman.

Keywords: *comparative philosophy; Hinduism; Atman; The Upanishads; Buddhism; Anatman; The Dhammapada*

I. Defining the Self, or the *Atman*

As understood in the *Isha Upanishad*, readers come to find a most revealing quote describing the absolute reality that is the Self, or the *Atman*:

The Self is everywhere. Bright is the Self, Indivisible, untouched by sin, wise, Immanent and transcendent. He it is who holds the cosmos together.[1]

---

[1] *The Upanishads,* trans. Eknath Easwaran (Tomales, CA: Nilgiri Press, 2007), 58.

Now, this Self, or *Atman* is that which serves to be the Godhead in the *Upanishads*.[2] That is because only something immanent, or within and throughout all reality can be everywhere the same, or that the Self, or the *Atman* is indeed that which binds all of existence together.[3] Furthermore, one may find in this quote that the Self, or the *Atman* as transcendent, or as above and beyond everyday reality in the natural order, is truly that which cannot be subject to division.[4]

Moreover, by being above and beyond ordinary existence, one may find that this transcendent Self, or *Atman* is exempt from affectability, preserving the Self, or the *Atman* as perfect, or untouched by sin, and thus wise, for nothing can taint or restrain the expression of its goodness, or wisdom.[5] Finally, if one considers another quote, this time from the *Shvetashvatara Upanishad* he/she may find something more revealing about the immanency of the Self, or the *Atman*. That is, let readers consider the following:

> The Lord dwells in the womb of the cosmos, The Creator who is
> in all creatures, He is that which is born and to be born; His face
> is everywhere.[6]

From this quote, one may find that the Lord, or the Self, or the *Atman* possesses a power that inhabits all that is, leading to the role of the Creator as that which becomes the cosmos itself.[7] One may also find in this quote that as such a Creator of the cosmos itself, and as within all creatures, who is findable everywhere, the Self, or the *Atman* is indeed immanent, or within and throughout all life.[8] Finally, one last implication of this quote alluding to the Self's, or the *Atman*'s immanency is that the Self, or the *Atman* is everlasting.[9] That is because as being that which will be alive, there is the implication that its birth was prior, an instant of the past, while the Self, or the *Atman* as being that which is indeed alive is an instant of its present, and as that which will be alive one may declare that as being an instant of the Self's, or the *Atman*'s future. Hence, as being throughout all time, one may verily claim that the

---

[2] Ibid., 37-39.

[3] Ibid., 58.

[4] Ibid.

[5] Ibid.

[6] Ibid., 165.

[7] Ibid., 167.

[8] Ibid., 58, 167.

[9] Ibid., 165.

Self, or the *Atman* is not only eternal, but also immanent, or everywhere in presence the same.[10]

Furthermore, one may declare that this eternal and all-pervading Self, or *Atman* as that which is in all things, animate and even inanimate, by which all that is finds their common source of existence, is also that which gave way to the life of its fragments, or oneself who is a being who contains an awareness of the *Atman* by being an *atman* oneself.[11] In other words, the Self, or the *Atman* who manifest as the first cause, or origin point of all life, is therefore also that which emerges to be in each person and it is everyone who houses in the depths of themselves this everlasting and immortal node of all that is and is in reality and existence itself.[12] Lastly, let us further explore the idea of each of individual as selves, or *atman*(s) as understood in the *Upanishads*.

II. Describing the self, or the *atman*

As mentioned, the concept of the *atman* also refers to the individual soul, or vitality, or life-breath that all people possess the power to be knowing of, in an aware way.[13] Now, some features of this self, or *atman* that is distinct from the absolute reality that is the Self, or the *Atman* is that *atman*(s), or individual selves exists trapped in a more microcosmic reality, in which they undergo certain processes of which the Self, or the *Atman* is exempt. In other words, unlike the pristine and perfect reality that is the Self, or the *Atman*, people's fragmentary existence as miniscule selves, or *atman*(s) comes with the plague of desire, craving, and yearning.[14]

As such, people should overcome this realm of *samsara*, or the illusory world of separation and suffering, in which they live, that reel births, deaths, and rebirths, to achieve reunification with their original and absolute source, which is the Self, or the *Atman*.[15] Finally, this release from *samsara* is the liberating state of *moksha*, or that permanent state of joy, or bliss achieved by living a life that uncovers levels of consciousness in a most perfect way.[16] However, to reach this state of joy is no easy task, and as stated in the *Taittiriya Upanishad*, one must first peel away his/her state of being, associated with what he/she takes to be the material body, which attends only to the outside

---

[10] *The Upanishads*, 35-37.

[11] Ibid.

[12] Ibid., 159.

[13] Ibid.

[14] Ibid., 6942-6943.

[15] Ibid.

[16] Ibid., 241.

world.[17] That is because an awareness that is only of the material body and its wants, leads to only pleasure, or a temporary state of relief from desire, craving, and yearning, which promises only a deeper state of craving once such a transient state of pleasure subsides.[18] Instead, one ought to seek joy, which in Hinduism is a permanent state of awareness in which the self, or the *atman*, finds the Self, or the *Atman* within.[19]

Next, even if one unveils and becomes impartial to the material body, there is still the consciousness of his/her vitality, or *prana*, that he/she must also tame.[20] Furthermore, one is then to quiet, or still the mind in such a way that he/she may further himself/herself, or his/her *atman* to progress toward everlasting happiness, or joy, and afterwards he/she is to then uncover intuition, or *buddhi*, to alas reach such joy.[21] Now, with this development toward joy, one should also note that he/she is doing away with karmic baggage that can delay him/her from achieving *moksha*, and that state of reunion with the Self, or the *Atman* that is the abode of living and perpetual joy.[22] In fact, this *Taittiriya Upanishad* even informs us of the following:

> The Self is the source of abiding joy. Our hearts are filled with joy in seeing him Enshrined in the depths of our consciousness. If he were not there, who would breathe, who live? He it is who fills every heart with joy.[23]

Lastly, let readers now explore the Buddhists understanding of the *anatman*, the antithesis of the self, or the *atman* found in Hinduism, as well as the problems even the Buddha faced regarding this concept of the self, or the *atman*.

III. Entering the Concept of the *Anatman*

As stated in Chapter 12 of the Buddhists text the *Dhammapada*, readers come to encounter a startling claim; namely, the concept of the *anatman*, or the "without a self" doctrine that the Buddha expounded in his teachings.[24]

---

[17] Ibid.

[18] Ibid., 241-242.

[19] Ibid.

[20] Ibid.

[21] Ibid.

[22] Ibid.

[23] Ibid., 254.

[24] *The Dhammapada,* trans. Eknath Easwaran (Tomales, CA: Nilgiri Press, 2008), 153.

That is, one may find that the *anatman* is the opposite concept of the Hindu Self, or *Atman* as well as the individual self, or *atman*.[25] One reason for the Buddha's claim is that people's personalities, as impermanent, generate a logical contradiction when individuals attempt to acknowledge such inherent impermanency with both an unchanging identity, or self, or *atman* as well as with an unaffectable eternal, and thus, immutable reality beyond this one, the Self, or the *Atman*.[26]

Moreover, to complicate matters further, readers may find that Easwaran, the commentator of this writer's version of the *Dhammapada*, brings to light the argument by the Buddha that a concept like the self, or *atman*, or the Self, or *Atman*, as permanent features of reality and existence, may just be another attachment that people should overcome to seek true lasting spiritual liberation, or what the Buddha calls *nirvana*.[27] In other words, one may find that the notion of the self, or the *atman* and the Self, or the *Atman* cannot be everlasting.[28] That is because to the Buddha, the individual self, or *atman*, when instructed to seek the Self, or *Atman*, to attain *moksha* and entry into the absolute reality that is the Self, or the *Atman*, what one finds behind such instruction is at least an engagement in desire.[29]

Accordingly, the Buddha sees it that the mistake of Hinduism is asserting such a self, or *atman* as well as a Self, or *Atman*. That is because if Hinduism seeks to extinguish desire for all to eventually join in *moksha* and thus identity and equality in the Self, or the *Atman*, then how can it be that individuals are to seek such a Self, or *Atman*, through striving for want of reunification with such an absolute reality. Would it not be the case that if one follows his/her want for reunification with the Self, or the *Atman*, then even this is merely another way in which *samsara* persists?[30] In the opinion of this present essayist, this is indeed the standpoint of the Buddha.

However, in this same Chapter 12 of the *Dhammapada*, readers encounter another startle; specifically, that it is pragmatic, or useful to assume the existence of the self, or the *atman*.[31] That is because the Buddha believes that by acknowledging the self, or the *atman* insofar as allowing people to be morally and ethically responsible for their intentions and deeds, one finds that he/she can strengthen aspects of himself/herself that are integral

---

[25] Ibid.

[26] Ibid.

[27] Ibid., 154.

[28] Ibid., 153 & 154.

[29] Ibid., 154.

[30] Ibid.

[31] Ibid., 155-156.

to his/her moral and ethical well-being, such as his/her willing to do right and refraining from performing wrong.[32] In other words, even though there may be no permanent self, or *atman*, it is still of worth to assume that the individual is the one who intends and performs his/her acts.[33] Finally, that is because people may further themselves on the quest to learn to unlearn, or to develop in such a way that all of this world's conditioning, goes nil, when they assume themselves to be manifesters of their reality of which they are ultimately responsible.[34]

IV. Asserting the Certainty of the self, or *atman*, and the Self, or *Atman* over the Buddha's *anatman*

One reason as to why one may claim that the Buddhist understanding of the *anatman* pales when compared to the surety of the Hindu concept of the self, or the *atman* as well as the Self, or the *Atman*, is that if one takes his/her existence, in the Hindu illusory world of *maya*, or the Buddhists illusory realm of *loka*, it is still the case that what he/she senses in either realm, must derive from a source that can sense, or oneself, or an *atman*.

Moreover, even if one peels away this self, or *atman*, as does the Buddha, he/she may claim that this too is an err on part of the Buddha. That is because one still must admit that when he/she is dismantling the self, or the *atman* even if it is for the sake of emptying, or voiding oneself, to be the clearest channel for raw consciousness to come forth, as akin to the beliefs of the Buddha, he/she must still admit that he/she is untangling something rather than nothing. In other words, one major flaw of the Buddhist idea of the *anatman* is that if there were genuinely no self at all, then how can it be that what one takes to be his/her perceptions are indeed his/her own? At the same time, an even more pressing flaw of the Buddhist idea of the *anatman* is how can it be that one ought to empty oneself for clarity of mind, while nevertheless failing to acknowledge that there must be a self, working toward its own emptying, that is previous to an emptied self, for that self to be emptiable?

Also, the idea that a perpetually abiding Self, or *Atman*, as being, in fact, impermanent, and hence defying the very truth of a Self, or an *Atman* cannot be the case either, thus defying the Buddha's assertion of the *anatman*. That is because in Hinduism although individuals may find that even the natural order and all of reality is impermanent, it is permanently impermanent, or that it is the very essence, or nature of the Self, or the

---

[32] Ibid., 156-158.

[33] Ibid., 155, 159.

[34] Ibid., 159-160.

Atman to renew itself through the periodic clearing away of all that is only for the reemergence of itself after such episodic conflagrations.[35]

One need only look to the idea of *Brahma* as that which creates all life, and *Vishnu* as that which sustains all life, and *Shiva* as that which is the destroyer of all life, that is, the Hindu Trinity itself, to know that these three guises of the Self, or the *Atman* calls forth change in a way that is everlasting. Thus, there is not a true *anatman* that can ever come to be, for, even the Self, or the *Atman* despite appearing to undergo change, is verily that which causes and ensures the eternity of existence through its self-regulating nature. In other words, and as the philosopher Leibniz once wrote "... the universe which will be changed but not destroyed," the Self, or the *Atman* alike may alter the cosmos, as effects of its nature, but it itself will eternally and unchangingly be.[36]

Also, readers may further challenge the Buddha's take on the self, or the *atman*, as well as the Self, or the *Atman*, by debasing the notion that if one strives for permanent joy and enlightenment for the sake of entering an eternal and absolute reality one is, in fact, pleasure seeking and attached to this world of delusion.[37] Now, the mistake of the Buddha here is that there cannot be a teleological story compatible with the Hindu *Upanishads*, for the Self, or the *Atman* is infinite and eternal, as stated above, and because of this the Self, or the *Atman* is immune from beginnings or ends, and thus it is already self-sufficient, and in no need of a purpose to fulfill.[38] Finally, if individual people are fragments of the Self, or the *Atman* as selves, or as *atman*(s), then how can it be that each possesses a purpose that each must fulfill, if the power to uncover the Self, or the *Atman* is within, and thus not an external goal that he/she must strive for, in a way that necessarily renders him/her attached and craving of spiritual liberation as well as reunification with the Self, or the Atman?

Hence, if one understands the *Upanishads* as a story of how he/she can uncover himself/herself, or *atman* to find the Self, or the *Atman* within, instead of a quest for achieving reunification with the Self, or the *Atman* in a purely desiring way, that situates itself with reaching a source outside of us then the Buddha is not indubitably correct about the accuracy of his concept of the *anatman*. That is because the process of self-discovery, of the self's, or the *atman*'s effort to raise to an awareness of the inner Self, or *Atman* within, is not an effort to attain something totally unique, or distinct.[39] Instead, the

---

[35] *The Upanishads*, 310-311.

[36] Ibid., 310-311; G. W. Leibniz, *Philosophical Essays,* trans. Daniel Garber, and Roger Ariew (Indianapolis: Hackett Publishing Co., 1989), 207.

[37] *The Dhammapada,* 154.

[38] *The Upanishads*, 58, 165.

[39] *The Dhammapada,* 154; *The Upanishads*, 165.

self's, or *atman*'s inner journey toward the Self, or the *Atman* is something individuals already harbor and although it is for us to come to realize, or recognize, it is still something that connects us all to the same common origin point that is the Self, or the *Atman*.

Accordingly, readers may begin to see how it is that the Buddha commits a straw man fallacy against the Hindu concept of the self, or the *atman* as well as the Self, or the *Atman*. For, the Buddha by understanding and claiming that those driven by purpose, to attain absolute reality, mistakenly takes this to be a desire for grasping to something that is separate, other, or estranged from oneself, when Hinduism instead expounds that absolute reality is that which one finds from within.[40] Lastly, readers should now consider other reasons as to why it is that the Buddha's teaching of the *anatman* is not as justifiable as it may appear to be to some, and instead, let us assert the concept of the self, or the *atman*, and the Self, or the *Atman* all the more.

Furthermore, readers may claim that the very concept of the *anatman*, or one who is without a self, or *atman*, defies the very concept of the enlightened, or The Awakened One, or of a Buddha himself/herself. In other words, if one acknowledges one who achieved Buddhahood as being a Buddha, then how can anyone ever establish the quality of Buddhahood characterized by a permanent state of bliss, clarity of mind, and beyond all conditioning?[41] That is, if people are truly *anatman*, or without a self, or absent of an *atman*, then why should they strive for Buddhahood if that too is merely a label and not a descriptive feature of permanent selves, or *atman*(s)?

Consequently, if readers embrace the notion of the *anatman*, as related to Buddhahood, then we are illogically asserting that one who is without a self, or an *atman* is now one who mastered himself/herself in such a way that that individual is free from all conditioning that person underwent, as a self, or an *atman*. In other words, the problem of the *anatman* and Buddhahood is how can it be that one who is without a self, or an *atman* can build such a self, or an *atman* that leads to an everlasting state of being that is Buddhahood. Finally, to further explicate this matter, readers must consider if it can be the case that an input that shares no likeness to its output, or the *anatman* as connected to Buddhahood, can ever produce such an output so different from itself.

Next, another problem that arises from the idea of the *anatman* when compared to the idea of the self, or the *atman*, as well as the Self, or the *Atman* is why should it be that we are to seek *nirvana*, if there is truly no self, in the eyes of the Buddha.[42] In other words, if one seeks a state of liberation from all conditioning that is permanent and unaltering, or *nirvana* itself, should it

---

[40] Ibid.

[41] *The Dhammapada,* 169-171.

[42] Ibid., 153.

not be so that everyone is already in a state of *nirvana*, if the conditioning that all understand as their own is instead just another mere illusion?[43]

As such, why should one attempt to achieve *nirvana* if there is no self to attain or experience such a state of undying spiritual emancipation? At least in Hinduism, although everyone ought to uproot the self, or the *atman* to reach, or disclose *moksha* and the Self, or the *Atman*, there is still a self, or an *atman* performing such a task that promises with it spiritual liberation for those who adhere to the *dharma*, or the spiritual laws, in this case of Hinduism, in a perfectly pure way.[44] Finally, this absence of a self in Buddhism, or the *anatman* only leads to questions that serve to be not completely resolvable which appears as bypassed in Hinduism by the affection, or embrace of the self, or the *atman* and the Self, or the *Atman*.

V. Conclusion

As stated in this present essay, readers encountered the Hindu and especially the Upanishadic understanding of the self, or the *atman* as well as the Self, or the *Atman*. Furthermore, readers also read an explanation of the Buddhists concept of the *anatman* and its appearance, as well as aporias as found in the *Dhammapada*. Moreover, this article concluded with challenges, to help debase the idea of an *anatman*, and therefore to assert and secure that the Hindu concept of the self, or the *atman* and the Self, or the *Atman* possesses more validity and a greater degree of justifiability than its Buddhists opposite. Finally, it is the sincere intention of this present writer that this article helps to fuel the ongoing debate regarding the notions of a core identity, inherent to each individual as opposed to a lack thereof, in both Eastern and Western philosophical circles alike.

References

Leibniz, Gottfried Wilhelm. *Philosophical Essays*. Translated by Daniel Garber, and Roger Ariew. Indianapolis: Hackett Publishing Co., 1989.

The Dhammapada. Translated by Eknath Easwaran. Tomales, CA: Nilgiri Press, 2008.

The Upanishads. Translated by Eknath Easwaran. Tomales, CA: Nilgiri Press, 2007.

---

[43] Ibid., 130-132, 153.

[44] *The Upanishads*, 241-243.

Reconsidering Burke's and Arendt's Theories on "The Rights of Man": A Surprising Plot Twist?

Ioannes P. Chountis

Universität Tübingen, Germany
E-mail address: giannischountis@gmail.com
ORCID iD: http://orcid.org/0000-0003-2910-2425

Abstract

Edmund Burke can perhaps be considered as the father of modern conservatism. Hannah Arendt was a very eclectic thinker who embraced ideas from the traditions of liberalism and republicanism. They both commented on the issue of the "Rights of Man" and rejected their abstract and metaphysical nature. And, it was Arendt who saw a 'certain pragmatism' in Burke's ideas. Is this coincidence of opinion a surprising plot twist? An unintentional 'alliance' against the naturalness of the "Rights of Man?" This paper first discusses the real relationship between Burke's and Arendt's theories on human rights. In the first part, the ideas of the two thinkers are presented and examined. In the second part, the main convergences and divergences are identified. Through a careful reading of the Burkean and Arendtian corpus, it is shown that Arendt agreed with Burke that human rights cannot be abstract or metaphysical. On the other hand, Arendt, being autonomous in her critique, argued for one universal and inalienable right, that is 'the right to have rights,' i.e. the right to belong to political community. In overall, the analysis endeavors to provide an answer to the question as to what degree did Arendt endorse Burke's theories on the "Rights of Man."

Keywords: *"Rights of Man;" Hannah Arendt; Edmund Burke; Conservatism; Republicanism*

I. Introduction

In recent years, human rights have been at the center of political and philosophical debates. Faith in liberal democracy and representative institutions seems to have been shaken by violations of fundamental rights, as happens with the case of refugees, minority groups and others. Such observation can be attested by the emergence of a vocal protest movement in the United States against racial discrimination and violence. Of course, the discourse on human rights is much older. Already since their enactment,

the Bill of Rights and the Declaration of the Rights of Man and of the Citizen provoked different interpretations regarding their nature and scope of application.

In the field of political philosophy and intellectual history, Edmund Burke and Hannah Arendt are two thinkers who belong to divergent philosophical currents and lived during different historical periods. Burke's philosophy correlates with the broader tradition of – British – conservatism while Arendt's multifaceted and complex philosophy can be linked up to a certain degree to republicanism. *Prima facie* establishing a common ground between these theories appears to be almost impossible. However, a careful consideration of their works may offer certain surprises; because both Burke and Arendt criticized the doctrine of the "Rights of Man," which was formulated according to the eighteenth-century notions of natural law. It should be noted that although they both rejected the abstract and metaphysical concept of the "Rights of Man," and some philosophical debts to Burkean thought are traceable in the Arendtian corpus, the German thinker was creatively autonomous in her critique and made sure to adapt it to the post-war context of the twentieth century.[1]

Indeed, Burke and Arendt formulated their theories against different historical backgrounds. On the one hand, Burke attacked the "Rights of Man" as formulated in the declarations of the French Revolution and espoused by radical English thinkers like Dr Richard Price and Thomas Paine.[2] On the other hand, Arendt wrote her critique of natural rights after the Universal Declaration of Human Rights had been adopted by the General Assembly of the newly founded United Nations in December 1948.[3] As of such, Burke presented his arguments during the culmination of what became known as the 'Age of Revolutions,' while Arendt drafted her objections against the backdrop of the decline of the nation-state after the end of World War Two. As it will be shown below, Arendt's critique of natural human rights followed, at least up to a certain degree, the arguments of Burke and Jeremy Bentham.[4] After all, it was the utilitarian philosopher who maintained that "rights" exist only "due to legislation."[5] At the same time, both Burke and Bentham agreed

---

[1] Arendt makes mention of Burke *inter alia* in her works *On the Revolution* and *The Origins of Totalitarianism*.

[2] R. R. Fennessy articulated this debate in *Burke, Paine, and the Rights of Men. A Difference of Opinion* (Hague: Martinus Nijhoff, 1963).

[3] Christoph Menke, Birgit Kaiser, and Kathrin Thiele, "The 'Aporias of Human Rights' and the 'One Human Right:' Regarding the Coherence of Hannah Arendt's Argument," *Social Research* 74, no. 3 (2007): 739.

[4] Menke, Kaiser, and Thiele, 742.

[5] Ibid.

that "laws are always passed for particular nation-states."[6] Arendt, following the same line of thought, argued in favour of the connection between human rights and laws on the one hand and the tie of legislation to a certain 'locale' on the other.[7] However, as it will be noted in the following sections, Arendt soon presented her alternative and potential solution to the problem posed by the 'English' critique of the "Rights of Man."

In the following sections, the fundamental arguments of Burke and Arendt against the theory of the "Rights of Man" are first analyzed, as recorded in their cardinal works, the *Reflections on the Revolution in France* and *The Origins of Totalitarianism* respectively.[8] Then, the convergences and divergences of the two theories are summarized, so as to provide some conclusions relating the position of rights in the philosophy of the two thinkers.

II. Edmund Burke on human rights

'Human rights' are intricately linked to *unhistorical* human nature. Therefore, they are natural rights "which do not exist only in law, but also independently, through binding precepts of morality that do not depend upon a legal code for their validity."[9] They were formulated as such in the Declaration of the Rights of Man and of the Citizen, issued in 1789 by the French National Assembly, which thus placed "natural rights at the center of the new system of government."[10]

For Burke though, all rights are 'derived from historical development and recognized through positive laws and customary practice;'[11] in other words, they originate from a particular tradition. Thus, it is evident that meant only within a certain political context, rights have their roots in the historically

---

[6] Ibid.

[7] Ibid.

[8] Excerpts from other texts are noted where it is deemed proper.

[9] Roger Scruton, *The Palgrave Macmillan Dictionary of Political Thought* (New York: Palgrave Macmillan, 2007), 602-603.

[10] John Morrow, *A History of Political Thought. A Thematic Introduction* (New York: New York University Press, 1998), 220. In her critique, Arendt not only mentioned the French but also the American Revolution to which Burke does not refer. According to Arendt, together with the French, the American Revolution established Human Rights as the basis of 'civilized societies;' Hannah Arendt, *The Origins of Totalitarianism* (New York: Cleveland, World Pub. Co., 1962), 293. Furthermore, the American revolutionaries primarily focused their demands on "life, liberty, and the pursuit of bliss;" Ibid., 295, 298.

[11] Morrow, 221. In his early work, *A Vindication of Natural Society*, Burke had already rejected the notion of natural society and the self-existent individuality of Man outside political society, because according to his analysis, human needs are met only within a specific political and social structure, see Scruton, *Dictionary,* 70.

shaped human society, such as the English nation in Burke's case.[12] For the Anglo-Irish thinker, this also presupposes that they are not abstractly worded and based on speculative *apriorisms*. Besides, Burke displayed in most cases a general suspicion, if not rejection, of any theory claiming universal application.[13] Furthermore, he argued that the theories of natural law and natural rights, which had been invoked by the enemies of arbitrary power in the seventeenth and early eighteenth-century, were the forerunners of the radical-revolutionary version of his contemporary France.[14]

Additionally, Burke maintained that laws sufficed to describe societal condition via a set of orders and prohibitions. He accused radicals like Thomas Paine of violent interference with the structure of society by the imposition of abstract theories on historical reality. For him, the result of these actions would be the eventual overthrow of all the pillars of political society.[15]

However, it should be underscored that Burke quite interestingly advocated natural law in line with the Aristotelian and Thomistic tradition, which was not based on transcendent principles but was rooted in political society. At first glance, this position might seem to constitute a *contradictio* to the offensive against the "Rights of Man." Nevertheless, one must bear in mind that the classical tradition of natural law did not focus on the right *per se* but on the linked duty, and that instead of Reason, it established Prudence as the foundation of rights.[16] It is in line with this tradition that Burke referred to what he labeled '*real* human rights' which had their roots in political society.[17] Besides, Burke's notion was not limited to rights protected by written law but extended to other privileges which in his time, at least, were not protected by specific rules but stemmed from tradition. To address this apparent antinomy, R. R. Fennessy pointed out that Burke had made a methodological distinction between rights as an individual privilege and rights in relation to others, endorsing the latter because only these exist within political societies.[18] This is the essence of Burke's real natural rights, and this

---

[12] Cf. Edmund Burke, "English Rights," in *Reflections on the Revolution in France*, ed. John Greville Agard Pocock (Indianapolis and Cambridge: Hackett Publishing Company, 1987), 28.

[13] Morrow, 372.

[14] Ibid., 224.

[15] Cf. his early satire, *A Vindication of Natural Society,* where Burke argued that Lord Bolingbroke's deism would, eventually, not only lead to the overthrow of revealed religion and the established church but of political society, too.

[16] On the essence of Prudence, Burke followed the Aristotelian tradition, according to which Prudence refers to the empirically acquired practical wisdom. On the other hand, the Platonic theory of Prudence is related to the knowledge of Ideas; Burke, 28, 30.

[17] Fennessy, 138-139.

[18] Burke shared Aristotle's view that virtue is always practiced in relation to someone else as

can further elucidate the reasons behind his fierce rejection of the "Rights of Man," condemning them as transcendent and unrelated to the state and society. Finally, Burke argued for the complexity of human nature and the various interpersonal relationships, that develop between the members of the body politic, and rejected the simplistic and vague wording of the French Declaration. In fact, he prophesied that such a proclamation could be used as a tool of despotism.[19]

In the following excerpt from his *Reflections*, Burke's perception of rights is stated quite clearly:

> Men cannot enjoy the rights of an uncivil and of a civil state together. [...] By having a right to everything they want everything. Government is a contrivance of human wisdom to provide for human wants. Men have a right that these wants should be provided for by this wisdom. [...] But as the liberties and the restrictions vary with times and circumstances and admit to infinite modifications, they cannot be settled upon any abstract rule; and nothing is so foolish as to discuss them upon that principle.[20]

In this context, he turned against Dr Price, Priestly and the French *philosophes*, accusing them of presenting metaphysical theories, which were not based neither on history nor on tradition. Of course, Burke never formulated his own comprehensive theory of human rights. On the one hand, this is due to the fact that he rejected the very existence of natural human rights and, on the other, it is owed to the structural and stylistic peculiarities of his works; the publication of his parliamentary speeches in pamphlets and the letter form of many of his other works – like the *Reflections* – inevitably resulted into a fragmentation of his thinking. As a result, Burke did not produce any treatise of political philosophy, that is a clear and 'watertight' framework of onto-political principles. Instead, he chose to compose texts with observations and thoughts on current issues, as is the case with the French Revolution.[21] Roger Scruton rightly noted that the complex and often literary style of Burke's thought had not particularly aided the Anglo-Irish thinker at a time when closed philosophical systems had been still prevalent.[22] In addition, this

---

opposed to Plato, who envisioned the inner harmony of the three parts of the soul; Ibid., 139.

[19] Burke, 101-102.

[20] Ibid., 52-53.

[21] Scruton, *Dictionary*, 69.

[22] Roger Scruton, *A Short History of Modern Philosophy: From Descartes to Wittgenstein*

topical nature of his writings creates additional problems of interpretation; often Burke, who had studied law, did not hesitate to deliberately distort the facts in order to provide for the validity of his allegations.

Overall, according to Richard Bourke's comprehensive analysis, Burke identified two cardinal dangers in the program of the French Revolution deriving from these 'abstract' and 'transcendental' rights. Following – utilitarian – jurisprudence, Burke underlined that the 'French rights' would eventually result in "self-government as a means of determining the shape of existing civil societies" and that they could "challenge the distribution of wealth in established societies."[23] Consequently, for Burke, these primordial rights were against the very essence of classical natural law.[24]

III. The Arendtian theory of rights

Arendt's theory of human rights is primarily based on an early article entitled "The 'Rights of Man,' What are They?" and the ninth chapter of the second part of the *Origins of Totalitarianism* under the title "The Decline of the Nation-State and the End of the Rights of Man."[25]

Arendt's primary thesis is linked to the belief that there is no practicality in realizing human rights and that the very essence of these natural rights is very different from what eighteenth-century thinkers believed.[26] Her theory was developed both as a criticism and an alternative to the notions of modern natural law and/or liberalism. As a result, it appears that Arendt took distance from eighteenth-century French revolutionaries and their conception of natural rights.

For Arendt there is only one fundamental right, that is the right to "belong to the political community," from which all other human rights derive.[27] In the post-war world, Arendt produced this notion in which a person, like the refugee, is deprived of a "place on the planet that makes their views important and their actions effective."[28] This led Arendt to believe that rights exist only within certain state entities and are protected by enacted laws.

(London: Routledge, 2001), 223.

[23] Richard Bourke, *Empire and Revolution. The Political Life of Edmund Burke* (Princeton, NJ: Princeton University Press, 2015), 574.

[24] Ibid., 574.

[25] Hannah Arendt, "The 'Rights of Man:' What are They?" *Modern Review* 3 (1949): 24-37; Arendt, *The Origins*, 268-302.

[26] Menke, Kaiser, Thiele, 740.

[27] Arendt, *Rights*, 37; Serena Parekh, *Hannah Arendt and the Challenge of Modernity. A Phenomenology of Human Rights* (London and New York: Routledge, 2008), 11.

[28] Arendt, *Rights*, 29. The loss of state is tantamount to the loss of all rights, Parekh, 18.

Endorsing Burke's position, Arendt agreed that human rights, as articulated in the eighteenth-century, were, indeed, an 'abstraction' and she argued in favor of the 'irreversible legacy' of rights inherited from one generation to the next, as Burke's 'rights of the English;' these are rights channeling from 'nations.'[29] Moreover, she believed that the "Rights of Man" were of the "order of an 'ought,' to which no 'can' correspond[ed]."[30] This highlighted their metaphysical and utopian nature in Arendt's perception.

The waves of refugees and 'stateless' people in the post-war world were important events which compelled Arendt to formulate her conception of human rights.[31] In the German thinker's argument, stateless people were those who had no recognized legal or political status.[32] Of course, the emergence of this phenomenon was a very significant global problem which went far beyond totalitarian regimes.[33] Besides, it was a phenomenon unknown to eighteenth-century philosophers and thinkers. Thus, it became clear in her mind that the loss of 'national rights' would lead to the consequent deprivation of all 'human rights.'[34] At that point the "world found nothing sacred in this abstract nakedness of being human," she argued.[35] Therefore, it can be said that for Arendt, having human rights means participating in a certain civic nation, which through written law protects the rights of its citizens.[36] It is in

---

[29] Arendt, *Rights*, 31; Parekh, 24; Bridget Cotter, "Hannah Arendt and 'The Right to Have Rights,'" in *Hannah Arendt and International Relations*, eds. Anthony F. Lang Jr, and John Williams, 95-112 (New York: Palgrave Macmillan, 2005), 96-97.

[30] Menke, Kaiser, Thiele, 742.

[31] Arendt, *Rights*, 31. Through her experience of being a stateless Jew and 'foreign enemy,' she realized how the twentieth-century bureaucracy was characterized by an absolute irrationality; Richard J. Bernstein, *Hannah Arendt and the Jewish Question* (Cambridge, MA: MIT Press, 1996), 75. Characteristically, she writes that Kafka and not Weber understood correctly bureaucracy's nature, Hannah Arendt, *Essays in Understanding, 1930- 1954*, ed. Jerome Kohn (New York: Harcourt, Brace & Co., 1994), 73-74. These experiences were mainly recorded in Hannah Arendt, "We Refugees," *Menorah Journal* 31, no. 1 (1943): 69-77.

[32] Hannah Arendt, *The Jew as Pariah*, ed. Ron H. Feldman (New York: Grove Press, 1978), 65; Bernstein, 77. Arendt drew her arguments from her critique of the legacy of the Enlightenment and classical liberalism, which both had underlined the importance of 'inalienable rights.'

[33] Arendt, *Origins*, 459.

[34] Arendt, *Rights*, 31.

[35] Ibid., 31.

[36] Bernstein rightly acknowledges that in the birth of modern nation-states an internal opposition prevailed between the two synthetic entities, the nation and the state. During the nineteenth - century, the internal contradiction of the Declaration of Human Rights between inalienable rights and the need to protect them by a state entity was alleviated. However, the undermining of the nation-state at the beginning of the twentieth - century together with the rise of imperialism and the First World War led to the disintegration of this guarantor of 'inalienable rights.' Then the invocation of Human Rights became politically weak and inapplicable, Bernstein, 79; Arendt, *Origins*, 291-292, 293.

this respect that Arendt's notion correlated with Burke's human rights with their inalienable character guarantee only the status of savages to people.[37] Nevertheless, as will be indicated next, Arendt accepted one inalienable right, i.e. the right to have rights, which eventually constitutes her theory significantly different from Burke's.

She further argued that even though people are not born equal, they become equal when they participate in political society, which in turn guarantees them equal rights.[38] It has been righty observed that Arendt "was deeply skeptical of all those tendencies in modern life that foster a false sense of social equality and homogeneity."[39] Sharing republicanism's view to a certain degree, Arendt offered the idea that there are no independent rights but only in relation to others.[40] In this, she agreed with Burke who had argued in favor of rights in relation to others and not as an independent privilege. What is more, she emphasized that although rights arise from political community, there must be either an institution of a federation of states or international law to protect and secure their implementation. Thus, *polity* becomes the cornerstone of the Arendtian conception of politics. Moreover, Arendt noted that the various declarations of human rights demanded equal rights for "something essentially nonequal: human beings as natural beings. There are only equal rights for political members, which are thus not human rights."[41] It is this loss of *polity* that expels one from humanity.[42] This idea of polity is interlinked with Arendt's notion of dignity, which meant being a member of a political community, i.e. possessing that right to have rights. Put in Aristotelian terms, dignity for Arendt is the ability of human beings to speak and accordingly be political animals. Thus, dignity and polity are not 'natural properties' but rather a 'politico-linguistic experience.'[43]

---

[37] Arendt, *Rights*, 32.

[38] Ibid., 33; Bernstein, 86; Arendt, *Origins*, 301.

[39] Bernstein, 86.

[40] Arendt, *Rights*, 34. As for Arendt's dialectical and rather eclectic relationship with republicanism and liberalism, it must be underscored that to this day it remains unclassified, because she agreed to some extent with both traditions. In addition to the aforementioned republican view, she advocated the liberal notion of a private sphere and negative freedom and in contrary to Rousseau and others she did not believe that it is the true nature of the individual to be a citizen above anything else. An extreme example of her consistency to liberal positions can be found in her article on mixed schools, in which she advocated for the parental right to choose whether to send their children to these schools or not, "Reflections on Little Rock," in *Responsibility and Judgment*, ed. Jerome Kohn, 192-213 (New York: Schocken Books, 2003).

[41] Menke, Kaiser, Thiele, 746.

[42] Ibid., 752.

[43] Ibid., 753.

In the controversy between Burke and Paine over the "Rights of Man," Arendt, in a perhaps surprising turn, sided with the former, arguing that "there is no period in history when the Declaration of Human Rights could find a response."[44] However, as already stated above, she was quick to remind that a universal and inalienable right did, in fact, exist. The right to belong to political community, although incomprehensible in Burke's time, became imperative in the post-war world.[45] Besides, the Declaration, in contrast to its American model, the United States Bill of Rights – which was endorsed by Arendt –, sough only to express 'positive, primary' rights in 'opposition to political status,' and replaced history with nature as a result.[46] This contributed to pre-political rights to livelihood becoming the cornerstone of the new revolutionary regime in France. At the same time, in Arendt's mind, this was also the inherent weakness of the new government.

IV. Convergences and differences

These are the most important points of the Burkean and Arendtian theories of the "Rights of Man." A superficial reading may lead to the erroneous conclusion that the two theories are completely identical. However, this is not true because Arendt never fully adopted the Burkean position, although she vindicated crucial aspects of his reading of natural rights. A cogent comparison of Arendt and Burke on the topic of the existence of natural rights can, perhaps, launch a broader discussion on the naturalness of rights It can be argued, that *in principio* both Burke and Arendt agreed that rights arise from political society and are not the product of nature. However, whereas Burke believed that they were the result of a particular community's tradition and history, Arendt maintained that their implementation can be enforced only by supranational institutions and general principles, like polity and dignity.[47] This is the first major divergence between the two theories, especially regarding Arendt's mention of the international law's critical role. There is another major difference between the two thinkers: Burke and Arendt interpreted differently the very concept of political society; Burke perceived civil society as a state edifice characterized by homogeneous ethnic features while Arendt described a community whose culture is purely political. This is crucial for their understanding not only of political community but for their notion of *polity* as well. This, eventually, extends into their cardinal point of difference regarding the existence of at least one universal or natural human

---

[44] Hannah Arendt, *On Revolution* (New York: Viking Press, 1963), 60.

[45] Ibid., 61.

[46] Ibid., 146.

[47] Parekh, 22-24.

right; Burke denied its existence while Arendt accepted that such a right could exist.

Furthermore, a critical feature in differentiating Burke from Arendt is the subject of universality. It is true that they both criticized the abstract and metaphysical nature of the "Rights of Man." But Arendt acknowledged one universal right without reducing it to the realm of the metaphysical. Thus, indeed the 'right to have rights' reveals a certain universality.[48] Bernstein noted that "there is clearly a universal thrust in her claim that every single individual has (or ought to have) the 'right to have rights.' But this right becomes concrete only in the life of a particular community."[49] To this point, of course, Burke would disagree arguing against the existence of any natural human right outside the nation-state.

If one were to analyze Arendt's phrase 'the right to have rights,' they would discover that the first part refers to the moral imperative of belonging to a political community while the second part refers to the right to equal participation in the public sphere, which in turn presupposes independence from arbitrary interference in the private sphere.[50] A careful examination of the Arendtian text would also highlight the lack of a particular subject – in other words, who is the holder of this right? In order to understand Arendt's point, one must return to her ontology. Because, based on the concept of humanity, the ontological foundation of her view is the *principium* of natality, which claims universal application.[51] In Arendt's thought, humanity approaches the notion of international law which should govern all human communities and affairs.[52] Put in such terms, there is a certain paradox in the Arendtian rights. On the one hand, they are established on the international principle of humanity which arises from earthly human condition and seeks universal application, while, on the other, they are concrete and not of an abstract nature. The universality of 'the right to have rights' is the result of the Arendtian ontology, as analyzed in *The Life of the Mind*, where the law of the Earth corresponds to plurality and is linked to the community and not the individual.[53] Whatever might be the true nature of Arendt's rights, it is

---

[48] Arendt, *Origins*, 296-297; Peg Birmingham, *Hannah Arendt and Human Rights. The Predicament of Common Responsibilit*y (Bloomington and Indianapolis: Indiana University Press, 2006), 1; Parekh, 29.

[49] Bernstein, 84.

[50] Seyla Benhabib, *The Reluctant Modernism of Hannah Arendt* (New York: Rowman & Littlefield, 2003), 56-57.

[51] Birmingham, 4, 6, 39.

[52] Here Birmingham identifies the influence of Augustine; cf. Ibid., 36.

[53] Hannah Arendt, *The Life of the Mind* (New York: Harcourt, Brace, Jovanovich, 1978), 20; Bernstein, 82. For Arendt, humanization begins with the integration of the individual into

evident that their very mention of universality renders them of a very different character from Burke's 'rights of the English.'

As for these 'rights of the English,' Arendt accepted them in principle, although she noted that this Burkean view almost reached the notion of 'a race of blue-blooded aristocrats.'[54] The basis of Burke's conception was for Arendt the ancestor of the nineteenth and twentieth-centuries racial theories. Of course, it must be said that Arendt did not make justice to Burke's ideas since they were by no means characterized by any hint of racism. On the contrary, Burke not only supported the rights of Catholics but also devoted almost two decades of his life vindicating Indians and prosecuting Warren Hastings. Finally, he belonged to the more moderate political connection of his time, the (Rockingham) Whigs.

Lastly, the issue of equality divides Burke's and Arendt's theories. The German thinker disagreed with Burke on the equality of citizens. Starting from the common point that humans are not born equal, the two thinkers followed different paths. As noted above, Burke never believed nor argued for the natural and legal equality of all people. Arendt on her part held the opinion that it is participation in political community that rendered people equal to each other, even if they had not been born equal.

Overall, regarding the similarities and divergences between the two theories, it can be argued that, perhaps, the most striking common feature is their rejection of the abstract, metaphysical, and inalienable "Rights of Man." This is a rather interesting plot twist and coincidence of opinion between two very different thinkers. Where one might expect Arendt to be in agreement with Paine she/he finds her supporting Burke's opinions. But this is, also, perhaps the only major point of genuine philosophical convergence.

As noted above, Arendt acknowledged the existence of one human right, that is the right to have rights, or, in other words, the right to belong to a political community. Consequently, this fundamental human right became a cornerstone for her notion of 'dignity.' Burke on his part never accepted the existence of such a fundamental 'human right.'[55] Furthermore, Arendt's only *real* human right was developed in an attempt to override this very 'English' critique of the "Rights of Man."[56] Burke, Bentham, and other critics of the

---

a community, where one actively participates. On the contrary, world-alienation, which characterized refugees, is for Arendt the 'hallmark of the modern age;' Hannah Arendt, *The Human Condition* (Chicago: University of Chicago Press, 1958), 254. Again, here the loss of participation in the community is tantamount to the loss of fundamental features of our humanity.

[54] Birmingham, 46.

[55] For more on Arendt's concept of dignity see Menke, Kaiser, Thiele.

[56] Menke, Kaiser, Thiele, 750.

French Declaration, argued, that all rights were based on laws passed in specific nation-states. As of such, no human right could exist outside these said political communities. But with Arendt's notion that "each individual human being, as members of an eventually universal and (quasi-) political community: the 'political entity' of 'mankind,'" this problem was solved.[57] This is where Arendt departed from the earlier English theory and formulated her own alternative of the 'right to have rights.'

V. Conclusion

In the second and third subjection a brief overview of Burke's and Arendt's theories against the "Rights of Man" was presented. Then, in the third part an attempt was made to identify those points of the Burkean argument from which Arendt took her insight against the "Rights of Man" but also the point of her departure from the 'English critique' of the French Declaration.

In conclusion, it should be underlined that Isaac is right to remind that Arendt recognized a certain pragmatism in Burke's theory. But that does not mean that she fully endorsed the philosophical basis of his claims.[58] Burke was a conservative and Arendt was a follower of republicanism and/ or liberalism. As of such the real relationship of their respective approaches is very complicated and by no means unambiguous.[59] What might possibly be closer to the truth is that Arendt, being an eclectic thinker, chose to support part of her arguments against the "Rights of Man" on Burke's opinions but at no point did she decide to embrace the whole of his philosophy. In fact, she quickly became autonomous in her critique and even proposed a solution to the problem of natural rights.

Burke proposed history and tradition, the 'rights of the English' as the only viable alternative to the French natural rights. Arendt agreed with him that the essence of the "Rights of Man" was rather abstract and metaphysical. Being of such nature they could not protect the victims of twentieth-century totalitarianism. But for her the historical 'rights of the English' were also inadequate to face the novel challenges of the post-war world. What Arendt counter-proposed was a unique, real human right, that is the 'right to have rights' or 'to belong to a political community.' Her alternative sought to alleviate both the problems of the "Rights of Man" metaphysical nature and of the respective 'English critique.' Finally, this partial coincidence of opinion

---

[57] Ibid., 751.

[58] Jeffrey C. Isaac, "Hannah Arendt on Human Rights and the Limits of Exposure, or Why Noam Chomsky Is Wrong about the Meaning of Kosovo," *Social Research: An International Quarterly* 69, no. 2 (2002), 511.

[59] Birmingham, 45.

between Burke and Arendt vindicates the subtle and complex nature of their thinking and the fact they both were quite creative in formulating their ideas.

References

Arendt, Hannah. "We Refugees." *Menorah Journal* 31, no. 1 (1943): 69-77.

Arendt, Hannah. "The 'Rights of Man:' What are They?" *Modern Review* 3 (1949): 24-37.

Arendt, Hannah. *The Human Condition*. Chicago: University of Chicago Press, 1958.

Arendt, Hannah. *The Origins of Totalitarianism*. New York: Cleveland, World Pub. Co., 1962.

Arendt, Hannah. *On Revolution*. New York: Viking Press, 1963.

Arendt, Hannah. *The Life of the Mind*. New York: Harcourt, Brace, Jovanovich, 1978.

Arendt, Hannah. *The Jew as Pariah*. Edited by Ron H. Feldman. New York: Grove Press, 1978.

Arendt, Hannah. *Essays in Understanding, 1930 – 1954*. Edited by Jerome Kohn. New York: Harcourt, Brace & Co., 1994.

Arendt, Hannah. "Reflections on Little Rock." In *Responsibility and Judgement*, edited by Jerome Kohn, 192-213. New York: Schocken Books, 2003.

Benhabib, Seyla. *The Reluctant Modernism of Hannah Arendt*. New York: Rowman & Littlefield, 2003.

Birmingham, Peg. *Hannah Arendt and Human Rights. The Predicament of Common Responsibility*. Bloomington and Indianapolis: Indiana University Press, 2006.

Bernstein, Richard J. *Hannah Arendt and the Jewish Question*. Cambridge, MA: MIT Press, 1996.

Burke, Edmund. *Reflections on the Revolution in France*. Edited by John Greville Agard Pocock. Indianapolis and Cambridge: Hackett Publishing Company, 1987.

Bourke, Richard. *Empire and Revolution. The Political Life of Edmund Burke*. Princeton, NJ: Princeton University Press, 2015.

Cotter, Bridget. "Hannah Arendt and 'The Right to Have Rights.'" In *Hannah Arendt and International Relations*, edited by Anthony F. Lang Jr., and John Williams, 95-112. New York: Palgrave Macmillan, 2005.

Fennessy, R.R. *Burke, Paine and the Rights of Men. A Difference of Opinion*. Hague: Martinus Nijhoff, 1963.

Isaac, Jeffrey C. "Hannah Arendt on Human Rights and the Limits of Exposure, or Why Noam Chomsky Is Wrong about the Meaning of Kosovo." *Social Research: An International Quarterly* 69, no. 2 (2002): 505-537.

Menke, Christoph, Birgit Kaiser, and Kathrin Thiele. "The 'Aporias of Human Rights' and the 'One Human Right:' Regarding the Coherence of Hannah Arendt's Argument." *Social Research* 74, no. 3 (2007): 739-762.

Morrow, John. *A History of Political Thought. A Thematic Introduction*. New York: New York University Press, 1998.

Parekh, Serena. *Hannah Arendt and the Challenge of Modernity. A Phenomenology of Human Rights*. London and New York: Routledge, 2008.

Scruton, Roger. *A Short History of Modern Philosophy: From Descartes to Wittgenstein*. London: Routledge, 2001.

Scruton, Roger. *The Palgrave Macmillan Dictionary of Political Thought*. New York: Palgrave Macmillan, 2007.

COVID-19 and Other Prevalent Diseases in Africa: A Pragmatic Approach

Cyril Emeka Ejike

Nnamdi Azikiwe University, Nigeria
E-mail address: cyril.ejike@gmail.com
ORCID iD: https://orcid.org/0000-0002-4728-195X

Abstract

The aim of this paper is to propose that the development and legitimization of African knowledge and validation systems on a pragmatic basis is an efficient and effective means of responding to a myriad of health problems plaguing Africans, particularly the COVID-19 pandemic. Whenever there is a novel disease outbreak, the norm is to wait for the development of scientifically proven vaccines for its treatment. However, the scientific validation of drugs is a rigorous and lengthy process, thereby inappropriate for dealing with health emergencies like the COVID-19 outbreak. The alarming rapidity with which the novel COVID-19 pandemic rages globally and decimates humanity has brought to the fore the need for Africa to look inwards in search of viable and efficient alternative approaches to the pandemic. In this paper, I examine pragmatism as a theoretical framework and relate it to proposed African epistemic and validation frameworks with a particular reference to homegrown orthodox and alternative/complementary medicines. I argue that the validation and approval of any knowledge claim based on pragmatism is a more expeditious mode of attending to COVID-19 and other prevalent diseases in Africa. The application of knowledge that brings practical success in dealing with health challenges in Africa without necessarily following rigid and lengthy scientific validation procedures will go a long way toward improving human conditions and well-being. I conclude that pragmatic considerations should ultimately inform local approval to homegrown African medicines for use in Africa.

Keywords: *Africa; COVID-19; diseases; pandemics; Pragmatism*

I. Introduction

The outbreak of the coronavirus disease (COVID-19) from the city of Wuhan, China in December 2019 has spread to other countries by leaps and bounds, resulting in a massive loss of lives and affecting individuals, families, communities, and countries economically, socially, politically, and psychologically. The pandemic outbreak and an alarming death toll of thousands recorded globally have disproved the invulnerability of the most developed nations, superpower nations in particular, as well as exposed the susceptibility of the developing nations, especially African ones. Currently, the priority of nations is to defend their own interests, with Africa being left to its fate.[1]

Given Africa's fragile economies and poor social protection systems, it is not absurd to think that the African continent will be the arena the virulent pandemic will play its final and enduring havoc with humanity. It is standard practice to require scientific evidence to corroborate any claim for the discovery of an effective vaccine or drug for the treatment of extant and novel diseases. However, the scientific process of the verification and validation of new vaccines/drugs for medical treatments is usually strict and lengthy.[2] Vaccines have to undergo extensive clinical trials. They normally require at least two years of animal testing to ascertain their efficacy and safety before they are deemed safe enough for clinical trials in healthy human volunteers to determine their effectiveness in treating people.[3]

Such a rigorous process of establishing truth claim does not fit the bill in the case of the outbreak of virulent diseases like the COVID-19 pandemic. One may pose several questions. Can African countries afford to wait for the development of scientifically proven COVID-19 vaccines amidst the surging scourge and death toll of the pandemic in Africa? Is it not suicidal to wait for the rigorous and lengthy scientific processes of the verification and validation of COVID-19 vaccines, while the pandemic is fast scourging and decimating Africans? Given that the chief purpose of establishing a civil society is to protect lives and property and improve human welfare, is it not expedient to adopt any viable and efficient approach to the COVID-19 pandemic to save lives?

---

[1] Charles C. Soludo, "Can Africa Afford COVID-19 Lockdowns?" *Proshare*, April 24, 2020, https://www.proshareng.com/news/NIGERIA%20ECONOMY/Can-Africa-Afford-COVID-19-Lockdowns----Chkwuma-Soludo/50636#.

[2] Cyril E. Ejike, "COVID-19 and African Traditional Medicines," in *COVID-19 and Afrocentric Perspectives: Health and Economic Implications*, eds. Ikechukwu A. Kanu, Chiugo C. Kanu, and Ejikemeuwa J. O. Ndubisi (Maryland: The Association for the Promotion of African Studies, 2021), 2.

[3] Ejike, 2.

Thus, looking inward in search of solutions to the pandemic is now at the forefront in the minds of well-meaning Africans. In response to the call for the adoption of a viable approach to myriads of health problems plaguing Africa, especially the novel COVID-19 pandemic, this paper aims at contending for the development, legitimization, and production of homegrown African medicines on a pragmatic basis as an ideal and positive response to COVID-19, subsequent pandemics, and other diseases that are rife in Africa. To this end, I will first explore pragmatism as a theoretical framework from which my arguments are developed. I will thereafter apply pragmatist theory in contending for the development and legitimization of Africa's epistemic frameworks with particular reference to homegrown African medicines. This is followed by demonstrating that a pragmatic approach to knowledge and ideas requires standard and functional healthcare systems and education to be successful. I will thereafter discuss ethical issues involved in the development, production, and distribution of synthetic COVID-19 vaccines. Finally, I will conclude that pragmatic considerations should ultimately inform the approval of homegrown African medicines and other medical products by national health authorities in African countries for use in Africa.

II. Pragmatism

Pragmatism is a philosophical system or theory propounded by Charles Sanders Peirce and developed and popularized by William James. John Dewey and Richard Rorty were also influential pragmatist thinkers. The theory has been known by various names, e.g., functionalism, instrumentalism, workability, experimentalism, and progressivism. Pragmatism holds that our knowledge, ideas, thinking, and propositions are true and meaningful if they have practical consequences. The utility of an idea, belief, or knowledge is a good measure of truth value. In other words, it is the practical usefulness of an idea or knowledge that makes it true and meaningful. Therefore, pragmatism uses practical consequences of knowledge and ideas as a standard for determining their values and truth.

Pragmatism was first used in 1878 in Peirce's article, *How to Make Our Ideas Clear,* where he states that "belief is a rule for action."[4] In this regard, Rorty asserts that the function of human mind is to produce practical ways of living, thus truth is simply "what passes for good belief."[5] Proponents of this theory agree that truth is the property of certain ideas,

---

[4] Charles S. Pierce, "How to Make Our Ideas Clear," *Popular Science Monthly* 12 (1878): 291.

[5] As quoted in Cheryl Misak, *Truth, Politics, Morality: Pragmatism and Deliberation* (London and New York: Routledge, 2002), 13.

but demand that such ideas must bear an action, that is, must be fruitful and useful in a sensible and practical way. James asserts thus:

> The pragmatic method in such cases is to try to interpret each notion by tracing its respective practical consequences. What difference would it practically make to anyone if this notion rather than that notion were true? If no practical difference whatever can be traced, then the alternatives mean practically the same thing, and all dispute is idle. Whenever a dispute is serious, we ought to be able to show some practical differences that follow from one side or to the other's being right.[6]

Pragmatism, for James, is therefore

> the doctrine that the whole 'meaning' of a conception expresses itself in its practical consequences either in the shape of conduct to be recommended or in that of experiences to be expected, if the conception be true....[7]

James explains that his pragmatic conception of meaning is grounded in Charles Peirce's work:

> Mr. Peirce, after pointing out that our beliefs are really rules for action, said that, to develop a thought's meaning, we need only determine what conduct it is fitted to produce: that conduct is for us its sole significance...To attain perfect clearness in our thoughts of an object, then, we need only consider what conceivable effects of a practical kind the object may involve – what sensations are we to expect from it, and what reactions we must prepare. Our conception of these effects, whether immediate or remote, is then for us the whole of our conception of the object, so far as that conception has positive significance at all.[8]

An idea or knowledge does not count, for the pragmatist, if it has no practical bearing on our experiential world[9] as, for James, "the possession

---

[6] William James, *Pragmatism* (Cambridge: Harvard University Press, 1975), 90.

[7] James M. Baldwin, ed., *Dictionary of Philosophy and Psychology,* vol. 2 (New York: The Macmillan Company, 1902), 321.

[8] James, 28-29.

[9] John A. I. Bewaji, *An Introduction to the Theory of Knowledge: A Pluricultural Approach*

of true thoughts means everywhere the possession of invaluable instruments of action."[10] Therefore, the focus of pragmatism is essentially the utility of ideas, knowledge, or beliefs. In discussing William James' pragmatic theory of truth, Schmitt asserts that an idea or belief is "true just in case it has practical utility in life or belongs to a system of beliefs that has practical utility."[11] Therefore, for the pragmatist, the test of truth is utility, workability, and successful results.[12]

For pragmatists, truth is something that happens to an idea and the idea becomes true when it produces a satisfactory result. While empiricists take every sense perception cognitively and continuously, whether it is practically useful or not, pragmatists accept only ideas that are practically useful. James explicates thus:

> True ideas are those that we can assimilate, validate, corroborate, verify. False ideas are those we cannot. That is the practical difference it makes to us to have true ideas; that, therefore, is the meaning of truth, for it is all that truth is known as. This thesis is what I have to defend. The truth of an idea is not a stagnant property inherent in it. Truth happens to an idea. It becomes true, is made true by events, its verity is in fact an event, a process: the process namely of its verifying itself, its verification. Its validity is the process of its validation.[13]

What the aforementioned assertions mean is that our ideas are true if they can be met with desired success or be successful in meeting our expectations. Pragmatism, obviously, does not discard science. In fact, it adopts scientific attitude in terms of verification in seeking true knowledge and in attending to human existential problems. Pragmatism underpins science in that it holds that true knowledge and ideas are hinged on practical success. However, its point of departure is that it takes a multi-faceted approach to verification and validation rather than being limited to rigid scientific verification and validation procedures. For instance, James asserts that the proposition that God exists is verifiable in the sense that we can ascertain whether it provides

---

(Ibadan: Hope Publications, 2007), 238.

[10] As quoted in Richard L. Kirkham, *Theories of Truth. A Critical Introduction* (Cambridge: The MIT Press, 1992), 92.

[11] Frederick F. Schmitt, "Truth: An Introduction," in *Theories of Truth*, ed. Frederick F. Schmitt (Oxford: Blackwell Publishing, 2003), 9.

[12] Ben O. Eboh, *Basic Issues in Theory of Knowledge* (Nsukka: UNN, 1990), 44.

[13] James, 97.

us with "vital benefits," that is, satisfies our spiritual or religious needs.[14] Such a proposition is true; it works satisfactorily, if it meets such needs. Therefore, the verification and validation of knowledge claims, ideas, or propositions must not rest on scientific testing or experimental procedures.

Pragmatism turns away from abstractions, a priori reasons, static knowledge, fixed principles, and closed systems, but turns toward concreteness, facts, and actions.[15] Peirce maintains that our pragmatic approach to scientific, philosophical, or theological questions should be:

> Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object.[16]

Thus, James views pragmatism as a method that has no dogmas and doctrines apart from its method. What pragmatists are driving at in this regard is that formulations in philosophy, theology, and science should be seen as only approximations (verisimilitude), rather than as absolute truths or knowledge, for such formulations offer us no conclusive answers or solutions to our existential problems. A single fixed formula makes an account of truth rigid, authoritarian, and doctrinaire in pragmatists' view. They therefore insist that there is no single fixed formula and absolute truth but many truths, as there are many concrete successful actions in the truth process.

Thus, James distinguishes between what he calls tough-minded and tender-minded approaches to truth. A tough-minded approach would consider more scientific behavior in the truth process, whereas a tender-minded approach would consider less scientific behavior in the truth process.[17] Our approach to knowledge should be multi-faceted, for we know things from many different perspectives. The value of any theory, for pragmatists, does not rest in its internal verbal consistency, but in its ability to solve human problems. Thus, James asserts that for a theory or an idea to be meaningful, we "must bring out of each word its practical cash value,"[18] that is, its experiential utility or usefulness, whether good or bad, the difference it makes for practical life.

---

[14] Anthony Harrison-Barbet, *Mastering Philosophy* (London: Macmillan Education, 1990), 80.

[15] Ome, and Amam, 330.

[16] Peirce, 291.

[17] Samuel E. Stumpf, and James Fieser, *Philosophy: History and Problems*, 6th ed. (New York: McGraw-Hill, 2003), 399.

[18] As quoted in Stumpf and Fieser, 398.

Dewey, in his brand of pragmatism known as instrumentalism, holds that thinking and ideas are instrumental in solving practical human problems.[19] He insists that inquiry into any knowledge claim should be empirical in method and practically motivated.[20] He explains further that the term pragmatic means "only the rule of referring all thinking, all reflective considerations, to consequences for final meaning and test."[21] The truth of any knowledge claim is thus based upon its usefulness. For Dewey, the best test of the value of any idea or theory is to ask:

> Does it end in conclusions, which, when referred back to ordinary life-experiences and their predicaments, render them more significant, more luminous to us and make our dealings with them more fruitful? [22]

For Dewey, the act that will bring about the most successful outcome is the most valuable. Hence, Dewey frowns upon any system of education that involves learning without doing or practice, as such education will have no practical usefulness. Dewey's instrumentalism is governed by the presuppositions of science, which recognizes the intimate connection between reflection and experiment, thought and action.[23] Human minds are basically problem-solving instruments and the best ways to discover the instrumental means for problems are experience and experiment.

Overall, pragmatists consider ideas and knowledge to be true and meaningful if they can help us make successful connections among various life experiences and can be brought to bear on our existential problems. Truth is what works; an idea, a belief, or knowledge is true and only true if it is functional, that is, if it yields a satisfactory result. Put differently, it is the functional values of ideas, beliefs, and knowledge that make them true.

III. Pragmatism and COVID-19 in Africa

As was explained in the preceding section, pragmatism does not preclude scientific requirements, albeit science may exclude certain things that evidently work. James maintains that a pragmatic approach to knowledge could be tough-minded and tender-minded. For example, while a tough-minded pragmatist would seek a scientific sort of analysis of the effectiveness

---

[19] Stumpf, and Fieser, 405.

[20] Ome, and Amam, 335.

[21] Ome, and Amam, 336.

[22] Stumpf, and Fieser, 405.

[23] Stumpf, and Fieser, 406.

of an herbal tonic from Madagascar in treating COVID-19 patients, a tender-minded one would only concern themselves with the efficacy of the medicine in determining its truth claim. For James, both approaches to truth are valid in their own ways, provided that the object of inquiry (Madagascar's herbal drink) fulfils its useful function. Therefore, a pragmatic approach to pandemics and other viral and common diseases in Africa must not involve rigorous and lengthy scientific validation processes.

Kuhn's seminal work, *The Structure of Scientific Revolutions*, has shown that science does not stand the test of time, as it undergoes periodic revolutions, which he calls a paradigm shift from which a new normal science emerges.[24] Truth is paradigmatic as it is determined by the prevailing paradigm. However, paradigms are always subject to change such that new truth emerges from the new paradigm. To this extent, truth is not always the same: what is true in a previous paradigm may not be true (at least in part) in a new paradigm. There is never a complete overlap between paradigmatic problems that can be solved by the previous and the new paradigms.[25] Kuhn's main contention here is that scientific knowledge does not rest on any foundation of a single absolute truth. There are more truths as there are more successful practical proofs, thus there will never be a complete collection of truths.

Given that scientific methodology is lengthy, rigorous, and does not offer us a cast-iron guarantee, it behooves Africa to be open to the application of what brings practical success efficiently without necessarily following rigid scientific procedures. It calls for the adoption of a pragmatic approach to pandemics and other diseases that are rife in Africa. African countries can achieve this by developing and legitimizing their own epistemic paradigm that will pragmatically attend to myriads of health problems plaguing their people. Misak explains that, for pragmatism, truth and objectivity are matters of what is best for the community of inquirers to believe, that is, what "best fits with the evidence and argument."[26] In the wake of the COVID-19 pandemic, Africans have demonstrated that their richly endowed medicinal plants could be exploited in attending to their multifarious health problems, if legislation enables this and they are given an environment to thrive in.

For example, Madagascar, one of the smaller African countries, has stolen a march on the West by developing an herbal tonic from a medicinal plant known as *Artemisia annua* for the treatment of COVID-19 patients to the consternation of the World Health Organization (WHO). The herbal remedy for COVID-19 produced by Madagascan herbal research institutes has been

---

[24] Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: The University of Chicago Press, 2012), 85.

[25] Ibid.

[26] Misak, 1.

practically tested and proved to be curative. The viability and efficacy of the herbal drink have led Presidents of some other African countries like Equatorial Guinea, Tanzania, Uganda, Egypt, Guinea-Bissau, and Senegal to recognize and promote it, while they ordered for it to be used for the treatment of COVID-19 cases in their respective countries.

Local medical researchers and scientists have also claimed to have discovered herbal medicines for the treatment of COVID-19. For instance, this is the case of the Iris Medical Foundation Drugs and Pharmaceutical founded by late Professor Paul Olisa Ojeih. This pharmaceutical company is committed to drug research, focusing on compounding drugs from plants and enzymes and synthesizing them into potent cures.[27] The company informed the Federal Government of Nigeria that it developed an organic phytomedicine, known as Venedi Elixir, which is effective in treating COVID-19. The alleged curative drug is derived from the enzymes of pharmaceutically engineered plants used to treat complex viral infections.[28] Moreover, in Nigeria, a Roman Catholic priest and one of the country's foremost traditional medical practitioners, Reverend Father Raymond Arazu has proclaimed that the Anambra Traditional Medicine Board headed by him has developed a cure for COVID-19.

Furthermore, a Benedictine monk and priest of the Roman Catholic Church, Reverend Father Anselm Gbenga Adodo, who is the founder of Nigeria's first alternative medicine and research laboratory enterprise in 1997 known as Pax Herbal Clinic and Research Laboratories in Ewu, Edo State of Nigeria, has announced that the research center discovered a clinical COVID-19 herbal drug and expressed their readiness to begin the mass production of the treatment drug immediately after its use has been approved.[29] Adodo also discloses that they have herbal medicines, which they have been producing for over 25 years, for many diseases plaguing Africans, such as malaria, typhoid fever, hypertension, tuberculosis, hepatitis B, diabetes, asthma, prostate problems, male and female infertility, etc.[30]

Rather than pay attention to these claims, the Nigerian government is waiting desperately for experimental drugs to be used in testing prospective

---

[27] Paul O. Ojei, "Coronavirus Pandemic: The Curative and Politics," *PM News*, May 5, 2020, https://www.pmnewsnigeria.com/2020/05/05/coronavirus-pandemic-the-curative-and-politics/.

[28] Ibid.

[29] Rasheed Sobowale, "Why WHO Suspended Chloroquine Clinical Trial for COVID-19 Treatment," *Vanguard Nigeria*, May 29, 2020, https://www.vanguardngr.com/2020/05/why-who-suspended-chloroquine-clinical-trial-for-covid-19-treatment/.

[30] Rasheed Sobowale, "EWU on COVID-19, SARS, Ebola: Inside Catholic Research Centre where Monks Cure with Herbs," *Vanguard Nigeria*, May 17, 2020, https://www.vanguardngr.com/2020/05/ewu-on-covid-19-sars-ebola-inside-catholic-research-centre-where-monks-cure-with-herbs/.

volunteer COVID-19 patients. This constitutes a master-slave mentality that shows strong preference toward being a comfortable slave, rather than being a free man;[31] this attitude stunts indigenous development in Africa. It could be recalled that when US President Trump asserted that hydroxychloroquine and chloroquine might help treat COVID-19, without any scientific evidence backing up the claim, many COVID-19 victims in Africa, including those that had not contracted the disease, rushed to purchase, and even store up the drug. In China, the US, and other countries where the drug has been used experimentally in COVID-19 patients, there is no satisfactory clinical evidence that chloroquine is effective in preventing and managing the pandemic.[32]

Observational research led by Mandeep Mehra of the Brigham and Women's Hospital in the US, which was published on May 22, 2020, and other numerous scientific studies, suggest that chloroquine and hydroxychloroquine are ineffective in treating COVID-19 and might aggravate the disease and increase the death likelihood of COVID-19 patients, given the potentially serious side effects, particularly arrhythmias (irregular heartbeat) both drugs can produce.[33] No wonder it was reported in March 2020 that three COVID-19 victims were hospitalized in the Lagos State of Nigeria after taking chloroquine.[34] This huge rush to obtain the drug was mainly owing to the assertion made by someone from the West, a President of a superpower nation for that matter, despite the position of the WHO that any medication, the efficacy of which has not been proved based on clinical trials, should not be used to treat COVID-19.

It is high time Africa refused to be needlessly tied to the West's apron strings and took its destiny in its hands. Madagascar's COVID-19 herbal remedy should serve as a real eye-opener for Africans to believe in themselves and their vast natural resources, and look inward in search of viable solutions for COVID-19 and other viral and common diseases in Africa. Other African governments should take their cue from the Madagascan government that employs its own knowledge and expeditious validation system to verify and validate the efficacy and safety of its COVID-19 herbal solution and thus endorse it for use without waiting for scientific validation processes. African countries should be committed to intense COVID-19 research and give attention to claims of breakthrough in the cure of COVID-19 by African

---

[31] Sunny Ikhioya, "The Master-Slave Syndrome," *Vanguard Nigeria*, May 27, 2020, https://www.vanguardngr.com/2020/05/the-master-slave-syndrome/.

[32] Stephanie Busari, and Bukola Adebayo, "Nigeria Records Chloroquine Poisoning after Trump Endorses it for Coronavirus Treatment," *CNN*, March 23, 2020, https://edition.cnn.com/2020/03/23/africa/chloroquine-trump-nigeria-intl/index.html.

[33] Sobowale, "Why WHO Suspended Chloroquine Clinical Trial for COVID-19 Treatment."

[34] Busari, and Adebayo.

scientists and alternative medical practitioners, rather than rely only on the WHO for COVID-19 remedies and co-opt their people to be used as experimental guinea pigs and lab rats for the WHO's untested and unproven COVID-19 vaccines in support of its solidarity trials.

Claims by local researchers and scientists to have developed curative medicines, whether modern, herbal, or alternative/complementary medicines, should be subjected to evaluation protocols, grounded on pragmatism, by national health authorities in different African countries; if the medicines successfully pass practical testing and verification, then they should be approved forthwith. Empirical evidence is not only obtained from lengthy scientific validation, but also from other practical experiences. Locally manufactured COVID-19 equipment and other medical equipment should be equally subjected to evaluation and verification on a pragmatic basis for possible approval for use in Africa. What we are driving at is that subjecting homegrown African medicines to safety, toxicological, and efficacy tests as well as clinical trials on volunteer patients, must not follow rigorous scientific validation processes. To rise to the myriads of health challenges confronting Africa, a viable and efficient mode of measuring and determining the effectiveness of homegrown medicines should be developed by the national health authorities in African countries. Homegrown medicines should be approved for use in Africa if they prove to be viable and efficacious after practical testing and verification. It is not necessary for Africa to wait for the scientific validation of its homegrown medicines by the WHO that seems to neither believe nor show enough interest in traditional or herbal African medicines.

The Madagascan government has demonstrated that rigorous and lengthy scientific validation processes are needless in the case of a health emergency, by approving the use of its discovered herbal tonic for the treatment of COVID-19 cases based on its own knowledge and validation system. Similarly, the Israeli health authority has approved for use a series of efficacious vaccines developed by the Israel Institute for Biological Research using its own validation system.[35] In response, the Israeli President is planning to build a plant for the production of vaccines for the treatment of COVID-19 patients. In recognition of the need for heterodox approaches to the pandemic ravaging the world, the House of Representatives in Nigeria has called on the Nigerian government to ignore caution by the WHO against the use of any drug on COVID-19 patients the efficacy of which is not scientifically proven, and approve and support

---

[35] Tordue Salem, "COVID 19: Ignore WHO, Go for Local Cures, Reps Tell FG," *Vanguard Nigeria*, May 13, 2020, https://www.vanguardngr.com/2020/05/covid-19-ignore-who-go-for-local-cures-reps-tell-fg/.

the use of alternative locally developed remedies for the management and treatment of COVID-19 in Nigeria.[36]

Although a theory yields certain practices by explaining the rationale behind the practices, sometimes theories do not precede practices. For instance, Africans' forebears used neem leaves (scientifically known as *Azadirachta indica*, but popularly known as *Dogoyaro* in the Hausa language) and other herbs to treat malaria for ages, but could not offer explanations for why the herbs worked. They did not know it was parasites that engendered the disease, how it got into the human blood system, and why quinine could cure the disease. However, today, we know that the disease is caused by malaria parasites, and that it enters the circulatory system when a female *Anopheles* mosquito infected with the parasite bites a person; the neem plant contains quinine that cures malaria and can also be used to treat similar diseases.

The practice of any discovery, such as the Madagascan COVID-19 herbal remedy, whose theoretical formulations (scientific explanations) are not yet developed, should be accepted on the grounds that it serves the purpose for which it is made, which in this case is to contain the COVID-19 pandemic. African herbal medicines and other homegrown medical products should not be dismissed because there is no available verifiable scientific knowledge of the whys and wherefores of their viability. The fact that a scientifically verifiable explanatory theory is not yet offered does not mean that it cannot be provided in the future. The efficacy of any medicine presupposes a workable theory behind it, albeit the explanation of such a theory may not be offered for the moment. Scientific explanations for why homegrown African medicines work can come later. Whatever passes the litmus test of pragmatism is a strong candidate for the test of scientificity. The success or effectiveness of the practical application of homegrown African medicines and other medical products should be the ultimate criterion for their endorsement for medical treatments in Africa. The fact that they pay off should be a justification for approving them, their scientific explanations pending.

IV. On Health

The global life or death situation brought about by the outbreak of the COVID-19 pandemic had started before the first COVID-19 case in Africa was recorded in Egypt on February 14, 2020. Prior to Africa's index case, the pandemic outbreak had become a cause for concern in the continent. The fact that the COVID-19 pandemic became a matter of considerable public concern was exacerbated by the weak healthcare systems of most African nations. Therefore, ultimately, Africa was apprehensive about its

---

[36] Salem.

level of preparedness and readiness to combat the virulent and deadly pandemic, especially when considering the fact that the pandemic had already overpowered robust health systems of some developed countries like Italy, Spain, and the United States.

Recently, African nations have somehow managed to tackle other disease outbreaks, such as Ebola virus disease, cholera, Lassa fever, and monkey pox outbreaks, without stepping up the medical treatments and preventive measures, or strengthening the respective health systems of the different nations. However, the COVID-19 pandemic is a completely different situation, as it overwhelms Africa and exposes its weak health systems. Thus, some African countries like Nigeria are already sitting ducks for the pandemic, given the plethora of understaffed, underequipped, and underfunded health care centers and hospitals that cannot provide the high quality and culturally appropriate healthcare services required to contain and manage the raging and ravaging pandemic.[37]

The United Nations Development Programme's (UNDP) report on current expenditure on healthcare suggests that the healthcare sector is badly underfunded in many African countries. For instance, healthcare expenditure in terms of the Gross Domestic Product (GDP) percentage spent on healthcare in 2015 was 2.5 in South Sudan, 3.3 in Eritrea, 3.6 in Nigeria, 3.8 in Papua New Guinea, 4.0 in Benin, Senegal, and Ethiopia, 4.3 in the Democratic Republic of Congo, 4.5 in Guinea, 4.6 in Mauritania and Chad, 5.2 in Madagascar, 5.4 in Côte d'Ivoire, Mozambique, and Burkina Faso, and 5.8 in Mali.[38] Therefore, it is not surprising that the COVID-19 pandemic has put a huge strain on the limited health services of African countries.

Many Africans do not have access to basic healthcare due to dilapidated hospital facilities and poor health service delivery. The low government expenditure on healthcare over the past years does not meet the global healthcare standards, thus African nations have low-rated healthcare systems. For instance, the healthcare system of Nigeria currently ranks 187th out of 191 healthcare systems globally.[39] Drug research, development, and production are very expensive. Pharmaceutical research institutes across Africa cannot raise the funds required for drug research and development without the support of African governments and private investments that presupposes enabling laws and creating a favorable environment.

---

[37] Sola Ogundipe, "How COVID-19 Rediscovered Nigeria's Health Care System," *Vanguard Nigeria*, May 29, 2020, https://www.vanguardngr.com/2020/05/special-report-how-covid-19-rediscovered-nigerias-health-care-system/.

[38] United Nations Development Programme, *Human Development Indicators and Indices: 2018 Statistical Update Team* (New York: United Nations Development Programme, 2018), 52-53.

[39] Ogundipe.

Recently, the US government approved a budget of $1 billion for AstraZeneca, a British-Swedish pharmaceutical company, to carry out research for a COVID-19 vaccine.[40] This is how a government's commitment to research and development is demonstrated. When drugs are developed and produced by local pharmaceutical industries, a high percentage of essential drugs will be available for domestic consumption, thereby reducing heavy dependence on the importation of drugs from foreign countries to meet local needs. This is necessary, especially in the case of a pandemic outbreak that can only be contained by lockdowns and border closure.

Some African countries like Nigeria that were affected the most by the COVID-19 pandemic, had a scarcity of essential drugs at the onset of the pandemic outbreak due to lockdowns and border closures in China and India, where most essential drugs are imported from. For a pragmatic approach to work, the healthcare systems of African nations must improve. It cannot be rightly gainsaid that Africa cannot respond to a pandemic and other prevalent African diseases, such as malaria, typhoid fever, cholera, Lassa fever, tuberculosis, diarrhea, small pox, hepatitis, yellow fever, measles, cancer, stroke, diabetes, and coronary heart diseases in a pragmatic way if healthcare, health workers, and health facilities are inadequate. This is mainly because the aforementioned are indispensable instruments for the practical application of homegrown African medicines. They are the means through which viable orthodox, alternative, and herbal medicines can be administered to people.

Therefore, the proposed pragmatic approach to Africa's multifarious health problems calls for the provision of adequate healthcare centers and hospitals, the employment of sufficient, experienced, and well-trained healthcare workers, and the provision of adequate and state-of-the-art medical facilities across states or regions in all African countries. Besides, government expenditure in healthcare as well as government and private investments in the pharmaceutical sector for the research, development, and production of drugs are critical to implement improvements in the various healthcare systems.

Monitoring and stemming the spread of a pandemic requires effective and early testing. Tedros Ghebreyesus, the Director-General of the WHO, discloses that "lack of testing is leading to a silent epidemic in Africa."[41] Testing early and often ensures that carriers of the pandemic are detected and isolated to minimize the rate of contagion. By March 2020, when the COVID-19 pandemic had reached epidemic proportions in Europe, the fatality

---

[40] Ibid.

[41] *Premium Times*, "WHO Fears 'Silent' Virus Epidemic unless Africa Prioritises Testing," May 26, 2020, https://www.premiumtimesng.com/news/top-news/394580-who-fears-silent-virus-epidemic-unless-africa-prioritises-testing.html.

rate of Germany was remarkably low (0.5%),[42] compared to those of Italy (10%), Spain (8.9%), and France (5%). Drosten, whose team of researchers developed the first COVID-19 test that was publicly available, attributes the low fatality rate of Germany to the ability of his country to test early and often.[43] He discloses that Germany has been testing approximately 120,000 people a week for COVID-19 since late February 2020 which helped the country stand out among the worst hit countries in Europe.[44]

To respond robustly and positively to COVID-19 and subsequent pandemics, testing sites must be created across states/regions of African countries. Africa must not wait until a new pandemic breaks out before it starts getting on the stick, for it is not good at flattening the curve. The most efficient means of achieving this is to set up ultra-modern molecular and diagnostic laboratories that have the capacity to undertake thousands of tests per day in all university teaching hospitals (UTHs) across states of African countries. The UTHs could easily be converted into testing centers during a pandemic outbreak, after receiving accreditation by the Center for Disease Control in each African country.

V. On Education

For a pragmatic approach to be effective, university lecturers must be encouraged and supported to be fully committed to research and development, through improved salaries and allowances as well as the provision of adequate research grants and educational facilities. Students, particularly those that take medical and pharmaceutical related courses, need a holistic experience and pragmatic education that involves learning and practicing, which will equip them with the necessary skills, knowledge, ideas, materials, and tools to excel in all spheres of their future careers and work in a pragmatic way in attending to future life challenges that may confront them and their people.

Africa needs medical and pharmaceutical lecturers and researchers in ivory towers to be at the forefront of research in viable homegrown medicines (modern or herbal) to tackle myriads of the health problems plaguing Africa. However, these ideals are not attainable unless there are improved conditions of service, an adequate provision of state-of-the-art laboratories and lab equipment, and an increase in governmental expenditure on research and development. Some groundbreaking medical research is lying

---

[42] Rob Schmitz, "Why Germany's Coronavirus Death Rate Is Far Lower Than in Other Countries," NPR, March 25, 2020, https://www.npr.org/2020/03/25/820595489/why-germanys-coronavirus-death-rate-is-far-lower-than-in-other-countries.

[43] Schmitz.

[44] Ibid.

dormant in university libraries across Africa. Enormous funds are required for the development and production of medical discoveries.

Regrettably, education is chronically underfunded by African governments. The United Nations Children's Fund recommends that developing countries should allocate no less than 15% of their annual budgets to education. However, over the past years, developing African countries have consistently failed to provide the bare minimum of allocation to maintain acceptable educational standards in Africa. Thus, the African education system is waning. In Nigeria, for example, there has been a steady decline in the education budget. According to the "2020 Budget Analysis and Opportunities" report, the budget for education was 12.3%, 9.3%, 7.3%, 7.1%, and 6.5% in 2015, 2016, 2017, 2018 and 2019, and in the pre-COVID-19 2020 budget, respectively.[45]

The 2018 UNDP report on government expenditure on education measures in terms of the GDP percentage allotted to education during 2012–2017 shows that the education expenditures of South Sudan (1.8%), Guinea-Bissau (2.1%), Madagascar (2.1%), the Democratic Republic of the Congo and Uganda (2.3%), Guinea (2.4%), Mauritania (2.6%), Cameroun, Gambia, and Liberia (2.8%), and Chad (2.9%) were negligible.[46]

VI. On Ethics

Morality is the basis of ethics and *eudaimonia* (happiness, welfare/human flourishing) is the standard of morality. Aristotle asserts that happiness is the supreme good for all human beings, albeit there is no general consensus as to what sort of life counts as happy or what happiness actually consists of.[47] However, in general, all human beings as moral agents have the right to life, liberty, and security, while the society of which they are parts has a moral obligation to protect these rights. Therefore, security and freedom to choose how to lead one's life as well as treating all humans with care and respect are basic principles of a good society.[48]

COVID-19 vaccine experimentation, development, production, and allocation as well as COVID-19 inoculation/vaccination raise certain ethical concerns. How could a limited supply of early vaccines be allocated or distributed fairly and effectively? Who should be prioritized in the vaccine

---

[45] *Vanguard Nigeria*, "How COVID-19 Can Help Nigeria Rethink Education Policy – AACS," May 27, 2020, https://www.vanguardngr.com/2020/05/how-covid-19-can-help-nigeria-rethink-education-policy-%E2%80%95-aacs/.

[46] United Nations Development Programme, 56-57.

[47] Aristotle, *The Nichomachean Ethics of Aristotle*, 10th ed., trans. F. H. Peters (London: Keegan Paul, Trench, Trübner & Co., 1906), 1094a.

[48] Barry Knight, *Rethinking Poverty: What Makes a Good Society?* (Bristol: Policy Press, 2017), 89.

distribution? What ethical values should guide the vaccine allocation to maximize benefits – the overall number of lives saved? Do the benefits of the synthetic vaccine outweigh the risks associated with its clinical trials and the pandemic inoculation/vaccination? Can we ever justify on moral grounds human lives wasted during inoculation? Are COVID-19 vaccines worth, in terms of their effectiveness, enormous public funds allocated for their research, development, production, and procurement?

As noted earlier, synthetic vaccines are required to pass rigorous safety and efficacy standards. It usually takes at least two years for the vaccines to undergo animal tests before clinical trials, thus vaccine research, development, and production could potentially take up to a decade. Experience from the 2009 pandemic influenza A (H1N1) virus as well as the onset of the COVID-19 pandemic shows that there is a proclivity for the world's leading economies to place their national interests in research, development, production, and distribution of vaccines above global interests.[49] Vaccine nationalism, which entails a desire and push by a nation to first get access to a vaccine supply and dictate vaccine production and distribution in its favor, raises some moral questions.

For instance, at the onset of the COVID-19 pandemic, superpowers, such as the US, China, and Russia, competed for the development of a COVID-19 vaccine and for who would become the first supplier. This intense rivalry among superpowers induced them to speed up their vaccine development and production with regard to clinical trials and pushed for quicker regulatory approval. For instance, the AstraZeneca and Moderna COVID-19 vaccines were approved for public use in November (which is less than a year after the pandemic outbreak) and December (which is just year after the outbreak of the pandemic), respectively. There are grave risks associated with hasty clinical trials and regulatory approval. The efficacy and safety scientific standards of COVID-19 vaccines might be compromised, thus constituting the vaccines unable to effectively treat the viral infection, protect people against it, and help achieve herd immunity; global herd immunity is achieved when a sufficient part of the population becomes immune to a virulent disease to such an extent that the spread of the disease is contained. If vaccines developed and produced in this manner turn out to be less effective, or have unconscionable harmful side effects, national vaccination plans would be complicated and global vaccine supply chains would be interrupted, thus frustrating further any globally coordinated efforts to contain the pandemic.

The AstraZeneca/Oxford vaccine is a case in point. Though the vaccine is widely used by most countries for large-scale vaccination programs, it

---

[49] Macro Hafner, et al., *COVID-19 and the Cost of Vaccine Nationalism* (Santa Monica, California, and Cambridge, UK: Rand Corporation, 2020), iii.

does not offer a guarantee of treating COVID-19. In a study, a volunteer suffered from a neurological condition known as transverse myelitis – a "possibly related severe side effect"[50] of the vaccine – after participating in a clinical trial. Furthermore, a 49-year-old nurse succumbed to severe bleeding disorders days after receiving the vaccine, which forced Denmark, Australia, Lithuania, and a few other countries to halt vaccinations with doses of the vaccine from the same batch.[51] What is more, the UK report on the AstraZeneca vaccine disclosed that 547 people had died from adverse side effects of the inoculation between January 4 and March 9, 2021.[52] Some other countries like France, Germany, Italy, Norway, and Sweden have suspended the use of the AstraZeneca shot due to reports of European Union (EU) countries citing its possible serious side effects, especially cases of post-jab hemorrhages.

In this regard, a rational mind may ask: Can we morally justify deaths and harm resulting from severe side-effects of the vaccine? Herein, we shall consider two ethical perspectives, namely, deontological and teleological theories. According to deontologists, certain actions, which border on human values like killing, are wrong ontologically, that is, in their very nature, and neither circumstances nor consequences can make them rights.[53] Peschke contends that "the judgement of the morality of an action is not possible without a careful study of the nature of being, that is without due consideration of the deontological factors involved."[54] In their view, the nature of a being and the nature of an action are ultimate criteria for making moral judgements. However, the circumstances and consequences (intended or anticipated) are also considered in making moral decisions. For the teleologists/consequentialists, the rightness or wrongness of an action is determined by the consequences (good or evil) resulting from the action,[55] regardless of the nature of an action or the means of achieving the end. This approach to making moral decisions considers an action as morally right if its good consequences outweigh bad ones. Teleologists want moral agents to focus only on the ultimate end or goal

---

[50] As quoted in *Vanguard Nigeria*, "Five Things to Know about the AstraZeneca/Oxford Vaccine," March 11, 2020, https://www.vanguardngr.com/2021/03/five-things-to-know-about-the-astrazeneca-oxford-vaccine/.

[51] *Medical Express*, "Five Things to Know about the AstraZeneca/Oxford Vaccine," March 21, 2021, https://medicalxpress.com/news/2021-03-astrazenecaoxford-vaccine.html.

[52] Nneoma Benson, "275 Dead: UK Report on AstraZeneca Vaccine Sends Message of Caution to Nigeria, Others," *The Whistle*, March 16, 2021, https://thewhistler.ng/275-dead-uk-report-on-astrazeneca-vaccine-sends-message-of-caution-to-nigeria-others/.

[53] Joseph I. Omoregbe, *Ethics: A Systematic and Historical Study* (Lagos: Joja Educational Research and Publishers Limited, 1993), 73.

[54] Karl-Heinz Peschke, *Christian Ethics*, 3rd ed., vol. 1 (Oregon: Wipf and Stock Publishers, 2012), 134.

[55] Ben O. Eboh, *Living Issues in Ethics* (Nsukka: Afro-Orbis Publishing Co. Ltd, 2005), 78.

of an act in making moral decisions. However, the ultimate end is very difficult to evaluate, as there are different definitions of ultimate goals resulting from different forms of teleology with different criteria and evaluation outcomes.[56] Besides, it is very difficult to determine all the proximate and remote consequences of an action and weigh them appropriately.[57]

Deontologists will argue that the victims of the harmful effects of the COVID-19 vaccine possess dignity and rights to life and security as others, and they did not implore to be killed. Hence, it is unethical to denude them of their right to life unsolicitedly. For consequentialists, the death of the COVID-19 vaccine victims is an unintended consequence of the vaccine that is meant to save lives, thus the benefits outweigh the risks. However, a risk-benefit analysis in vaccine experimentation and pandemic inoculation/vaccination is very difficult to undertake. Such a task would require detailed information about the consequences of the vaccine, empirical insights, and knowledge of all contingencies. Thus, the risk-benefit ratio seems to be indeterminate. Besides, due to severe side-effects associated with synthetic vaccines, COVID-19 vaccine experimentation and inoculation against the pandemic present more than minimal risks. A risk is said to be minimal if

> the probability and magnitude of harm or discomfort anticipated [...] are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.[58]

Given that the risks associated with new synthetic vaccine experimentation and inoculation against a viral infection are anticipated and this incurs more than minimal risks, one may be inclined to think that it amounts to treating human subjects, who are ends in themselves by virtue of their inherent dignity, as a means to an end. This constitutes an affront to human dignity. In any case, nations of the world have to balance the need to protect the greatest number of people with the need to adequately understand how a new synthetic vaccine will perform (including its likely harmful side-effects) and protect people when administered. Regrettably, national governments have failed to balance this need as regards the development, production, allocation, and administration of COVID-19 vaccines.

Again, COVID-19 vaccines need not only to be effective but also to be administered to large portions of the global population to achieve

---

[56] Peschke, 128.

[57] Ibid., 131.

[58] As quoted in Vasantha Muthuswamy, "Ethical Issues in Clinical Research," *Perspectives in Clinical Research* 4, no. 1 (2013): 10.

herd immunity. Wealthier countries have signed direct bilateral deals with COVID-19 vaccine manufacturers to secure a stock for their own population.[59] The saying that he who pays the piper calls the tune still holds true. For instance, the AstraZeneca vaccine, which was first approved for use in Britain that earlier ordered 100 million doses of the vaccine, has been constantly supplied by the British-Swedish firm to Britain to meet its demand, while the firm delays the delivery of doses of the vaccine to other EU member states, thus frustrating their vaccination programs.[60] In January, the firm announced that "it could only deliver one-third of the 120 million doses initially promised the 27 EU member states in the first quarter,"[61] which prompted the EU to invoke a Brexit deal protocol over export controls on COVID-19 vaccines. European Commission President, Ursula von der Leyen disclosed that the firm had delivered less than 10% of the doses ordered by other EU member states between December 2020 and March 2021, warning that other countries could block exports.[62]

Thus, the vaccine may not be adequately allocated to low-income and middle-income countries that house approximately 85% of the global population.[63] Wealthier countries might even hoard vaccine doses above and beyond their populations' needs; after all, they can afford it. As high-income countries scramble for limited vaccine supplies, low-income countries that could not fund vaccine production are abandoned to their fate. Consequently, early available vaccines are not equitably allocated, and poorer countries do not have access to them. However, what is the moral justification for spending humongous public funds on the research, production, and procurement of COVID-19 vaccines, if the global population is far from attaining herd immunity? Given the increasing health costs and financial constraints worldwide, traditional bilateral and multilateral African development partners do not have the financial wherewithal to wholly fund COVID-19 vaccination programs in Africa.[64]

Therefore, African governments still have to allocate millions of dollars for the vaccine procurement to vaccinate/inoculate their teeming population.

---

[59] Hafner, et al., iv.

[60] *Vanguard Nigeria*, "Five Things to Know about the AstraZeneca/Oxford Vaccine."

[61] Ibid.

[62] Ibid.

[63] Olivers J. Wouters, et al., "Challenges in Ensuring Global Access to COVID-19 Vaccines: Production, Affordability, Allocation, and Deployment," *Health Policy* 397, no. 10278 (2021): 1025.

[64] Danielle Serebro, "COVID-19 Vaccine Financing and Purchasing in Africa: Wherefrom the Money?" *CABRI*, February 19, 2021, https://www.cabri-sbo.org/en/blog/2021/covid-19-vaccine-financing-and-purchasing-in-africa-wherefrom-the-money.

If the two-fold objectives of the vaccine, namely, the direct protection of a sufficient percentage of the global population and the containment of the viral transmission cannot be achieved, on account of vaccine nationalism and global competition, this then amounts to a waste of public funds that could have been channeled into cushioning the adverse economic effects of the pandemic on Africans through the provision of carefully planned and coordinated fiscal and monetary stimuli and job creations to enhance people's well-being. COVID-19 vaccines are designed for the treatment of the vast majority of patients and the protection of the global population so as to achieve herd immunity and improve the well-being of the human race. However, if economic power determines who and when gets the vaccines as well as the quantity allocated, then we can safely assert that the two-fold objectives are far from being achieved.

Besides, this "my nation first" approach to COVID-19 vaccine production means that low-risk individuals in high income countries get vaccinated or inoculated first before high-risk individuals in low-income countries who might even die before the vaccine is made available in their countries. Therefore, vaccine nationalism or global competition tends to prevent the vaccine from reaching relatively high-risk individuals early. One cannot help but wonder if there is any moral justification for administering COVID-19 vaccines first to low-risk individuals rather than high-risk individuals, regardless of their countries. Also, if the vaccines are meant to save lives, is it morally right to give priority to low-risk individuals on the basis of vaccine nationalism or global competition? Avoidable deaths and serious harm ensue from failure to prioritize relatively high-risk groups – vulnerable people and front-line workers – in vaccine allocation, irrespective of their countries. This problem is compounded by inequitable access to early vaccination in economically disadvantaged countries after the late and insufficient supply of vaccines, as these are not accessible and available to all people and only the affluent could afford them. According to Wu et al., "allocation guidelines must balance the obligation to assist individuals most likely to benefit against the obligation to secure the greatest aggregate benefit across the population."[65] Vaccine allocation should also be guided by considerations of fairness, which entails taking into account categories of people such as frontline health workers and vulnerable people who are suffering from chronic diseases like asthma, diabetes, and heart diseases and thus are more susceptible to COVID-19 infection. Therefore, early vaccine allocation and use seem to be devoid of any value of fairness.

However, these ethical issues involved in synthetic vaccines do not arise in the development, production, and administration of herbal and alternative/

---

[65] Joseph H. Wu, et al., "Allocating Vaccines in a Pandemic: The Ethical Dimension," *The American Journal of Medicine* 133, no. 11 (2020): 1241.

complementary medicines in Africa. For instance, the gaping inequality in synthetic vaccine distribution, access, and use is bridged in the distribution of herbal and alternative medicines in Africa, owing to their availability, variety, accessibility to all, and low cost. Moreover, herbal and alternative medicines are natural remedies with minimal or no harmful side-effects and they have been shown to be effective in treating and managing viral infections at their early stages in the human body. They have the capability to boost and regulate an innate immune system to respond effectively to harmful alien antigens of a virus and prevent the virus from attaching its spike protein to human cells, thus forestalling the propagation and replication of the virus in the human host. For instance, a study by Ngcobo et al. conducted to evaluate the effects of African herbal tonics on immune and inflammatory responses using peripheral blood mononuclear cells, THP-1 monocytes, and bacteria-infected rats, showed that the tonics stimulate the secretion of cytokines, which interact with innate immune system cells to regulate the body's response to viral infections as well as inflammatory responses without any significant toxicity.[66]

It is no wonder that herbal extracts from *Andrographis paniculata* (commonly known as green *chiretta*) have been approved by the Food and Drug Administration (FDA) of Thailand to serve as alternative treatments for COVID-19, owing to their efficacy in curbing the viral infection and curtailing drastically the severity of inflammation. Thai's government official, Taweeslip Witsanuyotin, who is a spokesperson for the national COVID-19 response center, noted that human trials showed that if the herbal medicine is administered on a patient within hours of testing positive, his/her condition improves within three days of the treatment without side effects.[67] According to the chairman of the Bioresources Development Group, Professor Maurice Iwu, empirical data on InterCEDD, known as the "IHP Detox Tea," which contains *Andrographis paniculata* as its key ingredient, and is produced by the group, overwhelmingly revealed that the herbal drug is highly efficacious in treating COVID-19 at its early stage.[68]

Furthermore, herbal and alternative/complementary medicines are extremely versatile in terms of their potential to treat and manage all kinds of viral infections and other diseases at their early stages in the human

---

[66] Mlungisi Ngcobo, et al., "The Immune Effects of an African Traditional Energy Tonic In Vitro and In Vivo Models," *Evidence-Based Complementary and Alternative Medicine* 6310967 (2017): 10.

[67] As quoted in Chris Onuoha, "COVID-19: Iwu Restates Efficacy of Herbal Medicine," *Vanguard Nigeria*, January 10, 2021, https://www.vanguardngr.com/2021/01/covid-19-iwu-restates-efficacy-of-herbal-medicine/.

[68] As quoted in Onuoha, "COVID-19: Iwu Restates Efficacy of Herbal Medicine."

body; however, synthetic vaccines lack that versatility. A synthetic vaccine is basically developed and produced by attenuating or inactivating whole or fragmented structural proteins of a disease-causing virus that can induce immune responses and neutralize antibodies in infected humans. For instance, the AstraZeneca vaccine is made by modifying a portion of the spike protein (antigen) of the SARS-CoV-2 in order to stimulate the production of antibodies and an adaptive immune system that recognizes the novel coronavirus.[69] A synthetic vaccine is therefore designed to combat a particular pathogen and protect humans from it by developing adaptive immunity to an infection.

Accordingly, a synthetic vaccine within the context of viral infections is tailored to treat a specific strain of a virus and not all viral strains. Thus, the COVID-19 vaccine is prone to be less effective in treating other coronavirus variants. The emergence of a new coronavirus strain would therefore require further extensive research and studies to develop and produce a vaccine with all the accompanying extravagant financial demands and further clinical trials with great risks involved. For example, the Republic of South Africa (RSA) has paused the rollout of the AstraZeneca vaccine after a study showed that the vaccine offered "minimal protection against a mild and moderate new COVID-19 variant recently identified first in RSA."[70] On the whole, the moral and pragmatic values of herbal and complementary/alternative medicines lie in their efficacy, versatility, availability, accessibility, affordability, and minimal or no harmful side effects.

VII. Conclusion

In this paper, I argued that a pragmatic approach to COVID-19 and other common diseases in Africa is among the most efficient and effective ways of dealing with myriads of health problems plaguing Africans. Pragmatism does not negate science, rather it affirms it. It only differs from science in that it insists that a practically verified successful idea/knowledge need not pass through rigorous and scientific processes of validation before it can be recognized and endorsed. It adopts a scientific attitude toward ideas and knowledge, but at the same time remains open to other modes of verification that can establish efficiently the workability of any idea, belief, or knowledge. Such a flexible approach to knowledge ensures that the approval

---

[69] *Vanguard Nigeria*, "Five Things to Know about the AstraZeneca/Oxford Vaccine."

[70] *Vanguard Nigeria*, "UK Says Confident in COVID-19 Vaccines as S. Africa Pauses AstraZeneca Rollout," February 8, 2021, https://www.vanguardngr.com/2021/02/uk-says-confident-in-covid-19-vaccines-as-s-africa-pauses-astrazeneca-rollout/.

of the utilization of any knowledge or idea that is practically successful is not delayed or rubbished by rigorous and lengthy processes of verification and validation.

Therefore, adopting a pragmatic approach to knowledge does not imply that Africa discards science, rather it suggests that Africa is opening its mind to other practical and efficient modes of testing and verifying the workability of any knowledge claim for the possible utilization of such knowledge in attending to Africa's existential problems. Empirical evidence does not only rest on scientific validation, but also on other practical experiences. Accordingly, if the practical application of any idea is successful, then it is empirical. Any pandemic outbreak is a matter of life and death. It is suicidal to wait for the lengthy scientific validation of new synthetic vaccines. Hence, Africa's approach to COVID-19 and other pandemics should be predicated on what works efficiently.

Just as the outbreak of COVID-19 breaks all cultural, social, economic, political, and religious protocols precipitously, an approach to the pandemic should breach scientific protocol, provided that the approach is practically effective and efficient. An African proverb from the Igbo people of Nigeria that desperate situations call for desperate measures (*Anụ gba ajọ ọsọ, a jụjụa ya ajọ egbe*) gives credence to the aforementioned assertion. Therefore, it calls for the development and legitimization of Africa's epistemic system grounded in pragmatic principles. Pragmatic considerations should ultimately inform the approval of homegrown African medicines and other medical products by national health authorities in African countries for use in Africa. Homegrown orthodox and alternative/complementary medicines should be approved if they prove to be effective based on Africa's own knowledge and validation system. Ultimately, what matters is that the end-result is successful – the medicines work well in practice.

References

Aristotle. *The Nichomachean Ethics of Aristotle*, 10th ed. Translated by F. H. Peters. London: Keegan Paul, Trench, Trübner & Co., 1906.

Baldwin, James M., ed. *Dictionary of Philosophy and Psychology*, vol. 2. New York: The Macmillan Company, 1902.

Benson, Nneoma. "275 Dead: UK Report on AstraZeneca Vaccine Sends Message of Caution to Nigeria, Others." *The Whistle*, March 16, 2021. https://thewhistler.ng/275-dead-uk-report-on-astrazeneca-vaccine-sends-message-of-caution-to-nigeria-others/.

Bewaji, John A. I. *An Introduction to the Theory of Knowledge: A Pluricultural Approach*. Ibadan: Hope Publications, 2007.

Busari, Stephanie, and Bukola Adebayo. "Nigeria Records Chloroquine Poisoning after Trump Endorses it for Coronavirus Treatment." *CNN*, March 23, 2020. https://edition.cnn.com/2020/03/23/africa/chloroquine-trump-nigeria-intl/index.html.

Eboh, Ben O. *Basic Issues in Theory of Knowledge*. Nsukka: UNN, 1990.

Eboh, Ben O. *Living Issues in Ethics*. Nsukka: Afro-Orbis Publishing Co. Ltd, 2005.

Ejike, Cyril E. "COVID-19 and African Traditional Medicines." In *COVID-19 and Afrocentric Perspectives: Health and Economic Implications*, edited by Ikechukwu Anthony Kanu, Chiugo C. Kanu, and Ejikemeuwa J. O. Ndubisi, 1-16. Maryland: The Association for the Promotion of African Studies, 2021.

Hafner, Marco, Erez Yerushalmi, Clement Fays, Eliane Dufresne, and Christian van Stolk. *COVID-19 and the Cost of Vaccine Nationalism*. Santa Monica, California, and Cambridge, UK: Rand Corporation, 2020.

Harrison-Barbet, Anthony. *Mastering Philosophy*. London: Macmillan Education, 1990.

Ikhioya, Sunny. "The Master-Slave Syndrome." *Vanguard Nigeria*, May 27, 2020. https://www.vanguardngr.com/2020/05/the-master-slave-syndrome/.

James, William. *Pragmatism*. Cambridge: Harvard University Press, 1975.

Kirkham, Richard L. *Theories of Truth. A Critical Introduction*. Cambridge: The MIT Press, 1992.

Knight, Barry. *Rethinking Poverty: What Makes a Good Society?* Bristol: Policy Press, 2017.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press, 2012.

Medical Express. "Five Things to Know about the AstraZeneca/Oxford Vaccine." *Medical Express*, March 11, 2021. https://medicalxpress.com/news/2021-03-astrazenecaoxford-vaccine.html.

Misak, Cheryl. *Truth, Politics, Morality: Pragmatism and Deliberation*. London and New York: Routledge, 2000.

Muthuswamy, Vasantha. "Ethical Issues in Clinical Research," *Perspectives in Clinical Research* 4, no. 1 (2013): 9.13.

Ngcobo, Mlungisi, Nceba Gqalieni, Vinny Naidoo, and Protus Cele. "The Immune Effects of an African Traditional Energy Tonic In Vitro and In Vivo Models," *Evidence-Based Complementary and Alternative Medicine* 6310967 (2017): 1-14.

Ogundipe, Sola. "How COVID-19 Rediscovered Nigeria's Health Care System." *Vanguard Nigeria*, May 29, 2020. https://www.vanguardngr.com/2020/05/special-report-how-covid-19-rediscovered-nigerias-health-care-system/.

Ojeih, Paul O. "Coronavirus Pandemic: The Curative and Politics." *PM News*, May 5, 2020. https://www.pmnewsnigeria.com/2020/05/05/coronavirus-pandemic-the-curative-and-politics/.

Ome, Emmanuel M., and William Amam. *Philosophy and Logic for Everybody*. Enugu: Institute for Development Studies, 2004.

Omoregbe, Joseph I. *Ethics: A Systematic and Historical Study*. Lagos: Joja Educational Research and Publishers Limited, 1993.

Onuoha, Chris. "COVID-19: Iwu Restates Efficacy of Herbal Medicine." *Vanguard Nigeria*, January 10, 2021. https://www.vanguardngr.com/2021/01/covid-19-iwu-restates-efficacy-of-herbal-medicine/.

Peirce, Charles S. "How to Make Our Ideas Clear." *Popular Science Monthly* 12 (1878): 286-302.

Peschke, Karl-Heinz. *Christian Ethics*, 3rd ed., vol. 1. Oregon: Wipf and Stock Publishers, 2012.

Premium Times. "Who Fears 'Silent' Virus Epidemic unless Africa Prioritises Testing." May 26, 2020. https://www.premiumtimesng.com/news/top-news/394580-who-fears-silent-virus-epidemic-unless-africa-prioritises-testing.html.

Salem, Tordue. "COVID 19: Ignore WHO, Go for Local Cures, Reps Tell FG." *Vanguard Nigeria*, May 13, 2020. https://www.vanguardngr.com/2020/05/covid-19-ignore-who-go-for-local-cures-reps-tell-fg/.

Schmitt, Frederick F. "Truth: An Introduction." In *Theories of Truth*, edited by Frederick F. Schmitt, 1-38. Oxford: Blackwell Publishing, 2003.

Schmitz, Rob. "Why Germany's Coronavirus Death Rate Is Far Lower Than in Other Countries." *NPR*, March 25, 2020. https://www.npr.org/2020/03/25/820595489/why-germanys-coronavirus-death-rate-is-far-lower-than-in-other-countries.

Serebro, Danielle. "COVID-19 Vaccine Financing and Purchasing in Africa: Wherefrom the Money?" *CABRI*, February 19, 2021. https://www.cabri-sbo.org/en/blog/2021/covid-19-vaccine-financing-and-purchasing-in-africa-wherefrom-the-money.

Sobowale, Rasheed. "EWU on COVID-19, SARS, Ebola: Inside Catholic Research Centre where Monks Cure with Herbs." *Vanguard Nigeria*, May 17, 2020. https://www.vanguardngr.com/2020/05/ewu-on-covid-19-sars-ebola-inside-catholic-research-centre-where-monks-cure-with-herbs/.

Sobowale, Rasheed. "Why WHO Suspended Chloroquine Clinical Trial for COVID-19 Treatment." *Vanguard Nigeria*, May 29, 2020. https://www.vanguardngr.com/2020/05/why-who-suspended-chloroquine-clinical-trial-for-covid-19-treatment/.

Soludo, Charles C. "Can Africa Afford COVID-19 Lockdowns?" *Proshare*, April 24, 2020. https://www.proshareng.com/news/NIGERIA%20ECONOMY/Can-Africa-Afford-COVID-19-Lockdowns----Chukwuma-Soludo/50636#.

Stumpf, Samuel E., and James Fieser. *Philosophy: History and Problems*, 6th edition. New York: McGraw-Hill, 2003.

United Nations Development Programme. *Human Development Indicators and Indices: 2018 Statistical Update Team*. New York: United Nations Development Programme, 2018.

Vangard Nigeria. "How COVID-19 Can Help Nigeria Rethink Education Policy – AACS." May 27, 2020. https://www.vanguardngr.com/2020/05/how-covid-19-can-help-nigeria-rethink-education-policy-%E2%80%95-aacs/.

Vanguard Nigeria. "Five Things to Know about the AstraZeneca/Oxford Vaccine." March 11, 2020. https://www.vanguardngr.com/2021/03/five-things-to-know-about-the-astrazeneca-oxford-vaccine/.

Vanguard Nigeria. "UK Says Confident in COVID-19 Vaccines as S. Africa Pauses AstraZeneca Rollout." February 8, 2021. https://www.vanguardngr.com/2021/02/uk-says-confident-in-covid-19-vaccines-as-s-africa-pauses-astrazeneca-rollout/.

Wouters, Olivers J., Kenneth C. Shalden, Maximilian Salcher-Konrad, Andrew J. Pollard, Heidi J. Larson, Yot Teerwattananon, and Mark Jit. "Challenges in Ensuring Global Access to COVID-19 Vaccines: Production, Affordability, Allocation, and Deployment." *Health Policy* 397, no. 10278 (2021): 1023-1034.

Wu, Joseph, H., Stephen D. John, and Eli Y. Adashi. "Allocating Vaccines in a Pandemic: The Ethical Dimension." *The American Journal of Medicine* 133, no. 11 (2020): 1241-1242.

Should Skepticism Be Discredited?

Anthony Udoka Ezebuiro

University of Nigeria, Nigeria
E-mail address: anthony.ezebuiro@unn.edu.ng
ORCID iD: http://orcid.org/0000-0002-8231-5270

Obiora Anichebe

University of Nigeria, Nigeria
E-mail address: obiora.anichebe@unn.edu.ng
ORCID iD: https://orcid.org/0000-0002-4208-9360

Anthony Chimankpam Ojimba

University of Nigeria, Nigeria
E-mail address: anthony.ojimba@unn.edu.ng
ORCID iD: https://orcid.org/0000-0002-8188-8392

Abstract

In our day-to-day life and experiences, when one doubts or questions unusually, he is branded a skeptic and consequently resisted. Skeptics, over the years, are seen as people whose basic mood is that of doubt; those who deny absolutely that true knowledge is possible. Although this is not completely true of skepticism, the present work demonstrates, though arguably, that skepticism is more of a philosophical method of inquiry; an epistemological attitude towards knowledge but whose goal is indeed certainty, although it selects a serious doubt concerning all knowledge as the starting point of the inquiry into the possibility of true knowledge. It can rightly be said that the work displays the paradox of skepticism. The word 'paradox' originates from a Latin term paradoxum, which has a Greek association paradoxon, or paradoxos, signifying "conflicting with expectation." Thus, the word paradox signifies a tenet or proposition contrary to received opinions. It is a statement or sentiment that is seemingly contradictory or opposed to common sense and yet, perhaps true in fact. The need for this work is necessitated by the fact that in the present age, it has become no longer the case that the best way to certainty is only by accepting entirely all that one is told, especially when such comes from a sage or a tradition. Obviously, we live in a dispensation where almost every human situation challenges the human rational faculty hence the tendency to change facts and hang-on to lies generates serious fever in every thinking mind. The result of this work therefore is that imperatively, the work demands that whoever wants knowledge should proceed through doubt. The method through which this work arrives at this conclusion is the analytic process of discussion and presentation.

Keywords: *ataraxia; doubt; Nihilism; paradox; Skepticism; suspension*

I. Introduction

We live in a dispensation where almost every human situation challenges the human rational faculty. It is a world where the tendency to change facts and hang-on to lies generates serious fever in every thinking mind. Whether in religion, politics or socio-economic and cultural life, truth-telling has gradually become a thing of the past in the world that we live in. It has become a costly price for one to take a facelift what one's neighbor narrates, because what is solemnly handed down as truth may, after investigation, be an all-round package of distorted facts. It no longer sounds funny that the society in which we live has developed the habit of telling lies at all circumstances such that confidence is fading out in human interactions. Consequently, there comes the need to develop the skeptic's mind-set and attitude if we must live and interact happily with one another in the same society. Sextus Empiricus was aware of this situation years ago that he did not fail to sound the beauty and value of skepticism. Commenting on his view about skepticism, Samuel Stumpf has this to say:

> Skepticism originated in the hope of attaining mental peace or calmness. People have been disturbed by the contradiction of things and plagued by doubt as to which alternative they should believe. They were struck, however, by the different conceptions of truth different philosophers had proposed. They also noticed that people have discovered the truth (and these, the skeptics called dogmatists), those who confess they have not found it and also assert that it cannot be found (and this they also considered as dogmatic position), and finally those who persevere in the search for it. Unlike the first two, says Sextus, 'the Skeptics keep on searching.[1]

Based on these, skeptics thought that if they could, by investigation, determine truth from falsehood, they could then attain tranquility of mind. Skepticism, therefore, is not a denial of the possibility of finding truth, nor is it a denial of the basic facts of human experience. Rather, it is a continuous process of inquiry in which every explanation of experience is tested by a counter experience. The fundamental principle of skepticism, according to Sextus, is that to every proposition an equal proposition is opposed. It is a consequence of this principle, he says, that "we end

---

[1] Samuel Stumpf, *Philosophy: History & Problems* (New York: McGraw-Hill, 1994), 120.

by ceasing to dogmatize."[2]And this perhaps amounts to justification of belief according to Robert Audi.[3]

There is no gain-saying that skepticism is necessary if we must be freed from every form of dogmatism. To this extent, we must seek to deepen the skeptic's mind-set and attitude if we must conquer this ugly situation and get ourselves out of the mess it provides. Since the target of skepticism is to challenge the alleged grounds of accepted assumptions, in order to know whether the claims they make are indubitable or necessarily true, it follows then to say that skepticism is a method of inquiry; hence a skeptic is someone who is unsatisfied with what is given and still is looking for truth.

It is unfortunate to observe that many today have capitalized on the fact that skeptics deny what appears acceptable to others to say that skeptics are those whose basic mood is that of doubt. And following that, they call skeptics *doubting Thomases*. Also coupled with the concern that skeptics doubt and question 'extraordinarily' people say they must be resisted. But the truth remains that skeptics were far from denying everything including the evident of sense perception. Like Sextus would say, those who say skeptics deny appearances "seem to me to be unacquainted with the statements of our school."[4]This means that the skeptics did not question appearances but only the account given of appearances.

Since we believe that skepticism is a philosophical enterprise that meant something rather different, namely, seekers or inquirers of certain knowledge, in this paper, our attempt would be to demonstrate how this is possible. To be sure, the ancient skeptics were doubters, but they doubted in order that they may know. For instance, they doubted that Plato and Aristotle had succeeded in discovering the truth about the world, and they had these same doubts about the Epicureans and Stoics. But for all their doubt, they were, nevertheless, seekers after a method for achieving a tranquil life.[5]

It is based on this explanation therefore that skepticism, in our context, would demand that whoever wants true knowledge should proceed through doubt. Ordinarily, one would have thought that the best way to certainty is by accepting entirely all that one is told especially when it comes from a sage, or a tradition. But as a philosophical method of inquiry and epistemological attitude towards knowledge, skepticism has its goal as certainty, though it selects a serious universal doubt concerning all knowledge as the starting

---

[2] Ibid.*,* 120-121.

[3] Robert Audi, *Belief, Justification and Knowledge: An Introduction to Epistemology* (Belmont, CA: Wadsworth, 1988), 13.

[4] Ibid., 121-122.

[5] Ben Okwu Eboh, *Basic Issues in the Theory of Knowledge* (Nsukka: Fulladu Publishing Company, 1995), 16.

point of a theory of knowledge. History of philosophy is replete with the assumption that knowledge was possible and could be found. But skepticism questions such possibility and hence offers a methodology on how to behave without the criterion of truth.

II. Various Notions of Skepticism and Applications in History

Originally, in the history of philosophy, the word *skepticism* is understood as a Greek word, which is derived from the verb *skeptestai* and which means "to inquire," or "to investigate." It is also from the Greek noun *skeptikos*, which is equally concerned with investigation. However, in the history of philosophy, peoples' attitudes to skepticism depend more or less on their interpretation, understanding and applications of these basic words; hence there are variegated notions of skepticism and application in the history of philosophy.

What seems to be the first misconception of skepticism derived from misconstruing of the root meaning of the term 'skepticism' and these have led to some dangerous over-generalization where the skeptics are said to be denying almost everything. Obviously, the old Greek word, *skeptikoi* from which skeptics is derived, means something rather different namely, "seekers" or "inquirers" (of knowledge).[6]

A philosopher like Bittle would interpret the root word of skepticism to mean that "the mind cannot overcome doubt; that the human reason is not only perverted and diseased but is in itself fallacious, weak and unstable."[7] This for him means that the mind is incapable of attaining knowledge, i.e. real certitude in knowledge is impossible. Ben Okwu Eboh also thinks that

> What the sceptic is saying, in effect, is that the mind is incapable of attaining knowledge, that is, that real certitude in knowledge is impossible. In short, the sceptic holds that claims to knowledge are shaky because any supposed truth that is offered as a candidate for knowledge might conceivably be false. This is why, in the view of skeptics, the only logical and rational thing we have to do is to suspend our judgement always because of a real doubt as to the truth of our judgement.[8]

---

[6] Stumpf, 120.

[7] Celestine Nicholas Bittle, *Reality and the Mind* (Milwaukee: The Bruce Publishing Company, 1936), 26.

[8] Eboh, 16-17.

There is no gain saying that the above misconception is one of those ones which tend to ignore the vivid explanations of skepticism that the earliest skeptics themselves offered such as Pyrrho. Pyrrho was popular for his doctrine of *ataraxia*. *Ataraxia* is a Greek word which is literally translated as "imperturbability,""equanimity," or "tranquility." Although the word *ataraxia* originally meant "freedom from worry and anxiety," i.e., "a state of calmness of mind in the face of seemingly intractable disagreement," later application of the term by Epicurus and his group, and the Stoics made it acquired varied senses in accordance with one's philosophical theories. That is to say that the mental disturbance that prevented one from achieving *ataraxia* varied among the philosophers; hence each philosophy had a different understanding as to how to achieve *ataraxia*.[9]

The Pyrrhonian skeptics tried to avoid committing themselves on any and all questions, even as to whether their arguments were sound. However, for them, those who claim for themselves to judge the truth are bound to possess a criterion of truth. This criterion, then, either is without a judge's approval or has been approved. But if it is without approval, whence comes it that it is trustworthy? For no matter of dispute is to be trusted without judging. And, if it has been approved, that which approves it, in turn, either has been approved or has not been approved, and so on *ad infinitum*.[10]Skepticism for them therefore, was ability, or mental attitude, for opposing evidence both pro and con on any question about what was nonevident, so that one would suspend judgement on the question.[11] It was this state of mind that necessitated the state of *ataraxia*; a state of quietude, or unpertubedness in which the skeptic was no longer concerned or worried about matters beyond appearance.[12]

Ataraxia requires the suspension of judgment. Among the Pyrrhonists, *ataraxia* was necessary for bringing about *eudemonia* (happiness) for a person, representing life's ultimate purpose. Their method of achieving *ataraxia* was through achieving *epoche*.

Epoche is the suspension of judgement. And it is not the same as *ataraxia* although the latter is relationally induced or brought about by the former for the sake of *eudemonia*. The issue is that we are first brought to *epoche* and then to *ataraxia*.

For Sextus Empiricus, skepticism was not a denial of the possibility of finding truth, nor was it a denial of the basic facts of human experience.

---

[9] Stumpf, 120.

[10] Sextus Empiricus, *Against the Logicians*, trans. Robert Gregg Bury (London: W. Heinemann, 1935), 179.

[11] Richard Bett, *Sextus Empiricus: Against the Physicists* (Cambridge: Cambridge University Press, 2012), 221.

[12] Ibid., 121.

Instead, it was "the view that questions whether any of our beliefs can be supported by adequate or sufficient evidence."[13] This form of skepticism shows that skepticism is a continuous process of inquiry in which every explanation of experience is tested by a counter experience.

In the early hours of Platonic dialogues, Socrates was seen questioning the knowledge claims of others. And in the *Apology*, he stated that all he really knew was that he knew nothing. Obviously, this Socratic skepticism was not a complete denial of the possibility of true knowledge. Rather it was a method where the mind refrained from quick judgement or accepting anything which it was not too sure of.

Furthermore, Socrates' enemy, the Sophist Protogoras, also did contend that "man is the measure of all things, of what is, that it is, and what is not, it is not."[14] This thesis showed a kind of skeptical relativism. For it was taken that no views were ultimately true, but each was merely one man's opinion.

Gorgias wrote a book to prove that nothing exists, that even if anything were to exist, it would be impossible to know and communicate it. This does not still show that Gorgias totally denied the possibility of true knowledge. Unfortunately, many have misconstrued this to be that Gorgias totally denied the possibility of true knowledge. The truth is that Gorgias' form of skeptical nihilism questions the capacity of the human mind to comprehend the nature of reality. It looks like what Scott Aikin refers to as regress problem in epistemology which states that if one has good reasons to believe something, one must have good reason to hold those reasons are good. And for those reasons, one must have further reasons to hold that they are good, and so a regress of reasons looms.[15]

In Augustine's *Contra Academicos*, which was influenced by the works of Cicero and the Platonism of the Middle Academy, Augustine stated that "skepticism can be completely overcome only by revelation. And from this standpoint, philosophy was considered faith seeking understanding (fides quarens intellectu)."[16] This is really a Fideist voice, the type that will be heard from the Fideist of the Renaissance period.

Michael de Ayguem Montaigne is of the view that his own skepticism was a "New Skepticism." Stumpf states this thus: "Montaigne looked upon himself as an unpremeditated philosopher; one who was not confined

---

[13] Eboh, 16-17.

[14] Plato, *Theaetetus*, trans. Joe Sachs (Newburyport, MA: R. Pullins Co., 2004), 166d.

[15] Scott F. Aikin, *Epistemology and the Regress Problem* (New York: Routledge, 2011), 2.

[16] Juan Comesaña, and Peter Klein, "Skepticism," *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), ed. Edward N. Zalta, https://plato.stanford.edu/archives/win2019/entries/skepticism/.

intellectually to some rigid set of ideas within which his thoughts and life must be expressed."[17]

Furthermore, in his *Essays*, Montaigne confessed that even though he was one of those who normally condemned such stories concerning ghosts and prophesies, he has come to find out that "it is a foolish presumption to slight and condemn things as false because they do not appear to us as probable..."[18]

Consequently, in the follow-up work, *Apologie de Raimond*, Montaigne held that the reason why we should not condemn things as false just because they do not appear probable to us is that the criteria employed to determine standards of judgments are themselves open to question and doubt, unless God gives us some indubitable first principles and makes our faculties reliable.[19]

The truth Montaigne is saying is that unaided by divine grace, all of man's achievements even those of the most recent sciences become dubious. But contending that personal and cultural sentiments influence people's judgments and that the senses are unreliable, Montaigne further suggests that we should judge with moderation, reverence, and prudence and with greater acknowledgement of our ignorance and infirmity compared with the infinite power of nature.[20]

In the same vein, Francisco Sanches in his book *Quad Nihil Scitus* in 1581 used classical arguments to doubt science in Aristotelian sense, arguing that giving necessary reasons for causes would lead to infinite regress because true knowledge of the behavior of nature cannot be attained.[21]

The most fundamental skepticism in the modern time was launched by David Hume between 1711 and 1776. Before Hume was René Descartes (who published his *Meditationes* in 1641), and was known for his methodic skepticism or methodic doubt. Descartes' doubt had the chief aim of providing rules for clear and orderly thinking. It was an effort to help the mind overcome the deception of the senses. Earlier in his work, Descartes had lamented saying,

> Whatever I have up till now accepted as most true, I have acquired from the senses or through the senses. But from time to

---

[17] "Pyrrhonian Skepticism," in *The Cambridge Dictionary of Philosophy*, ed. Robert Audi, 738-741 (Cambridge: Cambridge University Press, 1999).

[18] Richard Popkin, *Philosophy Made Simple* (London: Made Simple Books, 1981), 168.

[19] Joseph Omoregbe, *Epistemology (Theory of Knowledge): A Systematic and Historical Study* (Lagos: Joja Educational Research and Publishers Ltd., 1991), 168.

[20] Stumpf, 120.

[21] *Encyclopedia of Philosophy*, ed. Paul Edwards (New York: Macmillan, 1967), 101.

> time, I have found out that the senses deceive. And it is prudence
> never to trust completely, those who have deceived us once.[22]

Consequently, Descartes formulated four methodic principles or rules which would always govern people's discussions and to help them minimize fallacies so that one would never misplace truth for falsehood or vice versa. The climax of Descartes' methodic doubt was the discovery of the indubitable truth: *Cogito ergo sum* – I think therefore I am (*Je pense dunc je suis*). This truth was the fertile ground upon which Descartes proved the existence of things, man and God included.[23]

One should recall that the British empiricists were generally known for their dictum: "*nihil est in intellectu quod prius non fuerit in sensu*" meaning that nothing is in the intellect that first was not in the senses. Through this dictum, they contended that experience is the origin of all knowledge. Hume was led to this form of Empiricist skepticism in the end, by his early faith in reason. He did not think that adherence to reason could lead the mind to any absolute truth. Hence, he taught that in nature, there existed no absolute principle derivable by reason upon which depends the meaning and the knowledge of reality.

For Hume, object in nature existed separately; the movement from what is (a matter of fact) to what we ought to do or required (the qualities we place on objects and actions) was a logical jump. Hence there was no necessary logical inference from what is, to what ought to be. There was no necessary connection between cause and effect. Thus,

> all our reasoning concerning causes and effects were derived from
> nothing but custom; and belief was more properly an act of the
> sensitive, than of the cognitive part of our natures.[24]

Further still, Hume stated that the principle of causality could neither be demonstrated nor known by intuition. The idea of cause for him, therefore, was derived from the principle of frequent association of things that generally go together. It was by this association that we knew that things were contingently caused by the other.

Hume was the most thorough-going skeptic among the empiricists that his sweeping doubts about causality, the self, substance and metaphysical knowledge, Kant says, woke him up from his dogmatic slumber.[25]

---

[22] René Descartes, *Discourse on Method and Meditations on First Philosophy*, trans. Donald A. Cress (Indianapolis: Hackett Publishing Company, 1999), 12.

[23] René Descartes, *The Philosophical Writings of Descartes*, ed. John Cottingham (Cambridge: Cambridge University Press, 1985), 13.

[24] Stumpf, 213.

[25] Immanuel Kant, *Prolegomena to Any Future Metaphysics*, trans. and ed. Gary Hatfield

Kant attempted to synthesize the rationalists' and empiricists' proposition into his synthetic a-priori knowledge. He posited space and time as the two a priori categories presupposed in knowing. Finally, Kant divided reality into phenomena and noumena, holding that while the phenomena (things-as-they-appear) are knowable, the noumena (things-in-themselves) are unknowable.[26]

The contemporary period further witnessed the linguistic skepticism of Fritz Mauthner, whose critique of language in *Analysis of Language* led to a total skepticism about the possibility of genuine language. For him, language was both social and individual, and showed only what linguistic conventions were used at a given time, and what features of experiences they named in various ways.[27]Each language, according to Mauthner, expressed a worldview (weltanschauung), and what was called language was always relative to this outlook.[28] This just looked like Ludwig Wittgenstein's language game theory,[29]which also got clearly spelled out in his other work: *On Certainty*.[30]

George Santayana was a naturalistic skeptic who in his book, *Skepticism and Animal Faith*, insisted that "nothing given, existed as it was given; all belief about what was given was open to question."[31]He wanted to carry skepticism even higher than Hume, hoping that when the full force of skepticism was realized, one could appreciate what was in fact absolutely indubitable.

Albert Camus was an existential skeptic influenced by the skepticism of Soren Kierkegaard, Leon Shestove and Frederick Nietzsche. Nietzsche's skepticism, regarding religion and objective values rejected the Fideist mentality of overcoming skeptical puzzles by Leap of Faith. Thus, Camus accepted Nietzsche's view of meaninglessness of the world because of the "Death of God." And being so skeptical about the possibility of metaphysical knowledge, like other skeptics, Camus contended that the human situation which involves a constant futile effort to achieve understanding and meaning in an unintelligible and meaningless world, was absurd.[32]

In summary, it is now made clear that skepticism, contrary to popular opinions, meant generally more than total denial of the possibility of true

---

(Cambridge: Cambridge University Press, 2004), 4:260.

[26] Ibid., 4:313.

[27] David Hume, *An Enquiry Concerning Human Understanding*, ed. Eric Steinberg (Cambridge, MA: Hackett Publishing Co., 1993), 7, 2, 59.

[28] David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge (Oxford: Clarendon Press, 1978), 55.

[29] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* (London: Routledge & Kegan Paul, 1981), 23.

[30] Ludwig Wittgenstein, *On Certainty*, ed. G. E. M. Anscombe, and G. H. von Wright, trans. Denis Paul (Oxford: Basil Blackwell, 1969), 2.

[31] *Encyclopedia of Philosophy*, 453.

[32] Ibid., 453.

knowledge. Instead, it meant more of inquiry, and doubt among the members of the Platonic Academy. It was true that Sextus Empiricus, Pyrrho and Montaigne conceived it as investigation; moderation; and suspension; and above all, as a rule of life, it still remains what Descartes saw it to be: a methodological attainment of certainty while Hume perceived it as a radical means of deconstruction and doubt.

The contemporary linguistic philosophers like Fritz Mauthner and Wittgenstein, on their side also saw skepticism as a tool of relativism where language was argued to have meaning only as it expressed a world-outlook or language game. Hence George Santayana could reason that it is a process of interpretation or Animal Faith. There is no doubt that Albert Camus was influenced by Nietzsche and thus was led to pessimism about the human situation while Russell moved from pessimism to Gnosticism.

The above explanations suggest that from antiquity, skepticism has never had a uniform practice although it stood for one and the same thing, which is suspension of judgement until all doubts have been cleared.

III. Categorization of Skepticism

There have been efforts to classify skepticism as 'absolute' or 'moderate skepticism.' This is because following the above explanations, some skeptics tended towards radical skepticism than the others; and so they should be called absolute skeptics and the other moderate skeptics.

a. Absolute Skepticism

Absolute skepticism is to be self-stultifying because it tends towards nihilism. The word *nihilism* is originally derived from the Latin word *nihil* which means "nothing."So, *nihilism* is the state of mind that doubts the existence of something or better put, nihilists are said to be those who doubt the reality of existence. In the *Webster International Dictionary*, *nihilism* is portrayed as

> a viewpoint that all traditional values and beliefs are unfounded and that all existence is consequently senseless and useless. In fact, it is a denial of intrinsic meaning and value of life...a doctrine that no reality exists.[33]

The nihilists were said to deny the relevance of traditional values like laws and customs; hence they argued that such values – as natural law – were unfounded. It was based on this that some skeptics were branded nihilists.

---

[33] Ibid., 453-454.

But the question one needs to ask at this point is this: if there were skeptics who truly denied that there were intrinsic values and meaning to life when they presented their arguments as meaningful, does that truly make them nihilists? The answer to this question is a capital "No," because nihilism is not defensible in the face of their denial of intrinsic value and meaning to life because to say that existence is senseless implies that their arguments were also senseless in so far as they were part of existence. In that case, there were no real nihilists. Instead, there were simply skeptics in the strictest sense of it.

Again, we should note that different nihilist positions existed. While some held variously those human values were baseless, others held that life was meaningless, and still there were those who believed that knowledge was impossible or that some set of entities did not exist. In whichever positions or forms there were, the same question and answer still applied: If there were skeptics who truly denied that there were intrinsic values and meaning to life when they presented their arguments as meaningful, does that truly make them nihilists? The answer again remains "No," because nihilism could not be defensible in the face of their denial of intrinsic value and meaning to life when they maintained that existence was senseless. And this would have implied that their arguments were also senseless in so far as they were part of existence.

Now the fact that radical skeptics like Nietzsche and Russell were normally quite notorious could not still account for the reason why many would mistakenly identify their skepticism with nihilism, i.e., as extreme position. Nietzsche's crisis of nihilism derived from two central concepts: the destruction of higher values and the opposition to the affirmation of life. His writing, which, according to Lawrence J. Hatab[34] contained significant references to nihilism,[35] issued a radical attack on traditional belief system, and often echoed many of the pronouncements of nihilism. Yet this would not make Nietzsche to be frequently taken to be a nihilist. Instead, he was a skeptic in the strictest sense of the term skepticism.

On this count, Richard Schacht[36] demonstrated that Nietzsche had a dual attitude towards nihilism. For him, the question of whether Nietzsche was a true nihilist must be answered in both ways of Yes and No; affirmatively, if nihilism meant a denial of traditional belief systems, and negatively if it meant the denial of any value, meaning or truth in the world. According to him, Nietzsche accepted a restricted form of nihilism that denied a realm of "true being" apart from this world and a transcendentally grounded system of

---

[34] Ibid.

[35] Stumpf, 214.

[36] Ibid., 215.

values. Also, Nietzsche saw complete nihilism (as here defined) as decadent, dangerous, and something to overcome. In other words, if Nietzsche saw complete nihilism as dangerous, decadent and as something to overcome, it means that he could not have been taken as a nihilist in that strict sense of the word. And since his version of nihilism was rooted in the Christian-moral tradition, it would be safe to say as Schacht did insist that Nietzsche cannot also be considered a nihilist in the strict sense of the term because even in his denial, he advocated for such doctrines as the will to power and eternal recurrence. And based on this, talking about absolute skepticism would amount to a mirage.

b. Moderate Skepticism

On the other hand, moderate skepticism was considered as constructive and served as both a philosophical methodology and epistemological attitude of doubts aiding knowledge. Also, moderate skepticism was thought as the moderate mood of doubt about some several, or single thing, but never everything. The moderate skeptics were said to be known by what they doubt and how long they doubted.

In Samuel Stumpf's *Introduction to Philosophy*, Sextus Empiricus was said to be a moderate skeptic; hence he argued that evident matters such as whether it was night or day raised no serious problems of knowledge. In this category were evident requirements for social and personal tranquility, for we knew that customs and laws bound societies together. But non-evident matters, as for example, whether the stuff of nature was made of atoms, some fiery substances, did raise some intellectual controversies.[37]

So, based on this, moderate skepticism was seen to be "partial," "sensory," "rational," and "methodic." It was also said to be relative. Under moderate skepticism were also ethical, religious, and other forms of skepticism, which restricted doubt to definite areas. But unlike what was called "nihilist skeptics," who "doubt almost everything," the moderate skeptic doubted only metaphysical knowledge since evident matters posed no puzzle. The empiricists and positivists were said to belong to this group since for them, opinions, statements and matters were to be doubted if and only if they were obscure.[38]

Many have categorized Descartes as a moderate skeptic because in his *Meditation on First Philosophy*, he had this to say: "It will not

---

[37] *Webster's Third New International Dictionary*, vol. 19 (Springfield, MA: Marriam Webster, 1987), 1528.

[38] John P. Dougherty, ed., *The Review of Metaphysics*, vol. 36 (New York: A.M.S., 1983), 846-847.

be necessary for me to show that all my opinions are false, which are nothing but illusions and dreams."[39]On another occasion, he added:

> Whatever I have up till now accepted as most true, I have acquired from the senses or through the senses. But from time to time, I have found out that the senses deceive. And it is prudence never to trust completely those who have deceived us once.[40]

But to affirm or deny that Descartes was truly a moderate skeptic would depend on how one is able to understand the fact that Descartes' philosophy was dominated by his personal quest for certainty. Although this was not a preoccupation peculiar or unique or him and/or his age, there were such traumatic transition periods in the history of understanding such that it became more obvious that old assumptions did not work any longer since they no longer fitted the experience of the world. So, it was at such time like this that philosophic mind as Descartes' were driven to critical reassessment of the very foundation of what he already knew.

So, Descartes' background as a geometrician really paved way for him in his search for the indubitable truth or certainty that he required. However, using the method of geometry to think about the world, Descartes found the foundation of such "self-evident" propositions upon which whole geometrical systems can be built. It was this "methodic" form of skepticism that led him to doubt everything – *de omnibus dubitandum;* suspending belief in the knowledge he learned from childhood. In his *First Meditation on the First Philosophy*,[41]he reiterated his firm doubt on all those things "which I allowed myself in youth to be persuaded without having inquired into their truth."[42]So Descartes' doubt was methodic; hence it served him as a deliberate strategy for proceeding toward certainty. In that case, and like the rest of others before him, the so-called absolute or radical skeptics – Descartes became a doubter not by nature, but by necessity. For what he really wanted was to be secure so he could stop doubting.

It was this methodic suspension of belief that really got him to the point where he could no longer doubt his existence. Then it became clear to him that what he couldn't doubt any longer was the fact that he was the same

---

[39] Lawrence J. Hatab, "Nietzsche, Nihilism and Meaning," *The Personalist Forum* 3, no. 2 (1987): 91-111.

[40] Omoregbe, *Epistemology*, 168

[41] René Descartes, "Meditations on First Philosophy," in *The Philosophical Works of Descartes*, vol. 1, ed. Elizabeth S. Haldane, and G. R. T. Ross (Cambridge: Cambridge University Press, 1911), 28.

[42] Stumpf, 120.

man doubting. And this realization led him to the knowledge of his own existence, for if he could doubt the existence of every other thing; he could not doubt his own existence, for he had to exist first before he could doubt. On this truth, Descartes became the author of his famous phrase in Western philosophy: *Cogito ergo sum*, or, originally, *Je pense, donc je sui* – "I think, therefore I am or (exist)!" This secure anchor, no doubt, became the basis of Descartes' philosophical system, and he proceeded to infer the rest of his "truths."

The lesson here is that both Nietzsche and Descartes would have differed in one way or the other, for they would have had their individual methods still within the skeptics' tradition, but this did not qualify one to become an absolute skeptic or nihilist and the other moderate skeptics. There is no reason to say either of them is a radical skeptics or moderate skeptic than the other. The truth is that both of them qualified as both moderate as well as absolute skeptics no matter what meaning we give to it. So there is no reason to see one as absolute or radical and the other as moderate. After all, nihilism would never have been defensible in the face of their denial of intrinsic value and meaning to life since to say that existence was senseless implied that their arguments were also senseless in so far as they were part of existence.

IV. Comments on the relationship between *Ataraxia* and/or '*Epoche*' and/or '*Aponia*'

Briefly, we shall comment on how *ataraxia* relates to *epoche* and *aponia*. There has been effort to equate *ataraxia* with the word *epoche* and *aponia*. This is wrong because they were not meant to be the same thing *ab initio*. While *ataraxia* is a Greek word literally translated as "imperturbability," "equanimity," or "tranquility," and which first appeared in the works of Pyrrho though subsequently used by Epicurus and the Stoics, it does not mean the same as *epoche* or *aponia*. *Ataraxia* refers to "freedom from worry and anxiety." In other words, it was "a state of calmness of mind in the face of seemingly intractable disagreement." Among the Pyrrhonists, *ataraxia* was necessary for bringing about *eudemonia* (happiness) for a person, representing life's ultimate purpose. The method of achieving *ataraxia* was through achieving *epoche*.

Epoche on the other hand, is the suspension of judgement according to Sextus Empiricus.[43] And it is not the same as *ataraxia*, although the latter is relationally induced or brought about by the former for the sake of *eudemonia*. The issue is that we are first brought to *epoche* and then to *ataraxia*.

---

[43] Jonathan Barnes, 'Introduction," in Sextus Empiricus, *Outlines of Skepticism*, trans. Julia Annas, and Jonathan Barnes (Cambridge: Cambridge University Press, 2000), xix ff.

For the Epicureans, the concept of *ataraxia* was highly valued because of how pleasure was understood as highest good. For them, those who achieved freedom from physical disturbance were in a state of *aponia*, that was understood as "the absence of physical pain." Therefore, the concept of ataraxia is thus far removed from the sense in which the Epicureans used the concept of '*aponia*' because those who achieved freedom from mental disturbance were in a state of *ataraxia*.[44]

This distinction is very important to our discussion because while *epoche* induces *ataraxia*, *ataraxia* is not the same as *aponia*; hence the "absence of physical pain" is not one and the same thing as "the absence of mental disturbance." Therefore, as *epoche* in Pyrrhonism it is indicated "a suspension of judgment or belief for the sake of inner peace, especially while faced with a precipice," the state of *ataraxia* was brought about by eschewing beliefs (dogma) about thoughts and perceptions;[45] hence the values of skepticism.

V. The values of skepticism

The values of skepticism are both theoretical and practical. Theoretically, the strength of skepticism lies not in whether it is tenable as a position but in the force of the arguments of its proposers against the claims of dogmatic philosophers. Popkin was said to have argued that without skepticism, probably we could not distinguish enthusiasm, prejudice, or superstition from serious or meaningful beliefs. Perceived in this direction, we can describe skepticism as an epistemological fiery furnace where opinions are purified like gold.

Again, Popkin was further said to have contended that skepticism was instrumental to the birth of the modern epistemology at the hands of Descartes who was referred to as a moderate and methodological skeptic. This point is made clearer, of course because while the metaphysical frame of the later rationalists like Leibniz and Spinoza was merely an advancement of Descartes', the all-important epistemological contributions of the British empiricists was a response thesis to Descartes. Kant admitted that Hume's skepticism woke him up from his dogmatic slumber. No wonder he is called the father of modern philosophy.

On the practical level, what strikes the mind immediately is the classical skepticism of Michel Eyquem de Montaigne who lived from 1553 to 1592 in France. For Montaigne, skepticism neither meant pessimism in all things as a mood, nor license as a rule to do anything one wants. Instead, it was a

---

[44] Frederich Nietzsche, *The Gay Science*, trans. Walter Kaufman (New York: Modern Library, 1968), 67; Richard Schacht, "Nietzsche and Nihilism," in *Nietzsche: A Collection of Critical Essays*, ed. Robert Solomon (Garden City, NY: Anchor Press, 1973), 165.

[45] Stumpf, 121.

source for a positive affirmation of all the facets of human life. That was why he advised people to start their philosophy of life by reflecting upon matters close at hand; such that, a good place to begin would be one's own personal experience, given that "every man carries within himself, the whole conditions of humanity."[46] For this reason, Montaigne felt that whatever proved useful to himself might also serve useful to someone else.

This frame of mind reflects Kant's categorical imperative: "Act only on that maxim whereby thou canst at the same time will that it should become a universal law."[47] This maxim, of course, points to the Golden Rule: "Do unto others as you would like others to do to you." And in order to live up to this rule, Montaigne considered "contentment" as basic in life. Contentment can only be achieved through mental tranquility, but mental tranquility itself is achieved, according to Montaigne, when people concern themselves with existential phenomena; leaving out metaphysical problems to wane and die on their own.

However, Montaigne regrettably pointed out that the saddest spectacle of all is to find people formulating final answers on questions that are far too subtle and variable for such a treatment. The final folly of this attempt to capture the perfect and permanent truth is the mind of fanaticism and dogmatism. By the above lamentation, Montaigne attacked both the system-building philosophers who claim to be the unriddlers of the universe, and also the religious fanatics who caused wars and fierce religious persecution in the bid to perpetrate one kind of absolutive law or the other.

For those who could perpetrate any kind of evil to humanity under any guise, Montaigne blamed such cruelty as fanaticism caused by lack of inner peace. He then believed genuinely that a mood of constructive skepticism could prevent such an outburst of cruelty, because, "in the true skepticism, human energies could be directed toward manageable subjects and purposes."[48] According to Stumpf, Montaigne "adopted as his own, the central insight of classical skepticism, using this formula: 'I stop-I examine-I take for my guide the ways of the world and the experience of the senses.'"[49] The above principle looks like the Socratic injunction, "Man know thy self, for an unexamined life is not worth living."

Another great figure is Socrates. The skepticism of Socrates was visible in the way he engages his listener to argumentation. Unlike the Sophists who tried to show that truth or knowledge was impossible, Socrates accepted the

---

[46] Descartes, *The Philosophical Works*,12.

[47] Immanuel Kant, *Groundwork for the Metaphysics of Morals*, trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 2006), 4:421.

[48] Ibid.,12.

[49] Descartes,*Meditations*, 23.

possibility of truth and tried to link knowing and doing. For him, knowledge was virtue, and ignorance was the cause of vice.

Socrates' engagement in the "dialectic" was never for end destructive of truth nor to develop pragmatic skills among lawyers and politicians, but to achieve concepts of truth and goodness. His clash with the Athenian government on account of being a "corrupter of the youth," and for which he paid with his life, got him the reputation of "an intellectual dealing in paradoxes and, worse still, of thinking freely on matters about which many Athenians believed that discussions should be closed."[50]He was regarded a true skeptic; hence he taught the youth to live authentic lives as he did.[51]

VI. Should Skepticism then be discredited?

Those who misunderstood the meaning and scope of skepticism thought it was opposed to knowledge hence it should be resisted or discredited. But contrary to them and from our discussion so far, skepticism is supportive of knowledge. Wittgenstein once thought he had detested skepticism without knowing he soaked himself deeply in constructive skepticism although he would still not like himself to be identified as a skeptic. This truth is contained in the work of Garfield when he admitted that Wittgenstein, of course, frequently denied that he was a skeptic. He writes: "Skepticism is not irrefutable, but obvious nonsense..."[52] But I would argue that the position Wittgenstein denotes by "skepticism" is what I am calling 'nihilism.'[53]In that sense, both skepticism and nihilism meant the same thing for Garfield.[54]

However, the type of response Wittgenstein repeatedly offered to the skeptical problems posed by nihilistic arguments was characteristically skeptical. The point is that one needs to be skeptical to doubt the certainty of skeptical arguments. Hence philosophers as Wittgenstein and others who put up healthy arguments against skepticism were simply being truly skeptical.

Obviously, the skeptics contributed a great deal to the development of epistemology in Western philosophy by challenging the claim to know and the basis of such knowledge. The critical and sometimes devastating challenges of the skeptics spurred the epistemologists on to continually re-examine the nature, the basis and the justification of knowledge.

---

[50] Jane Friedman, "Why Suspend Judging?" *Nous* 51, no. 2 (2017): 302-326.

[51] Jane Friedman, "Suspended Judgment," *Philosophical Studies* 162, no. 2 (2013): 165-181.

[52] Stumpf, 120-121.

[53] Ibid., 120.

[54] Jay L. Garfield, "Epoche and Śūnyatā: Skepticism East and West," *Philosophy East and West* 40, no. 3 (1990): 304.

Epistemologists had to do this in order to find answers to the challenges of the skeptics, and to refute them.

One of them was Goodman who criticized skepticism on the thinking that it obstructed knowledge because of its application of suspension therapy. Indeed, he writes:

> Like the positivist and the radical empiricist, he (the skeptic) is barred by his own principle from going beyond phenomena at all by way of interpretation. The result is that none of the conundrums or antinomies which arise in experience itself or our 'natural' responses to it can be confronted by him in any way...The net out-come of the skeptic's perfection of his critical capabilities is their complete suspension.[55]

This is a clear criticism but what he did not realize was the fact that while "suspension therapy" seem apparently to be negative, it gave room for investigation and convictions through which inviolable certainty could be attempted on something (if probable), instead of hastily condemning that out of ignorance or dogmatically accepting it out of myopia. It is therefore wrong to accuse the skeptics as unable to resolve most of the questions they generated; after all, skeptics did not think that in philosophy, questions were more important than answers.

The antimonies which Sextus Empiricus enjoined them to formulate were not meant to be resolved since that was the best way to show the dogmatists that they may not have found the complete truth as they claim. It is not also true that the skeptics were intellectually redundant and inactive in philosophic enterprise as infants, unphilosophic adults or common men. Skeptics were active men with strong intellectual and philosophic mind.[56]

VII. Conclusion

Man is a being constantly in search of true knowledge. Skepticism afforded man that single opportunity to sift knowledge before consuming. Hence skepticism was both a philosophical method and an epistemological attitude towards knowledge. No doubts, there were various skeptics in history. While some skeptics were seen as extreme, others were regarded as moderate. In whichever way or form one found himself, both extreme skepticism and moderate skepticism were one and the same. They were led by one single passion: investigation and inquiry into the truth of things before consuming.

---

[55] Stumpf, 215.

[56] Dougherty, 846-847.

Whether some doubts of the skeptics were directed to almost everything, including knowledge and existence, even to the point of denying them possible existence or directed only to the possibility of arriving at indubitable truth, skepticism remains a method of inquiry towards certain knowledge. The moderate skeptics said things existed and knowledge was possible, but the problem lied with discovering a reliable criterion of indubitable certainty. The so-called radical skeptics believed this too.

In the final discussion of this work therefore, it is now clear that skepticism should not be discredited. Instead, it should be applauded and accepted for what it is. The doctrine of skepticism demands that whoever wants certain knowledge should and must proceed through doubt. It is no longer the case ordinarily, that one would accept entirely all that one is told simply because it comes from a sage, or a tradition. On this note, the paper concludes that the best way to certainty or rather whoever wants knowledge should proceed through doubt.

References

Aikin, Scott F. *Epistemology and the Regress Problem*. New York: Routledge, 2011.

Annas, Julia, and Jonathan Barnes. *Pyrrhonian Skepticism: Sextus, Empiricus. Outlines of Skepticism*. Cambridge: Cambridge University Press, 2000.

Audi, Robert. *Belief, Justification and Knowledge: An Introduction to Epistemology*. Belmont, CA: Wadsworth, 1988.

Bittle, Celestine Nicholas. *Reality and the Mind*. Milwaukee: The Bruce Publishing Company, 1936.

Comesaña, Juan, and Peter Klein. "Skepticism." *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition). Edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2019/entries/skepticism/.

Descartes, René. "Meditations on First Philosophy." In *The Philosophical Works of Descartes*, vol. 1. Edited by Elizabeth S. Haldane, and G. R. T. Ross. Cambridge: Cambridge University Press, 1911.

Descartes, René. *Discourse on Method and Meditations on First Philosophy*. Translated by Donald A. Cress. Indianapolis: Hackett Publishing Company, 1999.

Descartes, René. *The Philosophical Writings of Descartes*. Edited by John Cottingham. Cambridge: Cambridge University Press, 1985.

Dougherty, John P., ed. *The Review of Metaphysics*, vol. 36. New York: A.M.S., 1983.

Eboh, Ben Okwu. *Basic Issues in the Theory of Knowledge*. Nsukka: Fulladu Publishing Company, 1995.

Encyclopedia of Philosophy. Edited by Paul Edwards. New York: Macmillan, 1967.

Friedman, Jane. "Suspended Judgment." *Philosophical Studies* 162, no. 2 (2013): 165-181.

Friedman, Jane. "Why Suspend Judging?" *Nous* 51, no. 2 (2017): 302-326.

Garfield, Jay L. "Epoche and Śūnyatā: Skepticism East and West."*Philosophy East and West* 40, no. 3 (1990): 285-307.

Hatab, Lawrence J. "Nietzsche, Nihilism and Meaning."*The Personalist Forum* 3, no. 2 (1987): 91-111.

Hume, David. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. Oxford: Clarendon Press, 1978.

Hume, David. *An Enquiry Concerning Human Understanding*. Edited by Eric Steinberg. Cambridge, MA: Hackett Publishing Co., 1993.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press, 2006.

Kant, Immanuel. *Prolegomena to Any Future Metaphysics*. Translated and edited by Gary Hatfield. Cambridge: Cambridge University Press, 2004.

Nietzsche, Frederich. *The Gay Science*. Translated by Walter Kaufman. New York: Modern Library, 1968.

Omoregbe, Joseph. *A Simplified History of Western Philosophy*, vol. 2: "Modern Philosophy." Lagos: Joja Educational Research and Publishers Ltd., 1991.

Omoregbe, Joseph. *Epistemology (Theory of Knowledge): A Systematic and Historical Study*. Lagos: Joja Educational Research and Publishers Ltd., 1991.

Plato. *Theaetetus*. Translated by Joe Sachs. Newburyport, MA: R. Pullins Co., 2004.

Popkin, Henry Richard. "Skepticism." *Encyclopedia of Philosophy*, vol. 7, 44-46. London: Made Simple Books, 1967.

Popkin, Richard Henry. *Philosophy Made Simple*. London: Made Simple Books, 1981.

Richard, Bett. *Sextus Empiricus: Against the Physicists*. Translated by R. G. Bury. Cambridge, Cambridge University Press, 2012.

Schacht, Richard. "Nietzsche and Nihilism." In *Nietzsche: A Collection of Critical Essays*, edited by Robert C. Solomon. Garden City, NY: Anchor Press, 1973.

Sextus Empiricus. *Against the Logicians*. Translated by Robert Gregg Bury. London: W. Heinemann, 1935.

Stumpf, Samuel. *Philosophy: History & Problems*. New York: McGraw-Hill, 1994.

Webster's Third New International Dictionary. Springfield, MA: Marriam Webster, 1987.

Wittgenstein, Ludwig. *On Certainty*. Edited by G. E. M. Anscombe, and G. H. von Wright. Translated by Denis Paul. Oxford: Basil Blackwell, 1969.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul, 1981.

No God, no God's Eye: A Quasi-Putnamian Argument for Monotheism

Åke Gafvelin

University of Cambridge, United Kingdom
E-mail address: ag2120@cam.ac.uk
ORCID iD: https://orcid.org/0000-0001-7564-3201

Abstract

In the first part of this paper, I present Hilary Putnam's famous BIV-argument against metaphysical realism and find that it relies on the rejection of a noetic-ray theory of reference (NTR). In the second part, I demonstrate two ways in which a metaphysical realist could save the NTR, and I develop Plantinga's claim that metaphysical realists can only rid of Putnam's concerns by adopting the thesis that the objects of our conceptual schemes roughly correspond to the furniture of reality. Thirdly, I argue that naturalism fails to explain such a metaphysically anthropocentric correspondence, and that monotheism is the only candidate that does so successfully. In the last part, I show that metaphysical realism in fact holds. The line of argumentation is two-fold: pragmatic and theoretical. If metaphysical realism does not hold, then normative considerations must guide theory-choice. But fundamental non-verbal normative disputes are not possible if metaphysical realism is false. Hence, there can be no non-equivocal counterarguments to the claim that metaphysical realism should be adopted if it is false. This amounts to a normative consideration in favour of metaphysical realism. Secondly, I employ a reversed Putnamian BIV-argument to show that metaphysical realism is true: if metaphysical realism is false, we cannot assert that it is false, as its denial is only possible from God's point of view. But we can assert that it is false. Hence, it is true. The conclusion, that God exists, is surely apt to generate the sort of incredulous stare that any metaphysically ambitious armchair argument is subject to. The fundamental point of the paper, however, is that no such stare has any epistemological status if God does not exist. No God, no God's eye.

Keywords: *antirealism; metaphysical realism; Plantinga; philosophical theism; Putnam*

I. The Issue with Metaphysical Realism

Metaphysical realism is the thesis that "the world consists of some fixed totality of mind-independent objects" such that "there is exactly one true and complete description of 'the way the world is.'"[1] In *Reason, Truth and History,* Putnam characterises metaphysical realism as

---

[1] Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 49.

committed to the following three theses[2]:

   1. *Independence*: The world is (largely) made up of objects that
   are mind-, language-, and theory-independent.
   2. *Correspondence*: Truth involves some sort of correspondence
   relation between words or thought-signs and external things
   and sets of things.
   3. *The Cartesianism Principle*: Even an ideal theory might be rad-
   ically false.

According to Putnam, these three principles presume a theory of reference
in which "occult rays – call them 'noetic rays' – connect words and thought-
signs to their referents."[3] Without a noetic-ray theory of reference (NTR),
Putnam argues, metaphysical realism would be impossible. As we shall see,
this follows from a line of reasoning that now is a philosophical classic:

   P1. If I am a brain in vat (BIV), I cannot assert/form the thought
   that I am a BIV.
   P2. I *can* assert/form the thought that I am a BIV.
   C. Therefore, by philosophical necessity, I am not a BIV.
   P3. If metaphysical realism is true, then I could be a BIV.
   C2. Therefore, metaphysical realism is false.

The justification for P1 goes as follows. Presumably, if we are BIVs, we have
never interacted with actual brains – only the *brain-looking* things com-
posed of electrical signals, created by the evil scientist. So how on earth
are we to refer to *real* brains – something we have never interacted with? It
seems that the BIV, in that case, would need an ability to 'magically' think
about objects it neither constructs nor interacts with. If we reject the NTR,
however, then no BIV could assert/form the thought that it is a brain in a
vat, and P1 is true.

   P2 is to be grasped by examining our own conceptual scheme. "'Ob-
jects' do not exist independently of conceptual schemes," Putnam writes,
rather, "we cut up the world into objects when we introduce one or another
scheme of description. Since the objects and the signs are alike internal to
the scheme of description, it is possible to say what matches what."[4] Only
if we believe in the NTR could we be radically wrong about what our words

---

[2] Tim Button, *The Limits of Realism* (Oxford: Oxford University Press, 2013), 7-10.

[3] Putnam, 51.

[4] Ibid., 52.

and thoughts refer to. Hence, if we reject the NTR, then we can be absolutely certain that we *can* form the thought that we are brains in vats.

The truth of P3 is less controversial. The *Cartesianism Principle* states that we could be radically wrong about nearly everything, and the BIV-case is just an instance of this larger scheme.

We hence see that the soundness of the argument (P1 and P2) turns on the NTR. If one postulates a noetic-ray reference relation, then a BIV *could* refer to things in "metaphysical reality," and neither being a BIV nor metaphysical realism would be a philosophical impossibility. That's great, you say, but why would you believe in the NTR? Apart from the seeming queerness of a noetic-ray, we have the *semantic worry* that the ray, granted that it exists, might refer uncontrollably. If metaphysical realism is true, how is it that we can talk about any one thing rather than any other? Without a proper theory of reference-fixation, metaphysical realists, in a very literal way, have absolutely no clue what they are talking about.

II. Fixing Reference

If we examine our own behaviour and mental content in isolation, reference is underdetermined. This follows from Quine's observation that a fully competent field linguist cannot determine, given a certain set of linguistic evidence, whether natives talk about rabbits when they say 'gavagai,' pointing to rabbits on grass, or, say, rabbits-on-grass.[5] Kripke's famous arguments in *Wittgenstein on Rules and Private Language* further establish that I cannot determine whether I have previously meant *quus*[6] or *plus* by 'plus' based on my own dispositions and mental states.[7] These indeterminacy problems fundamentally rest in the interplay between human finitude and the infinite nature of semantic content. "Rabbit" or "plus" have an *infinite* number of possible applications, whereas human beings only have a *finite* number of behaviours or mental dispositions. Hence, there are simply too few behavioural and mental facts about us to decide what we refer to.

So, if facts about *us* do not fix reference, perhaps there is something in the interplay between us and the world that does? Perhaps our non-linguistic *interaction* with rabbits precludes us from referring to gavagais (or similar

---

[5] See Willard Van Orman Quine, *The Pursuit of Truth* (Cambridge, MA: Harvard University Press, Revised Edition 1992), 31-37.

[6] The function quus (a,b) outputs plus (a,b) if a and b are both smaller than or equal to the largest number n previously used by S in an arithmetical computation. If a or b is larger than n, quus (a,b) outputs 5. Thus, there is no way, based on the previous behaviours of S, to tell which one S has used.

[7] Saul Kripke, *Wittgenstein on Rules and Private Language* (Cambridge, MA: Harvard University Press, 1982).

permutations). For abstract entities like plus, we would have to postulate an *intellectual* interaction of the sort that Plato imagined; an immediate grasp of the form of plus instead of quus. If that were the case, we wouldn't have to rely on facts about dispositions or behaviours to fix reference; simple interaction would do.

The first problem for the metaphysical realist here is that we seem to interact with gavagais (and countless other permutations) every time we interact with rabbits that happen to be on grass, making it impossible to decide whether we interact with gavagais or rabbits when we interact with rabbits on grass. If that is the case, reference is not fixed. A second, more pertinent issue, is that it looks like metaphysical realism has to go if interaction fixes reference. For then sceptical BIV-cases are impossible, as Putnam saw, and the Cartesianism Principle is false. Without further assumptions, non-linguistic interaction cannot do the job of fixing reference for the metaphysical realist. Here, we need to get our metaphysical hands dirty. Perhaps the world helps us a great deal in referring correctly?

This would be the case if the actual world is cut up in pre-existing objects that roughly match those of our conceptual schemes. Such a world would be inhabited only by rabbits, and literally no gavagais (or any similar permutations of similar objects); pluses, but no quuses. Let us call this the few-objects-solution. If the world is cut up in such a way, any *interaction* (causal or ostensive) with rabbits *would* fix reference, as I have literally not, at the same time, interacted with a gavagai (or any similar permutation). This theory can use interaction to fix reference without giving up the Cartesianism Principle. Radical skeptical scenarios are possible in worlds mostly consisting of objects that do *not* match those of our conceptual schemes, but we are (plausibly) not *actually* located in one of them.

A closely related alternative would be a Lewisian eliteness theory, on which the world *itself* connects our words and thought-signs with its objects.[8] The idea is that objects have more or less *elite* properties, and that it is easier to refer to an object the more elite properties it has. Elite properties are the most fundamental properties in reality; they are 'joint-carving,' as Lewis puts it. And we intuitively think that *rabbitness* carves reality at its joints to a higher degree than *gavagainess*; rabbits form a natural kind, we think, whereas gavagais make up an artificial kind, fabricated for use in thought experiments. So the solution to the first problem is that we refer to rabbits instead of gavagais because rabbitness is an elite property. The solution to the second problem is that elite objects are so easy to refer to that even a BIV could do it: the 'magnetic' pull from the elite objects outside the vat manages to

---

[8] David Lewis, "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61, no. 4 (1983): 343-377.

draw the noetic-ray to them in such a way that BIV's comes to be radically deceived. Thus, the Cartesianism Principle remains intact.

I will neither develop these theories in any more detail nor take a stance on which one is correct, but I suspect that the metaphysical realist must accept one of them or a combination of both. For either *we* fix reference in isolation, or the interplay between *us* and the *world* fixes reference, or the *world* fixes reference for us. The first alternative is implausible, the second needs the few-objects-solution in order to avoid indeterminacy, and the third needs to postulate elite properties. Indeed, Putnam introduces these very theories as metaphysically realist solutions to the problems he presents. The few-objects-solution should correspond to what he (pejoratively) calls "Medieval Essentialism," and eliteness theory would be what he (pejoratively) calls a theory of "Self-Identifying Objects."

If we nevertheless accept one of these theories, we would have to uphold a distinction between what we might call primary and secondary concepts. On the few-objects-solution, ‹rabbit› and ‹plus› are primary, because they correspond to actual objects in reality, and ‹gavagai› and ‹quus› would be secondary, as they are fictional linguistic constructions. On eliteness-theory, the primary ‹rabbit› and ‹plus› correspond to objects with many elite properties, whilst the secondary ‹gavagai› and ‹quus› refer to objects with very few elite properties. But can we uphold such a distinction? As Putnam notes, there would be an eerie symmetry between them;[9] we could define 'gavagai' relying on the primary concept ‹rabbit›, but we could also define 'rabbit' using secondary concepts. Assuming that the Oxford Dictionary gets the necessary and sufficient conditions of rabbits right, we define 'Gavagai' and 'Havagai' as follows:

> Gavagai = df. a gregarious burrowing plant-eating mammal, with long ears, long hind legs, and a short tail only existing *on* grass.
> Havagai = df. a gregarious burrowing plant-eating mammal, with long ears, long hind legs, and a short tail only existing *outside* of grass.

Thus, we are in a position to define rabbit:

> Rabbit = df. a gavagai or a havagai

Thus, given that we could define 'gavagai' and 'havagai' relying on the concept ‹rabbit›, but 'rabbit' relying on the concepts ‹gavagai› and ‹havagai›, how are we to determine which are primary? What if rabbits are secondary

---

[9] Putnam, 36-37.

linguistic constructions/less elite objects, and gavagai and havagai real/elite? If that were the case, the few-objects-solution or eliteness theory needs to explain why we do not quickly find ourselves in variants of the BIV-cases that Putnam wants to get rid of. Pose that I (for some reason) have only interacted with rabbits on grass. If gavagais and havagais are primary, then, 'rabbit' would refer to gavagais. But then all my current beliefs of the form 'rabbits *could* locate themselves on space-coordinate x' would be false when x is not on grass. Obviously, this is but one instance of a larger problem that could render nearly all of our beliefs false. Thus, the few-objects-theorist and the eliteness-theorist would need to answer the question:

> (i) Why is the world such that its primary objects roughly are those *we* think are primary?

The only answer to (i), as I can see it, would be to posit *metaphysical anthropocentrism*; the thesis that reality itself is carved out roughly along the lines that human beings carve it. If metaphysical anthropocentrism is true, then most of the objects we deem to be primary would be primary, and most objects we deem to be secondary would be secondary. (Note that metaphysical anthropocentrism does not entail that human beings are metaphysically *privileged* in any way; it could be that the order of the world just *happens* to coincide with the way we order things, or that the furniture of the world has been adapted to fit our schemes by a being vastly more metaphysically privileged than us. We will investigate these two possibilities in the coming part).

    Another way to phrase this view, close to Putnam's formulation, is that metaphysical realism requires that human beings potentially enjoy a God's eye point of view. Obviously, we *are* not omniscient, and we could in fact be radically wrong about everything, but we are *actually* set in a position so as to know a great deal about the objects of reality. This is why Putnam proclaims that the God's Eye point of view is the favourite point of view for the metaphysical realist. Otherwise, there is simply no way of fixing the "noetic ray" without at the same time allowing for extreme skepticism.

III. Metaphysical Realism with a Human Face?

Metaphysical realists must therefore be metaphysical anthropocentrists. In what follows, I argue that it is impossible to square metaphysical anthropocentrism with naturalism[10] and very easy to square it with some kind of monotheism. Therefore, unless a better alternative can be presented, which is doubtful, the metaphysical realist must ontologically commit to God.

---

[10] I will define *naturalism* as the thesis that there are no supernatural entities.

If our best scientific theories are correct, human beings have existed for about 200.000 years in a universe that came about 13.8 billion years ago. Further, human beings inhabit an extremely small slice of the universe, and could very well be but one member of a large set of intelligent species. Therefore, a naturalistic explanation of metaphysical anthropocentrism could not posit that the universe *itself* is carved in a way that fits human conceptual schemes; that would simply be absurd.

Instead, the naturalistic account must be that human beings have *evolved* to carve their conception of reality in line with reality itself, thereby answering (i). The problem here, however, is that human beings primarily have evolved to survive and reproduce. Thus, it simply does *not* matter whether one derives nourishment from a gavagai or a havagai or a rabbit, for as Quine has shown, these hypotheses are empirically indistinguishable (and thus, physically indistinguishable). For the same reason, we could not postulate a multiverse, where human beings, due to the anthropic principle, come to exist in a universe fine-tuned to our conceptual schemes. *Ceteris Paribus*, a universe inhabited by gavagais and havagais and a universe inhabited by rabbits are empirically indistinguishable. Therefore, human observers have the exact same observation-conditions in both, but only in one world would they carve reality as it is (assuming that they carve reality either along rabbit-lines or gavagai and havagai-lines). Hence, we lack evolutionary reason to think that our conceptual schemes match reality.

Moreover, we have positive reasons to think that they should *not* match reality on evolutionary grounds. In his recent work *The Case Against Reality*, MIT cognitive scientist Donald Hoffman presents and describes the *Fitness-Beats-Truth* theorem (FTB Theorem) in evolutionary game theory, according to which evolutionary strategies maximising fitness at the expense of correct representation always beat strategies accurately depicting reality. The conclusion is that any given perception almost certainly is non-veridical:

> Darwin's idea of natural selection entails the FBT Theorem, which in turn entails that the lexicon of our perceptions – including space, time, shape, hue, saturation, brightness, texture, taste, sound, smell, and motion – cannot describe reality as it is when no one looks. It's not simply that this or that perception is wrong. It's that none of our perceptions, being couched in this language, could possibly be right.[11]

Hoffman's conclusions are by no means uncontroversial, and evolutionary game theory is a young and emerging field. But I suspect that the FBT theo-

---

[11] Donald D. Hoffman, *The Case Against Reality: How Evolution Hid the Truth from our* Eyes (London: Allen Lane, 2019), 125.

rem points us in the right direction when constructing an evolutionary epis-
temology for metaphysics. If we have evolved primarily for survival, we have
reason to believe that our faculties do *not* mirror how reality carves at the
joints, as the cost of getting metaphysics correct outweighs any benefit as-
sociated with it.

Hence, naturalism seems like a no-go for an explanation of metaphysical
anthropocentrism. Instead, it looks like the metaphysical realist needs some-
thing along the lines of the Christian view, on which God creates the world in
an orderly fashion so that human beings, by virtue of their reason, can come
to know this order. As *John* 1:1 states: "In the beginning was the Word, and
the Word was with God, and the Word was God" – and *Genesis* 1:27: "God
created mankind in his own image." St. Augustine interprets these passages as
God taking unformed matter (nothing), molding it into determinate objects
(almost nothing) subordinated the Forms, so that human beings can interact
with them:

> For thou, O Lord, hadst made the world out of unformed mat-
> ter, and this thou didst make out of nothing and didst make it
> into almost nothing. From it thou hast then made these great
> things which we, the sons of men, marvel at.[12]

If something like St. Augustine's picture is right, metaphysical anthropocen-
trism has an explanation. The answer to (i) is that God carved out the world
in a way that roughly corresponds to how human beings come to carve it,
because God has a special relationship with us. This does not mean that we
know absolutely everything about the nature of reality, for we are limited
beings prone to epistemic error, and we can only get a grasp of The Good and
other privileged universals through their instantiations in discrete particulars
we meet in our sensory world. Nevertheless, it suffices to preclude BIV-cases
from holding in the actual world. And fortunately so, for such cases would
reduce metaphysical realism to ultimate absurdity, as Putnam rightly pointed
out.

The contours of this theistic answer to Putnam's critique of metaphysical
realism were already sketched by Plantinga in 1988:

> You might be inclined to accept (1) the Putnamian proposition
> that we do know that we are not brains in a vat, (2) the anti-Put-
> namian claim that metaphysical realism is true and antirealism a
> mere Kantian galimatias, and (3) the quasi-Putnamian proposi-

---

[12] Saint Augustine of Hippo, *Confessions and Enchiridion*, trans. Albert C. Outler (Philadelphia:
Westminster Press, 1955), 12:6; 209-10.

tion that if metaphysical realism is true and there is no such per-
son God who has created us and our world, adapting the former
to the latter, then we would not know that we are not brains in
a vat; if so, then you have a theistic argument.[13]

To my knowledge, no one has developed Plantinga's argument in any detail
except Daniel Bonevac,[14] who renders it deeply implausible. Because mean-
ing needs to be grounded in a "supernatural, infinite, eternal, necessary, ob-
jective, normative, and independent causal power," and only God has these
properties, we need to be ontologically committed to God; "if there were no
God, there would be no meaning," Bonevac concludes.[15] This is an interesting
argument, but the present issue is not whether there are meanings, but *which*
meanings our expressions carry. Even if we grant that God grounds meanings,
how does He know which one to assign a given expression? If there is some
fact in reality to settle it, which only God knows, then He is clearly not need-
ed. If there is *no* such fact, and God does guesswork, then the account clearly
relies on magic. Here, Bonevac[16] bites the bullet: "Any account of semantic
capacities must at some point resort to magic. And the best explanation we
have for that magic involves God." It confounds me how any magical phe-
nomena could have a *best explanation*. Does it then not seize to be magical?
I think we ought to reject magic and Bonevac's argument with it. An anthro-
pocentric world is to be postulated precisely because it is the only world in
which a metaphysically realist theory of reference does *not* need to rely on it.

Might there be non-monotheistic rivals that explain metaphysical an-
thropocentrism? I doubt it. Atheistic explanations within a naturalistic frame-
work would fail for the reasons outlined above. Hence, they might postulate
'non-theistic' *supernatural* forces/entities that carve reality and/or set human
beings in a relation to know it. But this is obviously at the expense of moving
too close to the God-idea; the account becomes one of theology rather than
atheology, thus undercutting itself. Hence, we are left with monotheism or
polytheism. However, the dilemma for polytheisms is that we either must
postulate alternate realities carved by different gods, in which case we have
an unacceptable relativism, or find ourselves governed by a plethora of ex-

---

[13] Alvin Plantinga, "Appendix: Two Dozen (or so) Theistic Arguments," in *Alvin Platinga*, ed.
Deane-Peter Baker (Cambridge: Cambridge University Press, 2006), 203-228.

[14] Daniel Bonevac, "(N) The Putnamian Argument, (O) The Argument from Reference, and (P)
The Kripke-Wittgenstein Argument from Plus and Quus," in *Two Dozen (or so) Arguments for
God*, eds. Jerry L. Walls, and Trent Dougherty, 214-234 (Oxford: Oxford University Press, 2018).

[15] Ibid., 228.

[16] Ibid., 227.

planatorily superfluous demi-gods that do not carve reality at its joints.[17] In neither case are we able to explain the comprehensibility of reality – indeed, that very notion finds its genealogical roots in monotheism:

> Modern science, from the time of Newton, and the founding of the Royal Society in London in 1660, assumes the existence of one world that reason could investigate. That stemmed from a theistic belief in the one God who had created it. Their belief that one mind permeated the universe gave early scientists in the modern age the confidence to assume that there was one rational structure built into the nature of things, and that one Reason had produced it. The fact they believed that humans were made in the image of the one God also gave them assurance that human rationality had the capability of unlocking, at least in part, the secrets of the physical universe. This gave answers to the question of why the physical world should behave a uniform way and why should it be accessible to human rationality.[18]

We conclude that metaphysical anthropocentrism relies on monotheism. Whether this monotheism conforms to any one of the multitudinous interpretations of Islam, Judaism, Deism or Christianity, however, is by no means settled. Such a fact, if we will ever come to know it, could only be established by work in philosophy and theology.

To summarise the argumentation thus forth. Part I stated that metaphysical realism needs a noetic-ray theory of reference, and that a noetic-ray theory of reference needs a theory of reference-fixation. Part II argued that the noetic-ray can be fixed only if the world has elite objects or contains few objects, and that both the eliteness-theory and the few-objects-solution require metaphysical anthropocentrism in order to avoid extreme scepticism. In this part, we stated that monotheism is the best, and probably the only, explanation for metaphysical anthropocentrism. Thus, the metaphysical realist must ontologically commit to God. No God, no God's eye.

IV. Interlude

---

[17] For example, pose that there is a set of gods who carve the world in different ways. If all gods are correct in their carvings, the world will fundamentally be in many different, contradictory ways, and an incoherent sort of relativism holds. So we might then privilege one scheme of description imposed on 'prime matter' by one god (or several gods who agree in their carving of it). But then it is unclear why we, equipped with Ockham's Razor, should believe in more than one God.

[18] Roger Trigg, *Monotheism and Religious Diversity* (Cambridge: Cambridge University Press. 2020), 15.

Before we move on, however, we must attend to Putnam's *Just-More-Theory manouvre* against metaphysically realist theories of reference-fixation. Model Theory shows that there are indefinite ways to make a theory true (and false). For example, the Löwenheim-Skolem Theorem can be used to show that one can make true all sentences in a physical theory by interpreting them as propositions about the natural numbers. Putnam has an informal proof in the Appendix of *Reason, Truth and History* showing that in all possible worlds "the cherry is in the tree" has the same truth conditions as "the cherry* is in the tree*," where 'cherry*' sometimes refers to cherries and sometimes to cats, and 'tree*' sometimes refers to trees and sometimes to mats. Thus, reference seems inscrutable, at least in relation to mere alethic considerations.

At this point, we might say (as we already have) that these problems are solved in a reality ordered by God, where queer entities like trees* and cherries* either aren't elite or do not exist, and interaction and/or magnetism fixes reference. The problem that Putnam raises, however, is that "Interaction and/or magnetism fixes reference in a reality ordered by God" is itself formulated in a language subject to permutation. Perhaps that very sentence is just stating a truth about natural numbers or cherries: there would be no empirically available way to know. And of course, we cannot stipulate that it refers to what we think it does, since this would just be to add *just more theory*, which in turn can be permuted.

Here, Button distinguishes two sorts of epistemic worries: Cartesian angst and Kantian angst.[19] Cartesian angst is the worry that we might have radically false beliefs about the things our statements refer to. Kantian angst is the worry that we might have radically false beliefs about *what* our statements refer to. But Kantian angst is literally an impossible situation to be in. If the worry is legitimate, one couldn't even describe it, as 'Kantian angst' might well refer to cherries or natural numbers. As there is no way of knowing whether one's theory of reference is right on metaphysical realism, and this warrants Kantian angst, which is incoherent, Button concludes that metaphysical realism itself is epistemically incoherent and ought to be discarded.

The way out of this is simple. The metaphysical realist should agree that we ought not believe in a philosophical position warranting Kantian Angst, because that would be epistemically incoherent. But she should affirm, in a Moorean fashion, that we ought to believe in Metaphysical Realism, because it is true. Therefore, Metaphysical Realism does not warrant Kantian angst — we could not be wholly deluded about the meaning and reference of our words, for that would make it impossible to rationally hold our position. This response, however, will require us to demonstrate the truth of metaphysical realism, and to that we attend in the following part.

---

[19] Button, 60.

V. Metaphysical Realism is True

I sense an incredulous stare. If we agree that God is what it takes to save metaphysical realism, why espouse it at all? It is sometimes said that one person's *modus ponens* is another person's *modus tollens*, so why not run with the following argument:

> 1. If metaphysical realism is true, then God exists.
> 2. God does not exist / 'God exists' is meaningless and thus not true.
> 3. Therefore, metaphysical realism is not true.

However, is such a *modus tollens* argument possible here? Recall the three theses of metaphysical realism. Independence states that the world is (largely) made up of objects that are mind-, language-, and theory-independent. Correspondence reads that truth involves some sort of correspondence relation between words or thought-signs and external things and sets of things, and the Cartesianism Principle is that even an ideal theory might be radically false. If we give up metaphysical realism, these theses would have to go,[20] and without them, what Putnam calls the *internalist perspective* would reign.[21] But I shall argue that this perspective is a philosophical impossibility. Therefore, metaphysical realism is true.

According to the internalist perspective,

> there is no God's Eye point of view that we can know or usefully imagine; there are only the various points of view of ac-

---

[20] Citing De Morgan, the anti-metaphysical realist could jettison only one of these principles. But any such position would be absurd. Consider *correspondence* and *independence*. Either one is left with a determinate set of mind-independent objects one cannot talk about (reject independence, keep correspondence), or one has to talk about a determinate set of mind-independent objects that do not exist (keep correspondence, reject independence). Consequently, one has to get rid of both if one is to get rid of one. Further, the Cartesianism Principle is equivalent to independence and correspondence. If one rejects correspondence and independence, then truth must be identified with some kind of coherence theory/ideal rational acceptability, in which case the Cartesianism Principle is false, and if one affirms correspondence and independence, then our statements and beliefs correspond to mind-independent reality and could be radically false.

[21] Like Putnam, I will assume a dichotomy between metaphysical realism and the internalist perspective. Button argues for a third position in between, but it is unclear what it exactly amounts to (see Button, 221). Further, his main reason for rejecting metaphysical realism is his disdain for magical theories of reference, but there is no reason to think that a metaphysically realist theory of reference needs magic if God exists. Hence, I will not entertain this third position in the current context.

tual persons reflecting various interests and purposes that their descriptions and theories subserve ('Coherence theory of truth;' 'Non-realism;' 'Verificationism;' 'Pluralism;' 'Pragmatism;' are all terms that have been applied to the internalist perspective...).[22]

Thus, the internalist rejects Independence, as the structure of the world is not something "out there" for us to discover, Correspondence, as truth must be identified with some kind of ideal rational acceptability/coherence, and the Cartesianism Principle, as we cannot be deluded about everything.

However, if we jettison these principles, it clearly seems impossible to *assert* that the internalist perspective is true. In doing so, the internalist is either stating a truth immanent to her own conceptual scheme, in which case she is making no more than a testimony of her own ideology, or talking about ultimate reality, in which case she is taking on God's point of view. In the first case, one seems to lack any reason to listen, and in the second case, one is no longer talking to an internalist.

Here, we could argue that it is fully intelligible to assert internalism without having to take on God's point of view, deflating the strong requirements for assertion assumed above. Richard Rorty offers Wittgenstein and Heidegger as examples of good role models here:

> This is an awkward, but not impossible, position. Wittgenstein and Heidegger manage it fairly well. One reason they manage it as well as they do is that they do not think that when we say something we must necessarily be expressing a view about a subject. We might just be saying something – participating in a conversation rather than contributing to an inquiry. Perhaps saying things is not always saying how things are. Perhaps saying that is itself not a case of saying how things are. Both men suggest we see people as saying things, better or worse things, without seeing them as externalizing inner representations of reality.[23]

As I can see it, however, Rorty's proposal is genuinely unintelligible. In order to say something of philosophical value, one *must* say how things are: I cannot say that "reality is nothing but a linguistic construction, but not *actually*." You cannot even criticise the content of this paragraph unless you think that it *actually* fails to capture something about the nature of assertion. "Just saying something," in the Rortyan sense, is incoherent. The internalist *must*

---

[22] Putnam, 50.

[23] Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton, NJ: Princeton University Press, 1979), 385.

take on God's point of view in order to assert or even think that internalism is true.

But perhaps the internalist does not *need* to do so; perhaps she could just start phrasing herself in new, interesting and fruitful replacements of earlier, metaphysically realist frameworks without explicitly stating that internalism is true. This is the general strategy of internalist philosophers like Carnap, Wittgenstein, Quine, Rorty, Nietzsche, Heidegger, Foucault and Derrida. Philosophy, conceived as a *Mirror of Nature,* is simply set aside, and internalistic languages that promote theoretical and/or practical aims are adopted.

This project, however, must rely on some kind of normative framework. It forces us to ask the question why we *ought* to choose any one language over any other – in particular, why we *ought* to choose an internalist language over a metaphysically realist one (say, Sider's *ontologese*). Here, Putnam seeks to ground the answer to such value questions in an account of human cognitive flourishing:

> Bereft of the old realist idea of truth as 'correspondence' and of the positivist idea of justification as fixed by public 'criteria,' we are left with the necessity of seeing our search for better conceptions of rationality as an intentional human activity, which, like every activity that rises above habit and the mere following of inclination or obsession, is guided by our idea of the good.[24]

We *ought* to reject metaphysically realist languages, then, because pluralism and diversity – sought to be reduced/removed by the metaphysical realist – "is part of the ideal" of human cognitive flourishing.[25]

The problem for this strategy, however, is that there is significant disagreement as to what the ideal of human flourishing is and how it relates to the idea of the good. Plato would disagree that conceptual pluralism or diversity is good, yet part of the idea of human flourishing, and so would other metaphysical realists in his footnotes. However, it is hard to see how there could be non-verbal, fundamental disputes about these issues if metaphysical realism is false. Arguing which definition of 'good' or 'human flourishing' is right would be like arguing with a BIV about the definition of 'brain;' in the best scenario, we would mutually give true analyses of 'brain,' 'human flourishing' or 'good,' in our respective languages, but we would not disagree.

This incommensurability of normative ideals is explored in Eklund's *Choosing Normative Concepts*, which lets us imagine a *Moral Twin Earth*, where *bad guys* use a concept ‹ought*›, such that bad things, according to

---

24 Putnam, 137.

25 Ibid., 148.

our moral concepts, ought* to be done (and we use a concept ‹ought›, such that bad* things, according to the *bad guys*' concepts, ought to be done).[26] If there is no *privileged* concept corresponding to 'ought,' and thus no privileged concept of the good, both parties can state seemingly incompatible truths using different nearby concepts, and there could be no real dispute. And of course, we could not spell out the disagreement with a notion of conceptual privilege or correctness relying on normative concepts (say, by the view that the fundamental dispute concerns what concepts we *ought* to use), since that would be to rely on the very concepts we are trying to choose. It seems that it is only if there is a joint-carving, *elite* concept of the good – such as a platonic form of the good – that we could even hope to account for non-verbal debates about matters of the good. In that case, we would utter incompatible statements about *one* form of the good instead of uttering compatible statements true internally to our various normative concepts. But to posit that the world itself privileges certain normative concepts over others, say, through a platonic realm of forms that we can collectively describe, necessitates the truth of metaphysical realism and the rejection of the internalist perspective, and is thus not an admissible option for the internalist.

To get around platonism (or any other metaphysics on which there is a privileged notion related to 'good' or 'human flourishing'), the only way out for the internalist seems to be to concede that disputes about theory-choice will ultimately be merely verbal, but to deny the importance of this fact. This line is taken by Thomasson, who argues that philosophical disputes are really forms of "meta-linguistic negotiation." This sort of negotiation is to be understood in the realm of "pragmatics – the ways in which speakers use these utterances to reinforce or alter the norms for using the terms in question,"[27] rather than the realm of semantics, evading the necessity of some privileged platonic form supplying diverging concepts with a common referent. In linguistic negotiation, participants have the goal of influencing each other to adopt certain ontological vocabularies, rather than stating truths using privileged concepts. And there is something very plausible with Thomasson's proposal: it is not as if we would lay down flat if *bad guys* came to earth to declare that they ought* to eat us, just because 'earthlings ought* to be eaten' comes out true. A dispute about whether we ought (or ought*) to be eaten seems to persist even if the referents of these concepts do not coincide: therefore, a correct account of disagreement does not need to suppose that they do.

---

[26] Matti Eklund, Choosing Normative Concepts (Oxford: Oxford University Press, 2017), 26.

[27] Amie Thomasson, "Metaphysical Disputes and Metalinguistic Negotiation," *Analytic Philosophy* 58, no. 1 (2017): 13.

However, the consequences of this account of disagreement for the pursuit of philosophy ought not be underestimated. On this picture, there cannot be any meaningful distinction between persuasion and argumentation, since arguments exist to reinforce or alter behaviour rather than guide us towards the truth. Taken to its logical conclusion, Thomasson's pragmatic theory of disagreement means that discursive (or non-discursive) *violence* is the only arbiter in questions of ontology, since one's success in meta-linguistic negotiation is directly proportional to the effectiveness of one's methods of rhetorical manipulation. It might be the case that the most *instrumental* way for me to get someone to stop using the concept C is to speak kindly to them (because people are more amenable to act or speak differently if they do not feel forced to do so), but there is no *principled* distinction between (what on the surface looks like) peaceful discussion and outright violence. This violent element latent in pragmatism was effectively brought to light by Russell in his infamous criticism of James' theory of truth, but his observations apply equally well here:

> If there is a non-human truth, which one man may know, while another does not, there is a standard outside the disputants, to which, we may urge, the dispute ought to be submitted. If, on the contrary, the only way of discovering which of the disputants is in the right is to wait and see which of them is successful, there is no longer any principle except force by which the issue can be decided.[28]

Let us then negotiate. I do not think that it is useful or good to adopt this concept of disagreement, because I think the violence it entails should be rejected on moral grounds. Thus, I wish to influence the reader to let go of the pragmatic account of disagreement that Thomasson is proposing, and I hope that some of the considerations lifted above help in doing so. Since a non-pragmatic account of disagreement could not even begin to account for normative disputes on internal realism (given Eklund's Moral Twin Earth-considerations), and since Putnam is correct in viewing normative disputes as fundamental to ontology once both metaphysical realism and positivism is given up, we thus stand without a good theory of what it would even mean for the internalist to disagree with the metaphysical realist if we accept internalism. They cannot say that internalism is true of the world *as such,* because then they would no longer be internalists. Neither can they say that internalism *ought* to be adopted without either begging the question against metaphysically realist normative concepts (understood semantically), or practicing a sort of linguistic violence

---

[28] Bertrand Russell, *Philosophical Essays* (London: Longman's, Green and Co., 1910), 126.

(understood pragmatically) that we ought to reject. In neither case do we have reason to listen to what the internalists are saying.

Thus, I think we *ought* to reject internalism even *if* it is true (whatever that would mean), because it is an impracticable doctrine that is not even assertible: and I can further see no non-equivocal counterargument to this normative thesis that does not assume the truth of metaphysical realism, because of its assuming there to be a common idea of the good that we are attempting to analyse, or some normative fact of the matter that we are trying to mirror. As pragmatic considerations really are the only standards by which to adjudicate disputes about theory-choice if we get rid of metaphysical realism, metaphysical realism is to be accepted even if it turns out to be factually incorrect (whatever that would mean). Thus, internalism is incoherent, and metaphysical realism is true.

I admit that this will hardly be convincing to someone not impressed by pragmatic modes of reasoning in metaphysics; indeed, it probably only clarifies the absurdity inherent in the whole pragmatist project. Thankfully, however, metaphysical realism is not ultimately to be adopted on practical grounds – surprisingly, it is true by the very argumentative scheme Putnam used against it. We have seen that internalists cannot coherently *assert* or externalise their "inner representations of reality" without ceasing to be internalists – indeed, that there is no coherent way they can formulate their position. Hence, we note that their situation is identical to that of a brain-in-a-vat: as the BIV, bereft a noetic ray, cannot be located in a vat if it understands it is in one, so the philosopher cannot be located in a sort of internal reality if she can form the belief that she is. *Internal realism takes possible brains out of vats only by putting philosophers' brains back in them.* But we have a sound argument against being a BIV if internalism is true, formulated by Putnam himself. We proceed to use it to prove metaphysical realism and the falsity of the internalist perspective:

> P1*. If the internalist perspective holds, I cannot assert/form the thought that it is true.
> P2*. I *can* assert/form the thought that the internalist perspective is true.
> C1*. Therefore, the internalist perspective is false.
> P3*. If the internalist perspective is false, then metaphysical realism is true.
> C*. Therefore, metaphysical realism is true.

We conclude that metaphysical realism is true. But as we have seen, metaphysical realism is an intelligible position only if some kind of monotheism holds. Therefore, God exists.

References

Bonevac, Daniel. "(N) The Putnamian Argument, (O) The Argument from Reference, and (P) The Kripke-Wittgenstein Argument from Plus and Quus." In *Two Dozen (or so) Arguments for God*, edited by Jerry L. Walls, and Trent Dougherty, 214-234. Oxford: Oxford University Press, 2018.

Button, Tim. *The Limits of Realism*. Oxford: Oxford University Press, 2013.

Eklund, Matti. *Choosing Normative Concepts*. Oxford: Oxford University Press, 2017.

Hoffman, Donald D. *The Case Against Reality: How Evolution Hid the Truth from our Eyes*. London: Allen Lane, 2019.

Kripke, Saul. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press, 1982.

Lewis, David. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61, no. 4 (1983): 343-377.

Plantinga, Alvin. "Appendix: Two Dozen (or so) Theistic Arguments." In *Alvin Platinga*, edited by Deane-Peter Baker, 203-228. Cambridge: Cambridge University Press, 2006.

Putnam, Hilary. *Reason, Truth and History*. Cambridge: Cambridge University Press, 1981.

Quine, Willard Van Orman. *The Pursuit of Truth*. Cambridge, MA: Harvard University Press, 1992.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Russell, Bertrand. *Philosophical Essays*. London: Longman's, Green and Co, 1910.

Saint Augustine of Hippo. *Confessions and Enchiridion*. Translated by Albert C. Outler. Philadelphia: Westminster Press, 1955.

Thomasson, Amie. "Metaphysical Disputes and Metalinguistic Negotiation." *Analytic Philosophy* 58, no. 1 (2017): 1-28.

Trigg, Roger. *Monotheism and Religious Diversity*. Cambridge: Cambridge University Press, 2020.

History of the Concept of Similarity in Natural Sciences

Virginia J. Grigoriadou

National Technical University of Athens, Greece
E-mail address: virgi.grigoriadou@gmail.com
ORCID iD: http://orcid.org/0000-0002-0940-1922

Frank A. Coutelieris

University of Patras, Greece
E-mail address: fcoutelieris@upatras.gr
ORCID iD: http://orcid.org/0000-0002-5998-2235

Kostas Theologou

National Technical University of Athens, Greece
E-mail address: cstheol@central.ntua.gr
ORCID iD: https://orcid.org/0000-0002-3883-601X

Abstract

The concept of similarity has been discussed by many scientists and philosophers since ancient times. Thales of Miletus, Euclid, Aristotle, Galileo, Newton, Edgar Buckingham, and the modern philosopher of science Susan G. Sterrett are examples of intellectuals who perceived and examined the concept of similarity, while many scientists incorporated it in their scientific methodology. The wide range and variety of definitions of similarity could result in confusion regarding the meaning of the concept, the role the similarity mechanism plays in scientific methodology, and the identification of scientific fields to which similarity could be applied. The main aim of this paper was to enhance the understanding of the notion of similarity. To this end, we examined the historical evolution of the concept of similarity and the utilization of the mechanism of similarity in various eras of human intellectual activity, ranging from antiquity to the present day. In this context, the research hypothesis we investigated was the existence of specific and distinct stages of evolution within the long history of the concept of similarity in parallel with the evolution of scientific thought. A core question that motivated our work was when and under which conditions did the transition from the "technocratic" utilization of similarity (i.e., the use of similarity as a solution for practical problems) to its theoretical documentation and its conscious and systematic use as a significant experimental tool occurred. Another important question examined was whether there was a certain era that favored the development of the concept of similarity more than other historical periods. In order to address this hypothesis and respond to these questions, we sought to trace the evolution of conceptualizing and using similarity in different spatial and temporal contexts, formed by the corresponding historical, institutional, religious, and social conditions as well as the characteristics of the scientific methodology established during the period the similarity concept evolved.

Keywords: *similarity; similar systems; analogy; scientific models*

I. Introduction

The mechanism of similarity is widely used in modern scientific methodology that is employed during the design of natural science experiments. The concept of similarity is characterized by a long historical evolution, which unfolds in parallel with the historical evolution of scientific thought from antiquity to current years. A significant number of philosophers and scientists from different scientific fields have approached the concept of similarity, which resulted in the existence of a wide range of definitions of the notion of similarity. In philosophy, similarity is defined as the existence of a common, similar, or analogous property or attribute between two or more objects, while in geometry it is assigned as an equal or proportional dimension.[1] In physics, similarity is considered as the ratio of specific relationships of specific physical quantities of two or more physical systems.[2] In engineering, similarity is perceived as a mechanism that operates on the basis of a set of rules, laws, principles, or mathematical relationships that are employed by the experimental technique of analogue models during the process of selecting or constructing the model and during the process of extending the conclusions from the model to the phenomenon, object, or system of interest.[3] The common ground between these different approaches of the concept of similarity is detected in the attempt to define it based on the ratio concept.

Despite the range of approaches on the concept of similarity, the study of its historical evolution reveals that during its evolutionary stages similarity is mainly associated with the fields that we nowadays collectively refer to as the natural sciences. Natural sciences have played a significant role in understanding and defining the similarity and similar system concepts and in utilizing the mechanism of similarity as a technique of experimental methodology, especially after the 17th century. The idea of similar systems is firstly detected in Galileo's experiments, while the concept of the similarity of physical systems or bodies is firstly defined by Newton in the second book

---

[1] Susan G. Sterrett, "Similarity and Dimensional Analysis," in *Handbook of the Philosophy of Science*, *Volume 9: Philosophy of Technology and Engineering Sciences*, ed. Anthonie Meijers (Amsterdam: North Holland, 2010), 799-801; Susan G. Sterrett, "Physically Similar Systems: A History of the Concept," in *Springer Handbook of Model-Based Science*, eds. Lorenzo Magnani, and Tommaso Bertolotti (Cham: Springer International Publishing, 2017), 384-386.

[2] Sterrett, "Similarity and Dimensional Analysis," 800-801; Sterrett, "Physically Similar Systems," 380-384.

[3] Susan G. Sterrett, "Physical Models and Fundamental Laws: Using One Piece of the World to Tell About Another," *Mind & Society* 3, no. 1 (2002): 56-58; Susan G. Sterrett, "Models of Machines and Models of Phenomena," *International Studies in the Philosophy of Science* 20, no. 1 (2006): 69-80.

of Principia.[4] Since the beginning of the 17th century, many approaches on the concept of similarity have been recorded in the field of natural sciences.[5] At the same time, the use of the mechanism of similarity was expanding in the natural sciences and engineering. One theory that can justify the significantly extensive utilization of the mechanism of similarity in the field of natural sciences in comparison with other scientific fields is the theory of determinism, according to which everything that happens in the natural world is determined completely by previously existing causes, which necessarily lead to the same result.[6] In this context, utilizing the mechanism of similarity is more secure and effective in describing, explaining, and predicting natural phenomena than, for example, social phenomena.

Modern scientists do not exploit the mechanism of similarity by accident, unconsciously, or in an exclusively technocratic manner. On the contrary, they understand the meaning and the role of similarity in modern scientific methodology. One core question that gave rise to the present approach is the following: when, under what conditions, and how was the transition from utilizing similarity as an exclusively practical technique to its theoretical documentation and its conscious and systematic utilization as an important scientific methodological tool completed? Another question that motivated our research was whether there was a certain period that favored the development of the concept of similarity more than other periods. These two leading questions are directly related to the concern about perceiving and defining the evolution of conceptualizing and exploiting similarity as a practical technique before the advent of episteme and natural philosophy, but mainly as an experimental technique of natural sciences. The main purpose of this work was to enhance the understanding of the concept of similarity by identifying the stages of its development in correspondence with the evolutionary stages of intellectual activity.

Based on the assumption that the concept of similarity evolved alongside scientific thought and acquired its modern meaning within the scientific methodology of natural sciences over centuries, we supported that the concept of similarity went through five distinct stages of evolution. Initially, we discerned the Egyptian stage, which corresponds to a generalized way of the utilization of similarity, thereby enabling ancient Egyptians to accomplish various architectural, medicinal, and mathematical feats. The second stage dates back to the Classical era, the era of the genesis of episteme and

---

[4] Sterrett, "Physically Similar Systems: A History of the Concept," 381-387.

[5] Ibid., 381-387.

[6] John Earman, "Το Πρόβλημα του Ντετερμινισμού στις Φυσικές Επιστήμες," στο *Εισαγωγή στη Φιλοσοφία της Επιστήμης,* επιμ. Αριστείδης Μπαλτάς, μτφ. Πάνος Θεοδώρου, Κώστας Παγωνδιώτης, Γιώργος Φουρτούνης (Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 1998), 319-320.

natural philosophy, when the notion of similarity appeared in philosophy, mathematics, music, and geometry and acquired increased methodological importance. The third stage is during the Dark Ages, which is characterized by the absence of experimental techniques or mechanisms, such as similarity in scientific methodology. The next stage (16th-19th centuries) coincides with the emergence of modern science when the concept of similarity gained new importance; during this time, similarity was expressed as a methodological idea of similar systems, mainly by Galileo who was probably the first to perceive the idea of similar systems and use it extensively in his experimental methodology, but also by Newton who was the first to define the term similar systems. Finally, the fifth stage corresponds to the period ranging from the 19th century to the present day. During this period, the mechanism of similarity has been accepted as a formal methodological tool of natural sciences, and the concept of similar systems has been examined and defined by a significant number of modern scientists, with the contributions of Buckingham and Sterrett being highly important approaches. In this study, we argued that the transition from the "technocratic" utilization of similarity to its conscious utilization could be traced to the classical era stage. The transition to the systematic use of the concept of similarity as a significant experimental tool can be traced after Renaissance. Finally, we identified the period characterized by a conscious and systematized effort pertaining to the theoretical documentation of the concept of similarity and the expansion of its application to more scientific fields as starting after the 19th century. Although all stages were important for the evolution of the concept of similarity, some periods favored its development and the extension of its application in several scientific fields; such a period began after the scientific revolution, when the experimental method of the 17th century was introduced and the transition from natural philosophy to science was completed.

II. The origins of similarity in Ancient Egypt

Several historians of science suggest that the origins of science can be traced to ancient Egypt, after 3000 BC. The Egyptians occupied themselves systematically in the fields of mathematics, astronomy, and medicine, thereby laying the foundations for the subsequent development of these scientific fields.[7] Examples of the utilization of similarity are found in Egyptian geometry and medicine.

Egyptian geometry was primarily developed to solve practical geometrical problems. An interesting example is the construction of the pyramids of Giza,

---

[7] David Lindberg, *Οι Απαρχές της Δυτικής Επιστήμης* (Αθήνα: Πανεπιστημιακές Εκδόσεις Ε.Μ.Π., 2003), 19.

which leads to the following reasonable question: how did the Egyptians manage to construct pyramids similar in shape but different in size? Ancient Egyptians calculated the area of flat shapes, such as the triangle, and the volume of solids, such as the pyramid. To calculate the volume of a pyramid, they multiplied 1/3 of the base area by height.[8] Thus, it is reasonable to believe that when Egyptians were designing the pyramids, they performed mathematical calculations that allowed them to obtain geometric similarity between the different pyramids.

Another field in which ancient Egyptians used the technique of similarity was medicine. The Egyptians obtained significant achievements in the field of medicine, as evidenced by the papyruses of Ebers, Edwin Smith, and Hearst as well as the London Medical Papyrus.[9] In these papyruses, therapeutic methods, techniques, and pharmaceutical prescriptions for the treatment of illnesses, fractures, or wounds are categorized and described in detail.[10] In the Ebers Papyrus, prescriptions and medicines for various illnesses and hygiene tips are categorized in 110 columns.[11] The Edwin Smith Papyrus contains an extensive text of 48 paragraphs that describes and classifies wounds and fractures alongside with their respective treatments.[12] However, how did the Egyptian doctors compile these lists? The details on the human body and its function lead to the conclusion that this knowledge was obtained from the systematic collection and analysis of experimental data. The similarity of symptoms or medical incidents and trials of similar therapies contributed to the description, explanation, and prediction of diseases. Moreover, archaeologists believe that ancient Egyptian doctors used animals as analogue models of the human body. This belief is mainly based on wall paintings of monuments depicting doctors examining dead animals, and it is reinforced by the discovery of a large number of mummified animals in Sahara in 2018.[13] The most important source of knowledge for Ancient Egyptians was the mummification of human bodies. Studying the anatomy of bodies enabled Egyptian doctors to get to know the human body, its skeleton, and its organs. All these facts lead to the conclusion that the ancient Egyptian doctors relied heavily on similarity, both while studying the human body and when categorizing the existing knowledge about it.

---

[8] Ibid., 20; Thomas Heath, *A History of Greek Mathematics, Volume 1: From Thales to Euclid* (Oxford: Clarendon Press, 1921), 122-123.

[9] Lindberg, *Οι Απαρχές της Δυτικής Επιστήμης*, 26; John F. Nunn, A*ncient Egyptian Medicine,* (Oklahoma: University of Oklahoma Press, 2002), 24-41.

[10] Lindberg, *Οι Απαρχές της Δυτικής Επιστήμης*, 26; Nunn, *Ancient Egyptian Medicine*, 24-27.

[11] Ibid., 30-31.

[12] Ibid., 25-30.

[13] BBC, "Egypt Animal Mummies Showcased at Saqqara near Cairo," accessed January 17, 2020, https://www.bbc.com/news/world-middle-east-50531808.

We do not know to what extent the concept of similarity was defined in Egyptian science; however, by studying the achievements of ancient Egyptians we can conclude that similarity had been used systematically in some cases as a means of categorizing knowledge as well as of describing, explaining, and predicting the world; whether this happened consciously or not, it was primarily aimed at solving practical problems.

III. The concept of similarity in classical antiquity (490-323 BC)

The precursor of modern science was episteme, which was born during the classical era and derived from the ancient Greek word *επιστήμη* (ἐπίστασθαι‹ ἐπίσταμαι: know, understand, be acquainted with).[14] The first to introduce the term "episteme" was Plato, while this concept was later defined more elaborately by Aristotle. Plato contrasts episteme with doxa[15] and through his dialogues he presents episteme as a condition more valuable, harder to achieve than doxa, and never false on contrary to doxa.[16] According to many intellectuals, Plato's concept of episteme resembles the meaning of knowledge; according to others, it refers to the process of understanding. In Plato's dialogue *Theaetetus*, episteme is defined as a true doxa with a logos:

$$\text{ἔστιν οὖν ἐπιστήμη δόξα ἀληθής μετά λόγου,}$$

while in his *Republic*, Plato claims through Socrates that:

episteme's object is what is.[17]

Perceiving Plato's episteme as a process of understanding is probably a more substantial approach; however, if we accept this approach, we are faced with an important question: what is the possibility of disseminating this kind of knowledge and how stable and objective could it be? The approach of Plato's student Aristotle came to solve this problem. Aristotle characterized episteme as a deductively valid system grounded in necessary truths about natures or essences and he distinguished it from techne, a kind of practical knowledge relating to what we nowadays call technology. Overall, it could

---

[14] George Henry, "A Greek-English Lexicon," accessed July 5, 2020, http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.04.0057:entry=e)pisth/mh.

[15] doxa‹ *ancient greek* δόξα (= a perception or belief) ‹δοκέω/ δοκῶ (believe, think, imagine, guess, assume, decide).

[16] Jessica Moss, "Is Plato's Epistemology About Knowledge?" in *What the Ancients Offer to Contemporary Epistemology*, eds. Stephen Hetherington, and Nicholas D. Smith (Oxfordshire: Routledge, 2019), 1-6.

[17] Ibid., 1-6.

be argued that the purpose of episteme during the Classical era was to explain the world but not to change it. In this context, natural philosophy appeared. The purpose of natural philosophers of the Classical era was not to predict or control the natural world, but to understand, describe, and explain it. In this respect, natural philosophy was different from modern natural science.[18]

Important intellectuals of this era approached the notion of similarity, which acquired increased methodological importance. The roots of the notion of similarity are found in the Pythagorean philosophers, who discerned a relationship between observable phenomena and ratios.[19] They correlated certain musical phenomena with specific length ratios of a lyre string. According to the Pythagoreans, these proportions are equal to the proportions of prime numbers. This observation led them to the conclusion that all physical phenomena could be understood or described in terms of ratios.[20] The analogies found in the study of harmony appeared in other mathematical representations, such as the Tetraktys, a ten-point triangular arrangement consisting of four columns containing one, two, three, and four points, respectively.[21] Moreover, the Pythagoreans traced a relationship among the first four numbers, the sum of which is 10 (1 + 2 + 3 + 4 = 10). From these first four numbers (1, 2, 3, 4), it is possible to construct certain ratios, representing the relationship between two notes, which in music attribute the harmonious musical intervals that Pythagoras first defined in numerical terms.[22] Through a series of experiments, Pythagoras observed that when two strings have the same length, they have the same pitch and the interval between the notes is called a unison.[23] If the length of one string is one-half that of the other string, its pitch is much higher, but they still sound consonant when played together. This interval is represented by the mathematical ratio 2:1 and is called octave [diapason (French) ‹ diapason (Latin) ‹ ἡ διαπασῶν (Ancient Greek)].[24] If the length of one string is two-thirds that of the other, the strings still sound consonant when played together, and this interval is called a perfect fifth, represented by the ratio 3:2 [perfect fifth ‹ *diapente, sesquialterum* (Latin) ‹ διά πέντε or dioxea ‹ δι οξείαν (Ancient Greek)]. Another

---

[18] John Reeves, "The Science and Religion Dialogue as Natural Philosophy," *Metanexus*, accessed July 5, 2020, https://www.metanexus.net/science-and-religion-dialogue-natural-philosophy/.

[19] Sterrett, "Similarity and Dimensional Analysis," 799.

[20] Ibid., 799.

[21] Ibid., 799.

[22] Heath, *A History of Greek Mathematics*, 76-86.

[23] Stephanie J. Shaw, *W. E. B. Du Bois and the Souls of Black Folk* (North Carolina: The University of North Carolina Press), 135-136.

[24] Ibid., 136.

Pythagorean interval was 4:3, which is called the perfect fourth [*Diatessaron, sesquitertium* (Latin) ‹ *δια τεσσάρων* (Ancient Greek)].[25]

The Pythagoreans also believed that numbers are related to geometric shapes,[26] owing to the use of sequences of shapes that represent integers, which are created by a specific procedure. In this manner, the unit is related to the point, the dyad to the line, the trinity to the triangle, and the quadruple to the tetrahedron.[27] According to the Pythagorean theory, each integer has a graphical representation. The relationships of analogy between the sides of the shapes that form the sequence are also correlated with specific numbers. Thus, the study of geometrical similarity was initially related to integer relationships.[28] A typical example includes square numbers, such as 4, 9, and 16, the side ratios of which are 2:2, 3:3, and 4:4, respectively, which are all squares, therefore geometrically similar.[29]

The concept of similarity is first detected in geometry in the theorem of the similar triangles by Thales of Miletus, a Greek philosopher and mathematician. According to this theorem:

> Two triangles are equal when they have one side and the angles adjacent to it, equal.[30]

Thales traveled to Egypt and was trained in mathematics by Egyptian priests-mathematicians. Of particular interest is the story in which Thales was able to calculate the height of the pyramids based on their shadow. According to Hieronymus, a disciple of Aristotle, Thales observed the length of the shadow of the pyramids just at the time when the height of our shadow is equal to our real height. The story is presented slightly differently by Plutarch, who in his dialogue between Nikoxenos and Thales presents Nikoxenos to praise Thales for his achievement in calculating the length of the Egyptian pyramids based on the length of their shadow and the shadow of a bar that had entered into the ground.[31] According to this assumption, Thales used the ratio of the sides of two similar triangles and calculated the height of the pyramids from the length of their shadow and the shadow of the bar, thereby impressing the

---

[25] Ibid., 136.

[26] Sterrett, "Similarity and Dimensional Analysis," 799.

[27] Ibid.,799.

[28] Ibid.,799.

[29] Heath, *A History of Greek Mathematics*, 76-86.

[30] Encyclopaedia Britannica, "Thales of Miletus," accessed November 25, 2019, https://www.britannica.com/biography/Thales-of-Miletus.

[31] Heath, *A History of Greek Mathematics*, 128-130.

Egyptian king Amasi.[32] Thales' work on geometric similarity was completed by the Greek mathematician Euclid. Euclid made an important contribution in terms of defining the concepts of ratio and proportion in his fifth book of *Elements.* According to Euclid:

> A ratio is a sort of relation in respect of size between two magnitudes of the same kind and magnitudes, which have the same ratio are called proportional.[33]

As Douglas Jesseph points out in his article "Ratios, Quotients, and the Language of Nature":

> A ratio is not a quotient formed by the division of one number by another, but rather a relation that holds between geometric magnitudes.[34]

Through his theories, Euclid succeeded in systematizing the existing knowledge, while, at the same time, laying the foundations for what would later be called geometric similarity.[35]

An important contribution to the development of the concept of similarity was that of Aristotle, who understood the concept and used it methodologically. The notion of similarity is found in Aristotle's distinction of the "being" in matter and form, which he defined as the sum of the attributes that each being has in common with other beings and integrates it into a class of similar beings.[36] Aristotle used the "form" in his attempt to describe and categorize animal species in a series of extensive zoological treatises, the most widely known of which is *Περί τα ζώα ιστορίαι* (Animal Histories). In this treatise, Aristotle carefully classified and described 500 species, which he distinguished mainly based on traditional classifications based on multiple features.[37]

We can suggest fairly certainly that during classical antiquity, similarity

---

[32] Ibid., 128-130.

[33] Jesseph Douglas, "Ratios, Quotients, and the Language of Nature," in *The Language of Nature*, eds. Geoffrey Gorham, Benjamin Hill, Edward Slowik, and C. Kenneth Waters, 160-177 (Minneapolis: University of Minnesota Press, 2016).

[34] Ibid.

[35] Sterrett, "Similarity and Dimensional Analysis," 799-800.

[36] Lindberg, *Οι Απαρχές της Δυτικής Επιστήμης*, 68-72.

[37] Ibid., 88-90; Αριστοτέλης, *Περί τα ζώα ιστορίαι, Βιβλία Α-Ε,* απόδ. Αλέξανδρος Βασιλειάδης (Θεσσαλονίκη: Εκδόσεις Ζήτρος, 2017); Αριστοτέλης, *Περί τα ζώα ιστορίαι, Βιβλία Κ-Ζ,* απόδ. Αλέξανδρος Βασιλειάδης (Θεσσαλονίκη: Εκδόσεις Ζήτρος, 2018).

played a significant role in the examination, description, and explanation of the world. Ancient Greek philosophers consciously incorporated the mechanism of similarity into their scientific methodology.

IV. The absence of the mechanism of similarity from the research methodology of the Middle Ages (500–1500 AD)

During the Middle Ages, the research focus was mainly on collecting, organizing, and critiquing the existing theoretical knowledge passed down from ancient Greek natural philosophers, in order to serve the purposes of ecumenical church. From 500 AD to 1000 AD, the political and social instability led to the decline of Western science.[38] In order to gain knowledge, the majority of scholars focused mainly on organizing and disseminating ancient Greek science theories and conclusions, but not on the research methodology or experimental techniques used by ancient Greeks.[39] However, during the Late Middle Ages, a number of researchers conducted experiments, but their findings were used to form descriptive encyclopedias rather than to explain or make predictions about natural phenomena.[40] Consequently, until 1200 AD the research activity was not characterized by well-organized and systematic experimentation[41] and the mechanisms, tools, and techniques, such as the mechanism of similarity, of the modern scientific methodology were not being used by the majority of intellectuals.

The appearance of the first universities in the 12th century, contributed to an increase in translations, ancient text critiques, and the organization and expansion of the existing scientific knowledge. After the 13th century, courses on Logic, Physics, Astronomy, Cosmology, and Mathematics were in the core of university education.[42] During this period, the first step of the transition from natural philosophy to science took place within universities. The concept of the scientific hypothesis was introduced into the research process.[43] When researchers were studying ancient texts, they formulated hypotheses in the form of questions, known as "Questions," and they answered them in the form

---

[38] Edward Grant, *Οι Φυσικές Επιστήμες τον Μεσαίωνα*, μτφ. Ζήσης Σαρίκας (Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 2013), 1.

[39] Herbert Butterfield, *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800)*, μτφ. Ιορδάνης Αρζόγλου και Αντώνης Χριστοδουλίδης (Αθήνα: ΜΙΕΤ, 2010), 79-82; Grant, *Οι Φυσικές Επιστήμες τον Μεσαίωνα*, 7-9.

[40] Butterfield, *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800)*, 80-81.

[41] Grant, *Οι Φυσικές Επιστήμες τον Μεσαίωνα*, 8.

[42] Ibid., 32-33.

[43] Ibid., 34-37.

of comments.[44] The introduction of hypotheses in the scientific methodology was an important contribution of the Middle Ages to the development of scientific thought, methodology, and the constitution of new science, of which the research hypothesis is an integral part. After the 14th century, the spread of nominalistic tendencies was gradually observed and the doctrine of "saves the phenomena" was back in the spotlight.[45] These circumstances, along with the strong criticism on Aristotle's natural philosophy, led to the next evolutionary stage of scientific methodology, which appeared during Renaissance.

V. The period of understanding and applying the concept of similarity in the Natural Sciences (16th-19th centuries)

After Copernicus and Galileo's discoveries of celestial bodies and their movements, the preceding scientific methodology was disputed and the ancient explanation of the universe began to collapse[46] and was replaced by new methods and explanatory principles. Eventually, this was followed by the period of the Scientific Revolution (1543-1687), during which the natural sciences advanced rapidly, and the need for a general scientific methodology emerged gradually.[47]

Owing to the Scientific Revolution, the late 17th and 18th centuries saw the appearance of the intellectual movement of the Enlightenment in England and France, respectively; this movement then spread to the rest of Europe. The roots of the Enlightenment are traced in the theory of rationalism, according to which knowledge can be acquired just through pure reason; in other words, the acquisition of knowledge is achieved through a more objective way of thinking that is free from prejudice or from unverifiable assumptions of religious revelation.[48]

The Scientific Revolution and the Enlightenment marked significant changes in the scientific methodology in terms of the perception and explanation of the world, thus laying the foundations for the formulation of the new science. The mechanistic idea,[49] the acceptance of logic as a basic tool of the correct method, and the exploitation of mathematics as the main technique of the experimental method are the three essential characteristics of

---

[44] Ibid., 34-37, 139-140.

[45] Ibid., 52-56.

[46] Richard S. Westfall, *Η Συγκρότηση της Σύγχρονης Επιστήμης,* μτφ. Κρινιώ Ζήση (Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 2008), 1-34.

[47] Butterfield, *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800)*, 79-96.

[48] Dorinda Outram, *The Enlightenment* (New York: Cambridge University Press, 1995), 1-11, 47-55.

[49] Westfall, *Η Συγκρότηση της Σύγχρονης Επιστήμης,* 35-116.

the scientific methodology after the 17th century.[50] These new conditions led to the development of the 17th-century experimental method that sought to turn to nature and directly examine it through systematic experimentation,[51] that is, through the directed and organized observation of the real world through experimental measuring instruments and the development of new scientific techniques utilizing mechanisms and models capable of contributing to the explanation and prediction of phenomena.

In the context of modern science, significant efforts have been exerted to define the concept of similar systems as it was developed after the 17th century and to work out an extensive exploitation of the mechanism of similarity in the natural sciences in the period of modernity (18th-20th centuries).

Galileo used the idea of similar systems in his attempt to explain particular behaviors of machines and structures in general. Galileo focused not only on geometrical similarity, i.e., on the similarity of the dimensions or structures, but also on the proportion of relationships between natural quantities. Galileo made his most important contribution to the development of the concept of similar systems with his pendulum experiments and his law of correspondence. Galileo observed that the quantities determining the behavior of a pendulum are characterized by a constant relationship, which applies to all pendulums. These quantities are the oscillation time and the length of the pendulum's string. According to his observations, the ratio of the length of the string to the frequency of the pendulum oscillations is constant and applies to every pendulum. This constant ratio constitutes a correspondence law, which correlates each of these two quantities of one pendulum with their corresponding quantities in another pendulum, thereby allowing Galileo to calculate the length of a pendulum's string from the number of oscillations of the two pendulums at a given time. The idea that each pendulum relates to another pendulum with a law of correspondence, forms the basis of the idea of similar systems.[52]

During the early 17th century, the application of the mechanism of similarity can be traced in experimental physics and, more specifically, in the study of "subtle" or "imponderable" fluids. The movement of electricity, heat, gravity, and magnetism, which have physical properties, but do not

---

[50] Butterfield, *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800)*, 79-96; Thomas L. Hankins, *Επιστήμη και Διαφωτισμός,* μτφ. Γιώργος Γκουνταρούλης (Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 1998), 1-10, 12, 25-32; Outram, *The Enlightenment*, 47-55.

[51] Butterfield, *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800),* 79-96; Hankins, *Επιστήμη και Διαφωτισμός*, 67-73; Westfall, *Η Συγκρότηση της Σύγχρονης Επιστήμης*, 35-36, 162-169.

[52] Sterrett, "Physical Models and Fundamental Laws," 57-59; Sterrett, "Physically Similar Systems," 384-387.

constitute regular material,[53] conveys their physical properties, but it does not carry mass. When researchers observed heat flowing from a hot to a cold object, they did not detect any changes in mass.[54] In order to describe and explain this movement, they compared its similarity to the motion of fluids. Until then, the concept of similar systems may not had been defined, but knowledge on subtle fluids allows us to infer that scientists had understood the role of similarity in the process of drawing scientific conclusions and had incorporated it into their scientific methodology when they considered that it would be useful.

In late 17th century, Newton in his second book of Principia, defined the concept of similar systems for first time in the history of the concept, as follows:

> Suppose two similar systems of bodies consisting of an equal number of particles, and let the correspondent particles be similar and proportional, each in one system to each in the other, and have a like situation among themselves, and the same given ratio of density to each other; and let them begin to move among themselves in proportional times, and with like motions (that is, those in one system among one another, and those in the other among one another). And if the particles that are in the same system do not touch one another, except in the moments of reflection, nor attract, nor repel each other, except with accelerative forces that are inversely as the diameters of the correspondent particles, and directly as the squares of the velocities: I say, that the particles of those systems will continue to move among themselves with like motions and in proportional times.[55]

In order to assess if two systems were similar, Newton focused on geometrical and structural (mass, density) similarities between two systems of bodies, the proportion of the movement between particles, and the movement duration.[56] In contrast to Galileo, who used the idea of similar systems as a specialized method aimed at explaining exclusively the behavior of pendulums, Newton presents the idea of similar systems as a method with general applications.[57]

---

[53] Hankins, *Επιστήμη και Διαφωτισμός*, 73-78.

[54] Ibid., 73-78.

[55] Sterrett, "Physically Similar Systems: A History of the Concept," 382.

[56] Ibid., 382-383.

[57] Ibid., 382-387.

Newton's approach was the starting point for the examination of the concept of similar systems, sparking a series of theories from researchers coming mainly from the fields of natural sciences and engineering. The term "similar systems" introduced by Newton was a reference point until the early 20th century.

It is clear that this period was characterized by extensive efforts to understand and define similarity. This proves that the Scientific Revolution and the Enlightenment contributed significantly to the development of the concept of similarity and to the utilization of the mechanism of similarity as an experimental technique of the natural sciences after the 17th century. However, it is worth noting that despite the significant changes in the scientific methodology developed during this period, the terms "science" and "scientist" did not appear until the 1830s, when they were first used in England; until then, the term natural philosophy was used instead.[58]

VI. The stage of the systematic utilization of the mechanism of similarity in the natural sciences (19th–21st centuries)

An important year for the development of the concept of similar systems was 1914, as it was then that Edgar Buckingham, an American physicist, proposed the term "physically similar systems" in order to replace Newton's previously accepted term "similar systems." His approach was as follows:

> Let S be a physical system, and let a relation subsist among a number of quantities Q, which pertain to S. Let us imagine S to be transformed into another system S' so that S' "corresponds" to S as regards the essential quantities. There is no point of the transformation at which we can suppose that the quantities cease to be dependent on one another: hence we must suppose that some relation will subsist among the quantities Q' in S', which correspond to the quantities Q in S. If this relation in S' is of the same form as the relation in S and is describable by the same equation, the two systems are "physically similar" as regards this relation.[59]

A common characteristic between Newton's and Buckingham's approaches of the concept of similar systems was the identification of a ratio between physical quantities or the relationship of physical quantities. While Newton defined similar systems on the basis of their similar structural characteristics

---

[58] Outram, *The Enlightenment,* 48-49.

[59] Sterrett, "Physically Similar Systems: A History of the Concept," 380-381.

(mass and density), Buckingham defined them on the basis of the proportional relationships observed between specific physical quantities of interest. Since 1914 the term "physically similar systems" introduced by Buckingham, has been widely accepted and used up to this day.

The systematic utilization of similarity extended significantly after the 19th century, mainly in the fields of Engineering and Physics. William Froude developed an interesting approach focusing on utilizing the similarity mechanism for ship design and construction. William Froude was an English engineer who got involved in hydrodynamics and ship design during the early 19th century. He utilized the concept of similar systems to solve major problems encountered in the construction of ships for the English Navy; these problems had to do with stability, ship speed, and the interaction between ships and water in motion or stillness.[60] The notion of similar systems in Froude, as in Newton, took into account correlating quantities in one situation with corresponding quantities in another situation.[61] In particular, Froude carried out experiments with ship scale models and extended the inferences of his experiments, through the appropriate calculations, to full-sized ships.[62]

VII. Similarity as a core mechanism of scientific models in modern science: Susan G. Sterrett's view

Susan G. Sterrett is a Professor of History and Philosophy of Science at Wichita State University in Kansas, US. While she initially studied Mechanics, later on her research interests focused on the field of History and Philosophy of Science. Her work focuses on issues related to the methodology of science, with her major contribution being highlighting the importance of similarity concepts and scientific models in the field of Philosophy of Science; the significance of such concepts has already been recognized in natural sciences and engineering.

According to Sterrett, the concept of similarity is powerful in the field of natural sciences and should be further examined and developed in other fields. Sterrett accepts the idea that the concept of similarity is related to the concept of ratio. She understands the concept of physical similarity as a generalization of the concept of geometrical similarity. While geometrical similarity is defined by the ratio of shapes or distance between two points, physical similarity is defined by the proportion of physical quantities pertaining to similar systems, such as time, mass, and force. In order to generalize the notion of similarity so as to apply it not only to geometry

---

[60] Ibid., 389-393.

[61] Ibid., 389-393.

[62] Ibid., 389-393.

but to natural sciences as well, the concepts of proportion and shape also had to be generalized.[63] Sterrett's significant contribution to the evolution of the concept of similarity is in highlighting the importance and the role of the scientific hypothesis in the light of which the similarity between two physical systems is determined. According to her, two systems can be characterized as physically similar when there is an analogy between specific relationships of corresponding physical quantities, which is always defined in the light of a scientific hypothesis.[64] This important observation by Sterrett contributes to a clearer definition of the concepts of similarity and similar systems, thereby placing her theory among the most important evolutionary stages of these concepts.

The majority of scientists working in the field of philosophy of science perceive scientific models as theoretical tools, which constitute an intermediate stage between theory and the real world.[66] These tools are formed by theory, laws, and principles that relate to the subject under consideration and they are used to draw conclusions about real-world situations.[67] Sterrett considers this approach as fragmentary, as it does not include a wide range of models, which are not theoretical tools of an intermediate stage, but parts of the real world, such as scale models in physics and mechanics or animal models in biology. She proposes the classification of scientific models in the categories of "realm of thought" and "using one piece of the world to tell about another." The first category includes models of abstract and mathematical structures as well as algorithms or mechanism descriptions. These tools are considered

Another important issue that concerned Sterrett was in what types of methodology is the similarity mechanism used and how are the criteria that determine the similarity between two bodies or systems selected.[65] She points out that since the beginning of the 19th century the mechanism of similarity has been associated with the concept of the scientific model, a core experimental technique widely utilized, especially in natural sciences. The importance of scientific models in describing, explaining, and predicting the natural world is recognized by researchers that are active in many scientific fields globally. Sterrett has examined extensively the utilization of the mechanism of similarity as the basic operating mechanism of scientific models.

---

[63] Sterrett, "Similarity and Dimensional Analysis," 800-801.

[64] Sterrett, "Models of Machines and Models of Phenomena," 69-80.

[65] Ibid., 69–80.

[66] Susan G. Sterrett, "Kinds of Models," in *The Multiple Meanings of Models* (John Hope Franklin Center: Duke University, 2003), 1-2, http://philsci-archive.pitt.edu/2363/.

[67] Sterrett, "Physical Models and Fundamental Laws," 56-59; Sterrett, "Kinds of Models," 1-2.

models in virtue of their relationship to some equations or formal scientific proposals.[68] Models that fall into the second category are parts of the real world. These models are commonly known as analogue models.[69] Analogue models are physical set-ups that are utilized as models of other physical set-ups, which researchers cannot observe because of their size as well as the space or time that separates them from them. The basic function of their mechanism is similarity, which is validated by a ratio of physical quantities or by a ratio of relationships observed between the physical quantities of two phenomena or objects. The analogue relationships between the model and the system of interest are based on the direction and purpose of the research, which are determined by the scientific hypothesis.[70] Similarity is defined by criteria that are determined by the phenomenon of interest and the problem to be solved. Therefore, the similarity between the model and the object of interest is usually not absolute, as it is defined in respect to a particular characteristic, which, in turn, is defined through the formulation of the scientific hypothesis.

Examples of analogue models are scale models that are extensively used in engineering and physics. Scale models are physical objects or systems, which are used to control or predict the behavior of a machine, an object, or a system of different dimensions. They are constructed in such a way that they are proportionate to an object in the physical world.[71]

Sterrett described the operation stages of scale models in order to present the utilization of the similarity mechanism in the context of this scientific technique. According to Sterrett, in the first stage, the researcher should study the physical quantities related to the phenomenon of interest. Then they should construct a physical state S2, which is similar to state S1, in the areas of their research interest. In other words, the researcher chooses the proportional relationship, which could correspond to their scientific hypothesis and constructs the model based on this relationship. This way, the researcher can define similarity based on their specific research hypotheses. Then, they develop the rules for transferring prices of quantities of S2 to S1 (principles, laws, and equations). Once the S2 model is constructed, the

---

[68] Sterrett, "Kinds of Models," 1-2, 9-11.

[69] Susan G. Sterrett, "Experimentation on Analogue Models," in *Springer Handbook of Model-Based Science*, eds. Lorenzo Magnani, and Bertolotti Tommaso (Cham: Springer International Publishing, 2017), 357-360.

[70] Sterrett, "Physical Models and Fundamental Laws," 59-63; Sterrett, "Models of Machines and Models of Phenomena," 69-80.

[71] Sterrett, "Kinds of Models," 1-3; Sterrett, "Physical Models and Fundamental Laws," 59-63; Sterrett, "Models of Machines and Models of Phenomena," 69-80; Sterrett, "Experimentation on Analogue Models," 360-362.

researcher measures the quantities, observes the behavior of the physical state, and draws inferences about the S1 state.[72]

We strongly believe that Sterrett's contribution is highly important because she opened a constructive dialogue in the field of philosophy of science on concepts, such as similar systems and scientific models that have been sufficiently examined, defined, and widely used in the experimental method of the natural sciences. Sterrett identified that the concept of similarity has been neglected in modern philosophical thought, thus managing to highlight the necessity for its further examination. Through her research, she laid the foundation for further investigation, with the main aim being to overcome problems, such as the inadequate understanding of similarity, similar systems, and scientific model concepts that sometimes lead to their fragmentary perception and their non-acceptance as formal scientific techniques by philosophers of science.

Working in this direction, Sterrett managed to contribute significantly to the sufficient definition and evolution of these concepts, with her main contributions being that she highlighted the importance and the role of the scientific hypothesis, in the light of which the similarity between two physical systems is determined, but also her observation, according to which the mechanism of similarity is the basic operating mechanism of scientific models. In this context, the mechanism of similarity could be understood as a set of rules, laws, principles, or mathematical relationships utilized by the analogue modeling technique in order to successfully validate a certain analogue relationship between the model and the system of interest in the context of a scientific hypothesis. This mechanism is utilized not only when the model is selected or constructed, but also during the process of extending the model's inferences to the object, system, or phenomenon of interest, always in light of the scientific hypothesis in question.

In this context, it becomes clear that Sterrett's contribution is not limited to her argumentation or her theories on the similarity, similar system, and scientific model concepts, which was undoubtedly important too. It could be argued that her most important contribution was highlighting how neglected these concepts are in the field of philosophy of science and how important is their further investigation. If the detection of existing knowledge during a research process is considered important, then the detection of absent knowledge should be accepted as a powerful motive able to motivate new research steps, reveal new research directions, and contribute to the development and evolution of science. We support that through the philosophical perspective, these concepts could be documented in a theoretical manner more sufficiently and recognized as formal techniques

---

[72] Sterrett, "Physical Models and Fundamental Laws," 56-58.

not only of the modern scientific methodology of natural sciences, but also in modern science overall.

VIII. Conclusions

The present historical review of the concept of similarity presented the evolution of conceptualizing and utilizing the mechanism of similarity as a practical and experimental technique, applicable to the scientific methodology of the natural sciences in various eras of human intellectual activity. The historical evolution of similarity was examined in the context of different historical periods, ranging from antiquity to the present day and is directly related to the evolution of scientific thought. According to this approach, the concept of similarity went through five distinct stages of evolution. The first stage corresponds to Egyptian science from 3200 BC to 1200 BC and it could be characterized as the beginning of the utilization of similarity, which enabled ancient Egyptians to achieve various architectural, medicinal, and mathematical feats. During this period, ancient Egyptians used similarity in a generalized manner, as a technique to categorize knowledge and contribute to the description, explanation, and prediction of the world, primarily aimed at solving practical problems. However, it is not clear to what extent the concept of similarity was defined in Egyptian science. The second stage corresponds to the Classical era, which is the era of the genesis of episteme and natural philosophy, when the notion of similarity appeared in philosophy, mathematics, music, and geometry and was perceived to be of increased methodological importance. During the Classical era, similarity was perceived and exploited consciously for the first time, while it was developed in the context of a more general attempt to describe and explain the world as viewed by ancient Greek philosophers.

The third stage was during the Dark Ages, a time of scientific stagnation. The medieval period proved unfavorable for the exploitation and development of experimental scientific techniques and mechanisms, such as the mechanism of similarity. It follows that during the Dark Ages, similarity was absent from scientific methodology. During the next evolutionary stage, after Renaissance, the concept of similarity gained renewed importance, this time as the methodological idea of similar systems. In particular, this was the period of defining and consciously utilizing similarity as an experimental tool of the natural sciences (late 16th century to early 19th century). Finally, the fifth stage corresponds to the period from the 19th century to the 21st century and constitutes the stage of the theoretical documentation and systematic application of the mechanism of similarity in the natural sciences as well as the extension of its application in many scientific fields. Two dominant theories on the concept of similarity originated in this period. The first is that

of Edgar Buckingham who introduced the term "physically similar systems," which is used up to the present day. The second is that of modern philosopher of science Susan G. Sterrett who highlighted the necessity to research further the concepts of similarity, similar systems, and scientific models in the field of Philosophy of Science, concepts whose significance had already been recognized in the natural sciences and engineering.

The study of the historical evolution of similarity clarifies that the transition from the "technocratic" exploitation of similarity to its theoretical documentation as well as to its conscious and systematic application as a tool of scientific methodology was not completed in a single evolutionary stage. On the contrary, it took many centuries for the concept to evolve in parallel with the evolution of scientific thought and to reach its modern significance and application within the scientific methodology of the natural sciences. The transition from the "technocratic" utilization of similarity to its conscious utilization can be detected in the classical era stage. The transition to its systematic use as a significant experimental tool is traced after Renaissance. Finally, the theoretical documentation of the concept of similarity and efforts to expand its application to more scientific fields, are traced after the 19th century. Although all stages contributed to the development of the concept of similarity, the period after the scientific revolution is considered crucial for the conceptualization and utilization of the mechanism of similarity. The changes that occurred in science after the Scientific Revolution and the Enlightenment played a decisive role in the evolution of the concept of similarity. The Scientific Revolution and the Enlightenment helped shape a new way of thinking that changed the way scientists research the natural world. The incorporation of systematic experimentation into scientific methodology resulted in the need to develop new scientific practices, including measuring instruments and the systematic exploitation of mechanisms and scientific models capable of contributing to the explanation and prediction of phenomena. These conditions contributed to the immediate adoption of the mechanism of similarity and to its systematic application in scientific models, which was greatly expanded from the 18th century onwards. Moreover, from Newton's concept of geometrical similarity to Buckingham's concept of physical similarity, and finally to the concept of physical similarity in the light of a specific research hypothesis in Sterrett's approach, these circumstances enabled the adoption of a multifaceted approach, a deeper understanding, and a more sufficient definition of the concept of similarity and its evolution.

Sterrett's significant addition contributes to a clearer definition of the concepts of similarity and similar systems. The emphasis she placed on the significance of the scientific hypothesis during the process of defining the similarity between two systems, rightly places her theory between the

important evolutionary milestones of concepts of similarity and similar systems. Except for this, Sterrett observed that the concepts of similarity, similar systems, and scientific models are neglected in modern philosophical thought and recognized the necessity to further examine them in the field of the philosophy of science. This view seems reasonable, as a more systematized philosophical research of these concepts could lead to a more comprehensive understanding, better clarification, description, and adequate theoretical documentation of them. A meticulous philosophical study of these concepts could reinforce the existing theory coming from natural science research and contribute to their safer and more efficient use as methodological tools and the expansion of their application into other scientific areas. Thus, Sterrett pointed out the absence of sufficient theories and knowledge regarding the concept of similarity in the field of philosophy of science, thereby provoking an open and constructive dialogue in this field.

References

BBC. "Egypt Animal Mummies Showcased at Saqqara near Cairo." Accessed January 17, 2020. https://www.bbc.com/news/world-middle-east-50531808.

Butterfield, Herbert. *Η Καταγωγή της Σύγχρονης Επιστήμης (1300-1800).* Μετάφραση Ιορδάνης Αρζόγλου και Αντώνης Χριστοδουλίδης. Αθήνα: ΜΙΕΤ, 2010.

Earman, John. "Το Πρόβλημα του Ντετερμινισμού στις Φυσικές Επιστήμες." Στο *Εισαγωγή στη Φιλοσοφία της Επιστήμης,* επιμέλεια Αριστείδης Μπαλτάς, μετάφραση Πάνος Θεοδώρου, Κώστας Παγωνδιώτης, Γιώργος Φουρτούνης, 317-369. Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 1998.

Encyclopaedia Britannica. "Thales of Miletus." Accessed November 25, 2019. https://www.britannica.com/biography/Thales-of-Miletus.

Encyclopaedia Britannica. "History of Technology Timeline." Accessed January 17, 2020. https://www.britannica.com/story/history-of-technology-timeline.

Grant, Edward. *Οι Φυσικές Επιστήμες τον Μεσαίωνα.* Μετάφραση Ζήσης Σαρίκας. Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 2013.

Hankins, Thomas. L. *Επιστήμη και Διαφωτισμός.* Μετάφραση Γιώργος Γκουνταρούλης. Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 1998.

Heath, Thomas. *A History of Greek Mathematics, Volume 1: From Thales to Euclid.* Oxford: Clarendon Press, 1921.

Henry, George. "A Greek-English Lexicon." Accessed July 5, 2020. http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.04.0057:entry=e)pisth/mh.

Ingold, Tim. *The Perception of the Environment. Essays in Livelihood, Dwelling and Skill*. London and New York: Routledge, 2000.

Jesseph, Douglas. "Ratios, Quotients, and the Language of Nature." In *The Language of Nature*, edited by Geoffrey Gorham, Benjamin Hill, Edward Slowik, and C. Kennedy Waters, 160-177. Minneapolis: University of Minnesota Press, 2016.

Jorgensen, Lise Bender. "Introduction to Part II: Technology as Practice." In *Embodied Knowledge: Perspectives on Belief and Technology,* edited by Marie Louise Stig Sorensen, and Katharina Rebay-Salisbury, 89-94. Oxford: Oxbow Books, 2012.

Lindberg, David C. *Οι Απαρχές της Δυτικής Επιστήμης*. Μετάφραση Ηλίας Μαρκολέφας. Αθήνα: Πανεπιστημιακές Εκδόσεις Ε.Μ.Π., 2003.

Moss, Jessica. "Is Plato's Epistemology About Knowledge?" In *What the Ancients Offer to Contemporary Epistemology*, edited by Stephen Hetherington, and Nicholas D. Smith, 68-85. Oxfordshire: Routledge, 2020.

Nunn, John F. *Ancient Egyptian Medicine*. Oklahoma: University of Oklahoma Press, 2002.

Outram, Dorinda. *The Enlightenment*. New York: Cambridge University Press, 1995.

Reeves, John. "The Science and Religion Dialogue as Natural Philosophy." *Metanexus*. Accessed July 5, 2020. https://www.metanexus.net/science-and-religion-dialogue-natural-philosophy/.

Shaw, Stephanie J. *W.E.B Du Bois and the Souls of Black Folk*. North Carolina: The University of North Carolina Press, 2013.

Sterrett, Susan G. "Physical Models and Fundamental Laws: Using One Piece of the World to Tell About Another." *Mind & Society* 3, no. 1 (2002): 51-66.

Sterrett, Susan G. "Kinds of Models." Paper based on a talk presented at an STS Interdisciplinary roundtable: "The Multiple Meanings of Models," John Hope Franklin Center, Duke University, 2005.

Sterrett, Susan G. "Models of Machines and Models of Phenomena." *International Studies in the Philosophy of Science* 20, no. 1 (2006): 69-80.

Sterrett, Susan G. "Similarity and Dimensional Analysis." In *Handbook of the Philosophy of Science, Volume 9: Philosophy of Technology and Engineering Sciences*, edited by Anthonie Meijers, 799-824. Amsterdam: North Holland, 2010.

Sterrett, Susan G. "Experimentation on Analogue Models." In *Springer Handbook of Model-Based Science,* edited by Lorenzo Magnani, and Bertolotti Tommaso, 857-878. Cham: Springer International Publishing, 2017.

Sterrett, Susan G. "Physically Similar Systems – A History of the Concept." In *Springer Handbook of Model-Based Science*, edited by Lorenzo Magnani, and Tommaso Bertolotti, 377-411. Cham: Springer International Publishing, 2017.

Westfall, Richard S. *Η Συγκρότηση της Σύγχρονης Επιστήμης.* Μετάφραση Κρινιώ Ζήση. Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης, 2008.

Αριστοτέλης. *Περί τα ζώα ιστορίαι, Βιβλία Α-Ε.* Απόδοση Αλέξανδρος Βασιλειάδης. Θεσσαλονίκη: Εκδόσεις Ζήτρος, 2017.

Αριστοτέλης. *Περί τα ζώα ιστορίαι, Βιβλία Κ-Ζ.* Απόδοση Αλέξανδρος Βασιλειάδης. Θεσσαλονίκη: Εκδόσεις Ζήτρος, 2018.

Religious Dimensions in Transhumanist and Posthumanist Philosophies of Science

Evaldas Juozelis

Lithuanian University of Health Sciences, Lithuania
E-mail address: evaldasjuozelis@gmail.com
ORCID iD: http://orcid.org/0000-0003-0463-496X

Abstract

The article discusses transhumanism and posthumanism as marginal trajectories of the modern philosophy of science, which, however, distinctly influence the mainstream narrative of science and societal relations. Among the decisive determinants of this impact is trans/posthumanism's para-religious content that replenishes a conceptualised process of cutting-edge scientific practices and ideals. In particular, transhumanism and posthumanism evolve as ideological exploiters of seemingly obsolete forms of religiosity, for they simultaneously exploit and reinvent the entire apparatus of the scientific, political, and moral activity in Western societies. Avant-garde secular worldviews tend to be religious in the sense that their ultimate quest is the transformation of humans into certain historical entities, which are capable of rearranging their own systems of order.

Keywords: *transhumanism; posthumanism; science; society; religiosity; Steve Fuller*

I. Introduction

In this article I view *transhumanism* and *posthumanism* as a certain *modus operandi* in contemporary philosophy of science, while I am trying to demonstrate that their quite often overt religious content plays a decisive role in their conceptualisations of cutting-edge scientific practices and ideals.

Initially, I will present definitions of religiosity. Setting aside the trivial notion of religion as a belief in supernatural powers (monotheism, polytheisim, spiritualism, etc.) that constitute moral principles and rituals, I rely on the definitions given by Irving Hexham and Clifford Geertz. By adding few elements from the concepts of civil religion by Enn Kasak and invisible religion by Thomas Luckmann, I aim to gain a solid ground in order to interpolate

scientific beliefs in a corporate modern worldview, where religious as well as scientific techniques of uniquely realistic meaning-formation overlap, for they are both sanctioned universally objective by culture and both transcend immediate human experience. Here, transhumanism and posthumanism (and their various subforms) unfold as ideological exploiters of an apparently exhausted phenomenon of 'religion/religiosity' and simultaneously (are prompted to) redefine the nature of science, human being, and prosperity.

Then, by briefly introducing transhumanist and posthumanist attitudes toward the prospects of human development, I will assume that the insufficient consideration of the ideological basis of scientific understanding undermines efforts of legitimating the scientific worldview, increases the risk of the negligent apprehension of human needs, and eventually compromises integrative models of science, technology, and society. These models take quietly their own ideas and ideals (responsibility, well-being, scientific progress, morality, etc.) for granted and unreflectively operate them as aims by providing unreliable arguments.

Therefore, aiming to highlight religious dimensions of these marginal scientific worldviews, I will unfold their scientifically engaged and ideologically contested self awareness as 'secularly religious' by arguing that what basically emerges, is constituted and evolved within the interactive 'post humanist/transhumanist' medium of cultural praxis. Religious dimensions of post/transhumanist praxis translate biological, social, and cultural distinctions into conventional categories. Consequently, habitual efforts to separate scientific knowledge from a broadly ideological environment, to interpret it as socially self justifying and organised acts based on rational decisions or individual capacities, are implicitly considered questionable and problematic.

II. The necessary extensions of religiosity

According to Irving Hexham, religion contains:

> intellectual, RITUAL, SOCIAL, and ETHICAL elements, bound together by an explicit or implicit BELIEF in the REALITY of an unseen world, whether this belief be expressed in SUPERNATURALISTIC or IDEALISTIC terms.[1]

Equally important is Hexham's note that precisely any definition of religion reflects 'a scholarly or a DOGMATIC bias' of the person forming the definition. One must keep that in mind.

---

[1] Irving Hexham, *Concise Dictionary of Religion* (Canada: Vogelstein Press, 1993), 186.

In his definition of religion, Clifford Geertz dispenses with the postulate of supernatural (seemingly obeying his own scholarly bias):

> a religion is: (1) a system of symbols, which acts to (2) establish powerful, pervasive, and long-lasting moods and motivations in men by (3) formulating conceptions of a general order of existence and (4) clothing these conceptions with such an aura of factuality that (5) the moods and motivations seem uniquely realistic.[2]

Initially, I refuse to judge the sufficiency of these definitions presented herein; consequently, I will not propose any definition of science and will not address the relationship between science and religion. My aim could be better articulated in showing how the transhumanist and posthumanist philosophers construct certain a) 'conceptions of a general order of existence,' that are followed by b) 'auras of factuality,' which convert their ideas into c) 'uniquely realistic' worldviews.

Additionally, a few more elements should be introduced for that purpose:

❑ Religion functions as a projected medium in order to transform human beings into participants of a specific historical-social course. If a component of human reality fulfils this function, it can be rightfully called 'religious.'[3]
❑ Certain beliefs in science resemble religious ones, but in a non-doctrinal sense: a belief is considered religious if it corresponds to Geertz's definition; a belief is considered scientific if it corresponds to intersubjective experience within the rules and context of the actual discipline.[4]
❑ The stronger the social regulation and pressure from society on science (to standardise thoughts and actions, to integrate individuals) is, the more it resembles religion and ideology.[5]

---

[2] Clifford Geertz, *The Interpretation of Cultures* (New York: Basic Books, 1973), 90.

[3] Thomas Luckmann*, The Invisible Religion*: *The Problem of Religion in Modern Society* (New York: MacMillan Publishing Company, 1967), 61.

[4] Enn Kasak, "Unperceived Civil Religion in Science," *Problemos* 80 (2011): 99-100.

[5] Serge Moscovici, "The New Magical Thinking," *Public Understanding of Science* 23, no. 7 (2014): 762.

III. Transhumanist and posthumanist revelations

Among the innumerable definitions of transhumanism, I personally prefer the following:

> Transhumanism is a class of philosophies that seeks the continued evolution of human life beyond its current human form as a result of science and technology guided by life-promoting principles and values. Transhumanism promotes an interdisciplinary approach to understanding and evaluating the opportunities for enhancing the human condition and the human organism opened up by the advancement of technology.[6]

Nevertheless, to understand better the transhumanist agenda, we have to look carefully at the *Transhumanist Declaration* (2009), particularly at the three (out of eight) following statements:

1. Humanity stands to be profoundly affected by science and technology in the future. We envision the possibility of broadening human potential by overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth.
6. Policy making ought to be guided by responsible and inclusive moral vision, taking seriously both opportunities and risks, respecting autonomy and individual rights, and showing solidarity with and concern for the interests and dignity of all people around the globe. We must also consider our moral responsibilities towards generations that will exist in the future.
7. We advocate the well-being of all sentient beings, including humans, non-human animals, and any future artificial intellects, modified life forms, or other intelligences to which technological and scientific advance may give rise.[7]

The aforementioned 'responsible and inclusive moral vision' and 'the well-being of all sentient beings' presently beg for a certain factual, sincerely transhumanist context. I take the opportunity here to speak of Steve Fuller,

---

[6] "Roots and Core Themes," in *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, eds. Max More, and Natasha Vita-More (Hoboken, NJ: Wiley-Blackwell, 2013), 1.

[7] "Transhumanist Declaration," *Humanity +*, accessed January 14, 2021, https://humanityplus.org/philosophy/transhumanist-declaration/.

the famous transhumanist philosopher, who is unique in providing such context.

In his book *Knowledge: The Philosophical Quest in History* (2015) Fuller cultivates the idea that science, if correctly understood, unequivocally demonstrates the divinity of human beings; or, to be more accurate, that humans, somehow, are not part and parcel of nature. Allegedly, this is the true mission of the unified institution of science and its religious core. In this context, taking science seriously means to endorse the purest 'Good News' of the transhumanist kingdom: 'Humans are gods in the making.'[8]

How so? Because humans invented science, and, inversely, they are themselves defined by science: scientifically speaking, if everything in its own existence could be resumed under Darwinian terms, we would not comprehend the Darwinian evolutionary theory. Therefore, it follows that human beings cannot be only Darwinian evolutionary subproducts. We are not natural human beings, so the argument goes, because science is not natural, and science explicitly is 'the dominant feature of the human being.'[9] That is how Fuller's transhumanism religiously generates 'the existential general order' with all the necessary decorum of the 'uniquely realistic aura of factuality.' A kind of magic, or the power of supernaturally excluding the human from the natural realm must certainly persist here.

Unfortunately, a few, minor though, issues also inevitably persist. First, a 'theological' theft of scientific autonomy: 'The best explanation for the shape and persistence of science's fundamental questions is theological.'[10] Fuller's transhumanism hurls him so far that in *Humanity 2.0. What It Means to be Human Past, Present and Future* he shamelessly seizes the 'unifying' scientific worldview in order to denounce the scientific attitude from within:

> I believe that Darwinism poses a much greater threat than Christianity or Islam to the future of humanity as a normatively salient category.[11]

Salient or not, this is just the way it is in the contemporary 'scientific worldview.'

The second minor religious issue is the Fuller's embezzlement of morality. 'Errors are unavoidable in the quest to extend human knowledge.'[12] One may

---

[8] Steve Fuller, *Knowledge. The Philosophical Quest in History* (New York: Routledge, 2015), 1.

[9] Ibid., 264.

[10] Ibid., 62.

[11] Steve Fuller, *Humanity 2.0. What It Means to be Human Past, Present and Future* (London: Palgrave Macmillan, 2011), 3.

[12] Ibid., 264.

wonder which are these unacceptable errors. Surprisingly, the transhumanist aura of the 'unique factuality' finds no such errors, because the scientific path, at least in the long run, is self-purifying and self-forgiving: 'In short, the march of progress is itself morally cleansing as we learn from our mistakes.'[13] I find the task of wreathing this sort of transhumanism with the 'responsible and inclusive moral vision' extremely difficult. Unless, as Barry Allen aptly remarks, 'human moral horizons can be sanctioned by human immortality.'[14]

The third (but inconclusive) magic trick of the Fuller's transhumanism lies in his academic arrogance:

> We must somehow believe that all the human and non-human lives lost through science-induced aggression, negligence, and obliviousness have contributed to a world that has maximised the welfare of more humans, understood as the highest form of life.[15]

Farewell to 'the well-being of all sentience.' At this point posthumanism as an ideological alternative comes into play. Posthumanism unambiguously associates human nature with the natural environment and is neither able nor willing to exterminate the divinely human 'essence'. That is, posthumanism finds nothing particularly special about human beings, and clearly declares the 'failure' in scientific terms. In this respect, posthumanism is:

> a break with humanism; it is a post-humanism. In recent years "posthumanism" served as an umbrella term for a variety of positions that reject basic humanist concepts and values. Above all, the construction of "human beings" is deemed to be ideologically laden, insufficient, dangerous, or paternalistic.[16]

The transhumanist visionaries help us, at least provisionally, to understand how a human being in his/her self-righteousness, arrogance and magnificence may incidentally transform from a dangerous idea into a dangerous entity. The aforementioned reasons are also sufficient to infer certain political agendas

---

[13] Ibid., 97.

[14] Barry Allen, "Review of *Knowledge: The Philosophical Quest in History*, by Steve Fuller," *Notre Dame Philosophical Reviews* 34, no. 3 (2015), https://ndpr.nd.edu/news/knowledge-the-philosophical-quest-in-history/.

[15] Fuller, *Knowledge,* 93.

[16] Robert Ranisch, and Stefan Lorenz Sorgner, "Introducing Post- and Transhumanism," in *Post- and Transhumanism: An Introduction*, eds. Robert Ranisch, and Stefan Lorenz Sorgner (Frankurt am Main: Peter Lang GmbH, 2014), 8.

from the sectarian transhumanist beliefs ('Left Creationism' in Fuller's case) and, more generally, from wider inclinations of the transhumanist movement to politically motivated institutions (e.g. *Humanity+*, *Institute for Ethics and Emerging Technologies*).

Posthumanism, on the other hand, 'serves as an umbrella term for ideas that explain, promote or deal with the crisis of humanism. So far, however, no common name for these critical discourses has been established.'[17] Unsurprisingly, the common denominator of the 'crisis of humanism' analysis lies upon the unorthodox yet paradoxical belief in the ephemerality of the real; hence, posthumanism distributes parareligious ideas of 'post-exclusivism' (or ontological depolarisation), 'post-exceptionalism' (or epistemological discontinuance), and 'post-centralisation' (or a sort of Nietzschean perspectivism).[18]

IV. In sum: Inconclusive return of the ideal

Posthumanist as well as transhumanist arsenals of the world imaginarium unequivocally target 'the dissolution of the idea of knowledge as a public good.'[19] However, the posthumanist vision of 'public good' is entirely different. 'Relational and multi-layered ways' of thinking, 'expanding the focus to the non-human realm in post-dualistic, post-hierarchical modes, thus allowing one to envision post-human futures, which will radically stretch the boundaries of human imagination'[20] from the transhumanist standpoint, end up only in an open revolt against 'scientific progress,' meaning that the 'normative regulation of both science and society has been effectively turned over to unconstrained markets.'[21] How it allegedly challenges the ideals of 'open society,' thus insinuating the deviously tyrannical character of the posthumanist agenda, remains a mystery.

Nevertheless, both narratives, quite obviously, exploit human religiosity, i.e. they construct alternative beliefs in the reality of an unseen world,[22] by simultaneously and inevitably exposing their ideological biases.

Inversely, societal regulations, pressure and expectations upon science, such as 'theological,' political, moral, etc. demands of 'public good,'

---

[17] Ibid., 14.

[18] Francesca Ferrando, "Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations," *Existenz: An International Journal in Philosophy, Religion, Politics and the Arts* 8, no. 2 (2013): 30.

[19] Fuller, *Knowledge*, 93.

[20] Ferrando, 30.

[21] Fuller, *Knowledge*, 208.

[22] Hexham, 186.

proportionally transform it into a religiously, perhaps even magically, arranged social system.[23]

Precisely, these external impediments disclose an opportunity to unveil the tacit aspects of scientific understanding that we may call a 'secular religiosity,' and an understanding of sorts as a post/transhuman condition. The aforementioned condition dictates that our 'avant-garde' worldviews are irrevocably oriented towards the religious urge to transform humans into certain historical entities, capable of rearranging their own present and future and constituting their own systems of order.

However, transhumanism and posthumanism as marginal philosophies of the sciences are too multifarious as intellectual/cultural movements; in fact, they lack the ideological backbone required to become 'systems of symbols.' Transhumanism is plainly incapable of offering any transparent criteria as to what ultimately *is* human, what being *human* actually means. Therefore, why should we expect any definite plan of human enhancement from transhumanism?

From the posthumanist perspective, the very idea of the 'improvement' of this obscure creature deserves sacramental denouncement via rituals of dehierarchisation, decentralisation, dehumanisation, if necessary – even descientification.

Acknowledgements

References

Allen, Barry. "Review of *Knowledge: The Philosophical Quest in History*, by Steve Fuller." *Notre Dame Philosophical Reviews* 34, no. 3 (2015). https://ndpr.nd.edu/news/knowledge-the-philosophical-quest-in-history/.

Ferrando, Francesca. "Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations." *Existenz: An International Journal in Philosophy, Religion, Politics, and the Arts* 8, no. 2 (2013): 26-32.

Fuller, Steve. *Humanity 2.0. What It Means to be Human Past, Present and Future.* London: Palgrave Macmillan, 2011.

---

[23] Moscovici, 760-761.

Fuller, Steve. *Knowledge. The Philosophical Quest in History*. New York: Routledge, 2015.

Geertz, Clifford. *The Interpretation of Cultures*. New York: Basic Books, 1973.

Hexham, Irving. *Concise Dictionary of Religion*. Canada: Vogelstein Press, 1993.

Humanity +. "Transhumanist Declaration." Accessed January 14, 2021. https://humanityplus.org/philosophy/transhumanist-declaration/.

Kasak, Enn. "Unperceived Civil Religion in Science." *Problemos* 80 (2011): 94-106.

Luckmann, Thomas. *The Invisible Religion*: *The Problem of Religion in Modern Society*. New York: MacMillan Publishing Company, 1967.

More, Max, and Natasha Vita-More. "Roots and Core Themes." In *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, edited by Max More, and Natasha Vita-More. Hoboken, NJ: Wiley-Blackwell, 2013.

Moscovici, Serge. "The New Magical Thinking." *Public Understanding of Science* 23, no. 7 (2014): 759-779.

Ranisch, Robert, and Stefan Lorenz Sorgner. "Introducing Post- and Transhumanism." In *Post- and Transhumanism: An Introduction*, edited by Robert Ranisch, and Stefan Lorenz Sorgner, 7-27. Frankfurt am Main: Peter Lang GmbH, 2014.

Proposing a Frame of Ethical Principles for Educational Evaluation in Modern Greece

Georgios Tsitas

University of Nicosia, Cyprus
E-mail address: gdtsitas@yahoo.gr
ORCID iD: http://orcid.org/0000-0002-0998-5315

Athanasios Verdis

National and Kapodistrian University of Athens, Greece
E-mail address: averdis@ppp.uoa.gr
ORCID iD: https://orcid.org/0000-0003-2354-1507

Abstract

In a country with a long philosophical tradition like Greece, the lack of Ethical Principles for educational evaluation is surprising. This article presents the reasons for such a gap within the general theoretical framework for educational evaluation, combined with major schools of thought on Ethics. The authors discuss the importance for educational evaluation and assessment and take a critical view of present ethical frames. They proceed to fill the gap by coming up with a list of twenty-seven Ethical Principles, the result of the varying consensus of sixteen Greek assessment experts, upon the researcher's initial proposals. The Delphi Method, that was employed to formulate the list, is described and the first complete Ethical frame of educational evaluation for modern Greece is proposed.

Keywords: *ethical; moral; deontology; educational assessment; evaluation; principles*

I. Introduction

In the age of "ethical otinanism," a Greek neologism used to describe the fact that moral and immoral, right and wrong, good and bad are frequently referred to as equivalent, in Greece we are still trying to apply a frame of educational assessment that will meet with all the current scientific standards of evaluation. Unfortunately, most of the efforts seem to overlook that such a project should also include a stable and clear ethical frame. According to Newman and Brown,[1] there are five gradient levels of Ethics concerning

---

[1] Dianna Newman, and Robert Brown, *Applied Ethics for Program Evaluation* (London: Thousand

evaluation: Rules, Codes, Standards, Principles and Theories. Apparently, one can simply choose the appropriate ethical level for one's purpose and to put it into practice to secure that a process is ethical. In reality, things seem to be much more complicated. Each country has its own way of solving moral issues about educational assessment. Some, like Greece, prefer simple but clear Rules in order to eliminate morally inappropriate behavior, especially during crucial exams for the assessed. Some other countries tackle the issue by setting up elaborate Codes of Ethics and Ethical Standards that bind everyone involved, notably the professionals. Very few countries, though, solve their problems using Ethical Principles, and even fewer bother to reveal or discover in which way all the above are connected or founded in Ethical Theories.

Relevant literature offers several reasons for this differentiation. First of all, the assertion that there are different levels of ethical behavior presupposes a semantic unanimity concerning words like "moral," "ethical" and "deontology." Unfortunately, this is not the case. In Greece, the word "ethical" represents a notion beyond the science of Ethics, or the quality of a person's character or even one's tendency or decision to act in the right way according to the values of each society.[2] Centuries of philosophical teachings, like Plato's and Aristotle's, and the exemplar of the Eastern Orthodox Church, have shaped the word "ethical" as experienced virtue as well.[3] On the other hand, words like "moral," and "morality" focus rather on the duties and the rights of each person, setting the appropriate limits to prevent harmful behavior,[4] thus acquiring a meaning nearer to "deontology"[5] which is used in Greek as a synonym to "Code of Ethics." Furthermore, the word "moral" per se does not exist in Greek, with the exception of words like "amoralism." The Greek word for "moral" is "ēthiko," not to be confused with "ethical" in English, which refers to ethics and morality. Consequently, one might expect, or even claim, that the linguistic wealth of the Greek vocabulary and tradition provides clear ethical directions towards any action for individuals in this country – educational assessment included. This is both arbitrary and misleading. Polysemy creates more problems than the ones it solves, because it must fit the "ethical culture" of each county, each social group or even each person. It is erroneous to consider Greek Ethical tradition superior to the rest of the moral culture in the western world, as it is a falsehood to think that

---

Oaks, CA:Sage, 1996), 21-23.

[2] Georgios Babiniotis, *Dictionary of Modern Greek* (Athens: Lexicology Center, 1998), 727; Peter Singer, "Ethics." Encyclopedia Britannica, accessed February 2, 2021, https://www.britannica.com/topic/ethics-philosophy.

[3] Perry T. Hamalis, "Eastern Orthodox Ethics," *International Encyclopedia of Ethics*, 2013.

[4] Kenneth Keniston, "Morals and Ethics," *The American Scholar* 34, no. 4 (1965): 628.

[5] Babiniotis, 727.

there is consensus in Greece about what is moral. Any evaluator or evaluated person in Greece has more than one ethical paths to follow and may still fit the "ultimate good person," who freely chooses the right thing to do: one may adopt the Aristotelian middle point between two extreme options, as described in the Nicomachean Ethics, or follow the much more austere and pious Pythagorean way.[6] One can even reach the same goal by adopting a paradigm from contemporary western philosophy, like the Bergsonian "Open morality," which goes for free and simultaneously exceptional persons.[7] Ethical culture is an additional factor for the aforementioned differentiation, because it influences the way ethical issues about educational evaluation are solved, not only between countries, but also within them.

As a result, the question that arises is on which foundations should a country's educational evaluation be built on (in this case, Greece) and what difference will it create for the evaluators, or the persons evaluated. These questions can be addressed if we keep in mind that assessment is not a theoretical process, like any other discussion on ethical issues. Instead, it is a purposive moral action with consequences on people's lives, such as academic, professional, social, and psychological. Some may believe in improvement through chastisement, revealing a juridical, forensic moral perception that has its roots in the Western Church. Others may prefer an evaluator who tries to "heal" their weaknesses, treating them like patients, in accordance with Greek Orthodox Church.[8] Nevertheless, the problem remains. No one can foretell for sure which moral approach best suits everyone who is evaluated. If one's personal moral system determines the way one copes with educational assessment, then we need many different ethical codes, customized on different individuals; a tremendously difficult, if not totally impossible task.

In view of such a differentiation on a national and individual level, is there any point in discussing about ethics in educational evaluation? If we pretend that morality does not matter on this subject, then we have to be ready as civilized societies to bear the consequences. The first consequence is the legitimizing of unfair practices. The second consequence is the possibility of killing education through assessment.[9] Instead of serving education,

---

[6] Irini-Fotini Viltanioti, "Porphyry's Letter to Marcella: A Literary Attack on Christian Appropriation of (Neo-) Pythagorean Moral Wisdom?" in *Pythagorean Knowledge from the Ancient to the Modern World: Askesis, Religion, Science*, eds. Almut-Barbara Renger, and Alessandro Stavru (Wiesbaden: Harrassowitz Verlag, 2016), 168.

[7] Jozef Maria Bochenski, *Europäische Philosophie der Gegenwart*, trans. Christos Malevitsis (Athens: Dodoni Publications, 1985), 148.

[8] John Meyendorff, *Byzantine Theology: Historical Trends and Doctrinal Themes* (New York: Fordham University Press, 1974), 176, 195-196, 215, 226.

[9] Richard Pring, *The Life and Death of Secondary Education for All* (New York: Routledge,

evaluation can be used distortively to impede on any one of education's goals: personal and social enhancement, critical thinking, or creativity, to name but a few. The third consequence, which we can call "the evaluative paradox,"[10] is that a system does not ultimately practice what it preaches, because it is not ethically reliable. Kaptein has articulated four prerequisites of a reliable system: legitimacy, honesty, meticulousness, and justice.[11] Some people and some societies may not truly embrace educational evaluation and propositions for its reliability, but they acknowledge its necessity for society per se and therefore demand an ethical way of practicing it.

The last reason for the differentiation between countries in how they deal with moral issues in educational assessment has to do with the phenomenon of the evaluation per se, with its graduations, its fields, and its pivots [Table 1]. In other words, it is the differentiation in morality of each country that leads to different solutions, but it is also the variety of educational evaluative issues that seek a convincing answer within the range of all moral scope.

Table 1. Levels and ways of Correlation between Ethics and Educational Evaluation

Graduations of Ethics in Educational Assessment	Fields of application of Educational Evaluation	Ethical pivots of Educational Assessment
Ethical Rules	Assessment of students	
Codes of deontology		Evaluator
Ethical Standards of Educational Evaluation	Assessment of educational personnel and Institutions	Evaluation
		Evaluated
Ethical Principles	Program Evaluation and Evaluation of educational Systems - Policies	
Ethical Theories		

The simplest and most common way of controlling unfair practices in educational evaluation are the Ethical Rules. Both in Greek and non-Greek literature as reflected in the works of Gipps, Dimitropoulos, Konstantinou, Kassotakis, and Kapsalis and Chaniotakis,[12] there are numerous Rules

---

2013), 124-136.

[10] The paradox is that even if the moral and merit person thrives in meritocracy, the same person can be easily elbowed by an immoral person in a corrupt system.

[11] Muel Kaptein, *Ethics Management: Auditing and Developing the Ethical Content of Organizations* (Dordrecht: Kluwer Academic Publishers, 1998), 7, 32.

[12] Efstathios Dimitropoulos, *Educational Evaluation: The Evaluation of Education and the Educational Project. Part I.* (Athens: Grigoris Publications, 2007); Caroline Gipps, *Beyond*

which cover all three pivots of educational assessment: the evaluators, the evaluated and the evaluation per se. In most cases rules are simple ordains that can be set by anyone: a teacher, a school, a parent or even the student, the government, or an educational institution. Their nature is practical, their tone is directional focusing on the do's and the don'ts of each particular situation, and their main disadvantages are their multitude and contradictions. Nevertheless, scholars are unanimous on one point: assessment must fulfill certain scientific standards, like validity, reliability, objectivity and utility, and everybody involved must act accordingly to meet this goal. If an action is seen as a threat to these criteria, then a Rule can be formulated and applied.

Codes of Ethics represent a more systematic effort of dealing with ethical issues in educational evaluation. In Greece there is no official Code. Instead, there are Oaths[13] of high ethical commitment but low specificity, and a proposal of Ethical Code for the evaluation of the students by the Assessment guru in Greece, Michalis Kassotakis.[14] Internationally, although sometimes Codes and Standards are considered the same thing,[15] there are many Associations, Committees and Councils which have set their Codes of Ethics, or have collaborated in order to compose one, concerning different aspects of educational Evaluation (measurement, testing etc.).[16] In any case, Codes provide us with a frame of the appropriate evaluative behavior, that is much more than a collection of ethical rules, and can be official or unofficial, local, national or both.[17] The contribution of the Codes of Ethics[18] in educational assessment is that a) they make clear that educational assessment entails all the steps, from designing an assessment to the announcement of the results, b) they bind most the evaluators, c) they focus on the scientific training and qualification of the evaluators, so that discriminations of any

---

Testing: Towards a Theory of Educational Assessment (London, Washington: The Farmer Press, 2003); Achilleas Kapsalis, and Nikos Chaniotakis, *Educational Evaluation* (Thessaloniki: Kyriakidis Bros - Publications SA, 2015); Michalis Kassotakis, *Assessing Student Performance* (Athens: Grigoris Publications, 2013); Charalampos Konstantinou, *The Evaluation of Student Performance as Pedagogical Logic and School Practice* (Athens: Gutenberg, 2007).

[13] The foundation of professional ethics is considered to be the Oath of Hippocrates.

[14] Kassotakis, *Assessing Student Performance*, 57-60.

[15] Newman and Brown, *Applied Ethics,* 22.

[16] Helen Simons, "Ethics in Evaluation," in *Handbook of Evaluation, Policies, Programs and Practices,* eds. Ian Shaw, Jennifer Greene, and Melvin Mark (London: Sage Publications, 2006), 247.

[17] Cynthia Schmeiser, "Ethics in Assessment," *EDO-CG-95-23, ERIC Digest* (USA, 1995), https://www.counseling.org/resources/library/ERIC%20Digests/95-23.pdf.

[18] Two typical examples are the Code of Professional Responsibilities in Educational Measurement, 1995, by National Council on Measurement in Education (NCME), and the Code of Fair Testing Practices in Education, 2004, by the Joint Committee of Testing Practices.

kind or biases are minimized, scientific protocols are observed, subjectivity is eliminated, and the rights of the evaluated are protected and guaranteed.

Standards of Educational Evaluation include both ethical and practical specifications laid down by organizations,[19] revealing their interest and commitment in the proper implementation of educational assessment. Ethical Standards cover all the fields and all the pivots of the educational assessment. They are easily revised, adjusted according to current demands, or combined with Ethical Rules, Codes or Principles. In spite of the fact that they are very detailed, their most fundamental truth is that in order for an evaluation to be ethical, it must respect scientific knowledge, legislation and the human rights of all the involved parties, including both evaluators and evaluated. Among the deficiencies of Standards is the lack of internal hierarchy that would showcase the most important ones. Furthermore, their adoption or rejection is a matter of personal choice and personal ethics. In addition, in order to ensure their independence, Standards are quite costly for independent evaluators. This has raised questions among some whether their true beneficiaries are the evaluators and the evaluated or, as Lyons και Hall[20] claim, those who "shell" the tests. There is one more issue that complicates matters; Ethical Standards do not always seem to be really ethical. In some cases, the actual word is missing or deliberately effaced. Greece lacks Ethical Standards. The simple act of translating Standards designed for other educational systems demands adaptation to Greek realities, which is a quite complicated procedure.

Subsequently, a new question arises: since there are so many Ethical Rules, Codes, Standards, why do we need Ethical Principles for educational assessment? Perhaps the answer lies in the lack of consensus that calls for an overarching ethical framework to ensure the capacity to make judgments about ethical assessment practices,[21] or the fact that the boundaries between those terms are not always clear-cut. Principles for Fair Student Assessment Practices for Education in Canada,[22] for instance, are meticulously detailed. They do not have the generalized character or simplicity of such canons as:

---

[19] Schmeiser.

[20] Susan Lyons, and Erika Hall, "The Role of the Standards for Educational and Psychological Testing in Establishing a Methodology to Support the Evaluation of Assessment Quality," *Center for Assessment.* 2016, 1, accessed July 25, 2017, https://www.nciea.org/sites/default/files/publications/Standards_in_Assessment_Quality_Eval_042016.pdf.

[21] Susan Green, Robert Johnson, Do-Hong Kim, and Nakia Pope, "Ethics in Classroom Assessment Practices: Issues and Attitudes," *Teaching and Teacher Education* 23 (2007):1000.

[22] *Principles for Fair Student Assessment Practices for Education in Canada*, Edmonton, Alberta: Joint Advisory Committee (1993), https://www.wcdsb.ca/wp-content/uploads/sites/36/2017/03/fairstudent.pdf.

"Do no harm," "Avoid score Pollution," "Assess As You Would Be Assessed."[23] On the contrary, they include specific descriptions of behaviors that should be avoided.[24] Consequently, we may perhaps articulate the substance of evaluation as: "all equal, all different in educational assessment." This phrase serves as differentiator between Ethical Principles and all the other ethical graduations of educational assessment.

Ethical Principles should be specialized in individual ethical propositions and, vice versa, ethical propositions should be able to produce generally formulated, comprehensive and simple Ethical Principles.[25] This seems to be the best way to avoid strong contradictions between the above ethical propositions. If, for instance, we try to condense all the ethical propositions that focus on the priority of protecting the evaluators, then not only can we have a new Ethical Principle, "evaluator comes first," but we can also place it at the top of all evaluative principles. Nevertheless, the possibility of seeking Ethical Principles only in simple declarations might prove misleading. Sometimes, they might have the format of Standards[26] or Rules. Other times, they may lurk in the purpose of a research, like Kunnan's search[27] for fairness and validation, or in the expression of a scientific opinion like the one expressed by Gipps: "The greater the consequences of test use, the more important it is to be concerned about issues of validity and fairness in relation to test use."[28] There are cases where the Principles emerge from "dead ends." In order to overcome conflicting phenomena between Principles or other ethical issues, American Evaluation Association encourages evaluators to consult with colleagues on how to best identify and address them, because they deem Evaluators responsible "for undertaking professional development to learn to engage in sound ethical reasoning." But who is the one to judge or to define which ethical perspective is "sound?" There is always the possibility,

---

[23] Green et al., "Ethics in Classroom Assessment," 1000-1001.

[24] "Assessment methods should be bias-free from factors extraneous to the purpose of the assessment. Such factors include culture, developmental stage, ethnicity, gender, socio-economic background, language, special interests, and special needs… All students should be given the same opportunity to display their strengths." *Principles for Fair Student Assessment Practices for Education in Canada*, 5-6.

[25] Masoomeh Estaji, "Ethics and Validity Stance in Educational Assessment," *English Language and Literature Studies* 1, no. 2 (2011): 91-92, according to whom principles provide professionals with guidance upon which they can make choices.

[26] See American Evaluation Association, "Guiding Principles for Evaluators," last modified August 2018, https://www.eval.org/p/cm/ld/fid=51.

[27] Antony John Kunnan, *Fairness and Validation in Language Assessment: Selected Papers From the 19th Language Testing Research Colloquium*, *Orlando, Florida* (Cambridge: Cambridge University Press, 2000), 3-5.

[28] Gipps, *Beyond Testing*, 57.

one's personal ethics to be much more "sound" than the directions given by the Principles of an Association. If, for instance, somebody has "diakrisis,"[29] which means the charisma of judging correctly and fairly, knowing always how to treat people and how to handle truth according to uniqueness and the endurance of each person, a virtue which is considered by the Greek Orthodox Tradition as the peak of all virtues, what is the need of any other Ethical Principle? Of course, someone may claim, that "diakrisis" could have been an Ethical Principle per se, but unfortunately there is no bibliographical precedent and no way to guarantee that an evaluator possesses it. So, the Principle of personal responsibility and collaboration between evaluator should be considered to remind us that it is each person's ethical quality that determines the ethical quality of every evaluative action. Finally, some Principles are reflected in simple words that express virtues, personality characteristics, human values or rights. Newman and Brown[30] use such comprehensive words: Autonomy, nonmaleficence, beneficence, justice, and fidelity. In the Greek literature, Dimitropoulos[31] sets three main Ethical Principles: a) Educational Evaluation must be holistic, must include all parts and not be fragmentary; b) differentiation according to the purposes and c) bidirectionality (those who evaluate must also be evaluated themselves). However, this last Principle may contradict the demand for specialization among evaluators. Teachers, being the experts, may evaluate students, but is it possible for students to do so for their teachers? Such difficulties instruct us that an overarching Ethical Principle of Everything in Evaluation seems, at least for now, as far-fetched as The Theory of Everything, due to the subjective nature and the inherent imperfections of assessment. Nevertheless, Principles can not only cover all the fields and the pivots of assessment offering useful ethical guidance in contradictions, dilemmas, and conflicts, but they can also be used with no clear Ethical Theory backup. Moreover, they are the "bridge" between the Theories and all the rest of the ethical graduations because they can both specialize or summarize them. Finally, they are not attached to financial interests, because, as statements of general value, they are not subject to copyright or other restrictions. After all, it only takes a word, e.g., integrity, to compose them or a lot of personal effort to collect them, unless a researcher gathers them for the sake of the rest.

　　Ethical theories are based mostly on the views of eminent Christian and secular philosophers through the centuries, who have tackled a variety of

---

[29] Daniel Coriu, "The Path from the Natural to the Spiritual Diakrisis Through Askesis in the Views of St. Apostle Paul and Elder Joseph the Hesychast," *International Journal of Orthodox Theology* 9, no. 4 (2018): 152-175.

[30] Newman and Brown, *Applied Ethics,* 37-54.

[31] Dimitropoulos, *Educational Evaluation,* 349-351.

issues but not the one about educational assessment. Therefore, it is our task to interpret or to translate their thinking so that it fits both our school and the challenges of current evaluation. Even so, there is always the danger of arbitrary interpretations, conclusions or moral misdirections. Still, Ethical Theories can provide a completely new way of dealing with ethical matters by revealing perspectives that can help us understand evaluation clearer and deeper.

Hedonism[32] (Aristippus of Cyrene, Epicurus), for instance, could be used as a justification to eliminate educational evaluation altogether or to be indifferent to it, on the grounds that the beneficial results of assessment are, most of the times, long-term or ultimate, demanding the sacrifice of immediate pleasure which must be considered morally superior. Yet, such an interpretation sets aside epicurean "phronesis." In other words, it is a partial view of Hedonism that neglects other equally important aspects of this school of thought. Objections of this kind can always be raised for all Ethical Theories, but the fact that each one of these Theories provides a different ethical background for educational assessment has its value.

The *Theory of Instinct*[33] (Sophists, Protagoras, Hobbes) stands as a justification not only for educational evaluation, but also for the competitive spirit it entails, since it emphasizes long-term earnings and takes into account personal interest.

The *Theory of Categorical Imperative*[34] (Kant) reminds us the importance of "knowledge," namely the improvement of evaluative methods and techniques. Apart from that, a very strict and rigid educational assessment can be as acceptable as its abolition, as long as each version can be proved reasonable and consistent with Kant's Principle of Universalizability.

Utilitarianism[35] (Bentham, Mill, Aristotle) may consider educational evaluation to be the guarantee for meritocracy, a state conducive to general happiness. Yet, there are pitfalls in the theory. It is unclear how assessment can be a source of happiness for those who fail due to

---

[32] Kurt Lampe, *The Birth of Hedonism: The Cyrenaic Philosophers and Pleasure as a Way of Life* (Princeton; Oxford: Princeton University Press, 2015); Norman Wentworth DeWitt, *Epicurus and His Philosophy* (University of Minnesota Press, 1954).

[33] Richard Bett, "The Sophists and Relativism" *Phronesis* 34, no. 2 (1989): 139-69; Howard Warrender, "Hobbes's Conception of Morality," *Rivista Critica Di Storia Della Filosofia* 17, no. 4 (1962): 434-449.

[34] Allen W Wood, *Kant's Ethical Thought* (Cambridge: Cambridge University Press, 1999).

[35] Martha C. Nussbaum, "Mill between Aristotle & Bentham," *Daedalus* 133, no. 2 (2004): 60-68.

reasons irrelevant to their "value," or whether happiness is a matter of quantity or quality. In other words, utilitarianism finds acceptable and sufficient to conduct an evaluation which is beneficial for majority of population, even if it is very harmful for the minority of it.

Intuitionism[36] (Moore, Shaftesbury, Ross) seems too vague to provide any ethical direction to the way educational assessment is implemented. Yet, each of its philosophers supports an idea that can be useful and enlightening for the science of evaluation. Moore, for instance, focuses on the intuitive awareness of goodness. Ross talks about moral "duties," and Shaftesbury develops the idea of moral sense. It is up to us to combine their beliefs in order to create an educational evaluative system that reaches its ethical peak, bearing in mind that intuition might be a matter of talent – and, as a consequence, an act that can be deemed "moral" only by some "authorities" – or a matter of education and ethical standards of the social milieu.

Emotivism,[37] (Ayer, Stevenson, Hare) as a meta-ethical theory, can be interpreted in a way that totally justifies educational assessment or in a way that does not justify it at all. After all, assessment seems to have at least one thing in common with ethics. Both can be considered as expressions of approval or disapproval, which might influence other people's views. Nevertheless, there is always the possibility for somebody to theoretically embrace an ethical principle but fail to act accordingly. So, any attempt to use "universalizability" as a solution to the failings of ethics or evaluation, may not have the desired results.

Ethical Theories present all the different perspectives of ethical thinking and action. Their contribution to our struggle for a more ethical educational evaluation is that they provide ethical principles, directions, rules, conclusions, or even doubts, questions and objections. All the above not only help us understand the difficulty of reaching our almost utopian goal (or pretending to have achieved it), but also help us realize how important it is to respect and combine different ethical backgrounds or cultures.

In conclusion, in Greece we lack a scientific ethical frame of educational assessment that meets the challenges of all evaluation's Fields (students, personnel, institutions, policies) and Pivots (Evaluator, Evaluated, Evaluation per se). Among the Graduation of Ethics in Educational Assessment, only Ethical Principles seem suitable to fill the gap, because of their quality to

---

[36] Michael Huemer, *Ethical Intuitionism* (New York: Palgrave Macmillan, 2005); William Donald Hudson, *Ethical Intuitionism* (London: Macmillan, 1967); Philip Stratton-Lake, ed., *Ethical Intuitionism: Re-Evaluations* (New York: Oxford University Press, 2002).

[37] Stephen Satris, *Ethical Emotivism* (Dordrecht: Martinus Nijhoff, 1987).

"generalize" or "specialize" all the rest of the Graduations and, consequently, to cover all aspects of assessment. This research focuses on composing and setting such Ethical Principles, hoping that this may contribute crucially to a more ethical exercise of evaluation in contemporary Greece.

II. Methods and Research

In order for *Ethical Principles* to be composed, the *Delphi Method*, has been used. This research method can be considered a mixed one, a quantitative or a qualitative.[38] Although the *Delphi Method* has a long history, especially in the USA, where it appeared during the 1950s[39] inspired by the Oracle of Delphi in ancient Greece, it is not a very common method in Greece in general, and in educational research in particular. Nevertheless, it seems to be the perfect choice for research like the present author, because of its variations. In some cases, the main goal of *Delphi* may be to speculate what is likely to happen in the future, namely the possible, and in other cases to formulate what we hope will happen, namely the optative.[40] In addition, it is considered suitable for issues of ethical business that include ethical dilemmas and ask for consensus.[41]

The most crucial in the *Delphi Method* is the "*experts*," who are meant to play the role of Pythia. These experts must be truly authorities in their fields. Their heterogeneity and anonymity are also of high importance if we are to guarantee that the most eminent ones will not unwillingly impose their views and that people with different characteristics and opinions will express them equally.[42] As for the number of the experts,[43] in most cases it ranges from ten

---

[38] Dia Sekayi, and Arleen Kennedy, "Qualitative Delphi Method: A Four Round Process with a Worked Example," *The Qualitative Report* 22, no 10 (2017): 2755.

[39] Harold A. Linstone, and Murray Turoff, "Introduction," in *The Delphi Method, Techniques and Applications,* eds. Harold A. Linstone, and Murray Turoff (2002), 10.

[40] Muhammad Imran Yousuf, "Using Experts' Opinions Through Delphi Technique," *Practical Assessment Research & Evaluation* 12, no. 4 (2007): 2.

[41] Leire San-Jose, and José Retolaza, "Is the Delphi Method Valid for Business Ethics? A Survey Analysis," *European Journal of Futures Research* 4, no. 19 (2016): 1,12.

[42] Megan Grime, and George Wright, "Delphi Method," in *Wiley StatsRef: Statistics Reference Online,* ed. Paolo Brandimarte, Brian Everitt, Geert Molenberghs, Walter Piegorsch, and Fabrizio Ruggeri (New York: John Wiley & Sons Inc., 2016), 2, 3; Chia-Chien Hsu, and Brian A. Sandford, "The Delphi Technique: Making Sense of Consensus," *Practical Assessment, Research, and Evaluation* 12, no. 10 (2007): 2; Linstone and Turoff, "Introduction," 65; San-Jose, and Retolaza, "Is the Delphi Method Valid," 3, 5; Yousuf, "Using Experts," 1, 3.

[43] Mohammed Alyami, *Modification and Adaptation of the Program Evaluation Standards in Saudi Arabia* (PhD diss., Western Michigan University, 2013), 43: 10-20 experts; Grime, Wright, "Delphi Method," 2: 5-20 experts at the most; San-Jose, and Retolaza, "Is the Delphi Method Valid," 6, 7; 10-20 experts; Sekayi, and Kennedy, "Qualitative Delphi Method," 2757: 20-30 experts; Grime, and Wright, "Delphi Method," 2: 5-20 experts at the most.

to thirty persons. In our case, the Greek experts of educational assessment who were invited to participate were nineteen. All of them had at least a Master's degree in educational evaluation. The number of the experts who responded was sixteen. Nine of them were men and seven were women. Seven of the experts were teaching at a University (44%), three held a PhD, three were PhD candidates, and six had a relevant Master's degree. Nine of the experts, were also members of the Greek Society of Educational Evaluation (GSEE). Only one of the experts had also a scientific specialization in the field of Ethics.
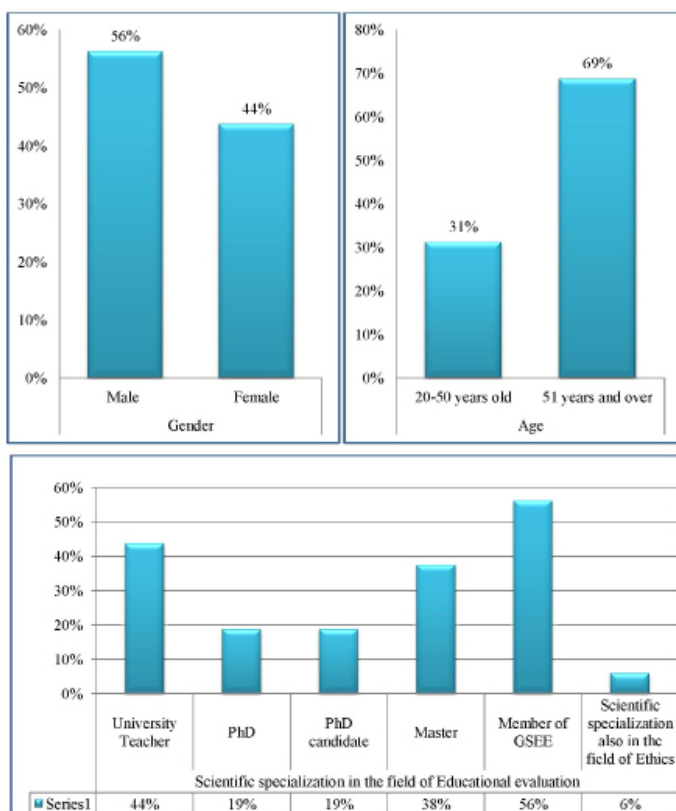

Figure 1. The panel of the Experts

In order to ensure even more the sample's heterogeneity, the first part of the given questionnaire included nine Likert Scale questions and one of multiple choice about ethical issues of educational assessment. All the experts agreed that *"Assessment as an action has great ethical importance."* Most of them agreed that *"The ethical quality of the evaluators and the evaluated has determining role in educational evaluation."* The experts seemed

to disagree on the following sentence: *"Scientific progress on techniques and tests of educational evaluation is sufficient to eliminate phenomena of ethical diversion."* Nine experts disagreed, three expressed neutrality, and four agreed. All but one disagreed on *"The use of unfair means by the evaluated, as a counterpoint for unfair evaluation."* Six experts agreed on the assertion that *"Assessment of learning, assessment for learning, improvement, accountability or excellence is ethically equivalent."* Among those who disagreed, there were five who believed that assessment for improvement is morally superior, three considered assessments for learning to be superior, one (assessment for effectiveness and one restated the initial item, claiming that morally superior is assessment "as and for learning." Twelve experts agreed that *"In Greece, there are ethical as well as unethical conducts concerning educational evaluation."* Two of the experts disagreed, and the rest two didn't express agreement or disagreement. The next sentence divided the experts. Eight out of sixteen believed that *"The evaluated who fails, should endure the consequences of his actions instead of being treated with clemency,"* whereas five believed the opposite and three avoided expressing any opinion. Eleven experts agreed that *"Assessment as an action reflects the ethical level of its society."* Only one disagreed, and five kept a neutral stance. Six experts believed that *"It is very difficult for assessment to work ethically and with meritocracy in a society of ethical crisis, no matter the number of the implemented ethical valves."* Seven disagreed and three neither did they agreed, nor they disagreed. Finally, experts thought that *"To be ethical as an evaluator or an evaluated is equivalent..." "chiefly to follow ethical standards, then to be ethical himself and final to be lawful,"* (four experts,), *"chiefly to be ethical himself, then to follow ethical standards and finally to be lawful,"* (five experts, and *"to all the above equally"* (seven experts) [see Figure 2 on the next page].

The process[44] of the *Delphi Method* is quite simple and includes a series of rounds, each one of which aims for the highest consensus among the panel experts. At first, a questionnaire is created based on the opinions of the experts on the given issue. Sometimes, like in this case, the questionnaire is created by the researcher himself. Unfortunately, the combination of specialization on both fields, Educational Assessment and Ethics, is very rare in Greece. Consequently, the initial sentences could have been proven scientifically weak, superficial or arbitrary, unless grounded in sound theoretical knowledge, a task that I undertook for my inquiry. Nevertheless, it is very hard to distinguish between the Principles which are my own creation, like the "*The Principle of the Inverted Pyramid,"* and those which are based on previously existing Principles, because in both cases the fermentation of ideas, proposals and beliefs was

---

[44] Alyami, "Modification and Adaptation," 43-44; Grime, and Wright, "Delphi Method," 3; Hsu, and Sandford, "The Delphi Technique," 2-3; Yousuf, "Using Experts," 2.

deep and continuous. Then, the experts are asked to express their agreement or disagreement, to rephrase or to make any other corrections on each item of the questionnaire. The items of high consensus – the higher, the better – are considered to have accomplished their goal and they are not subject to
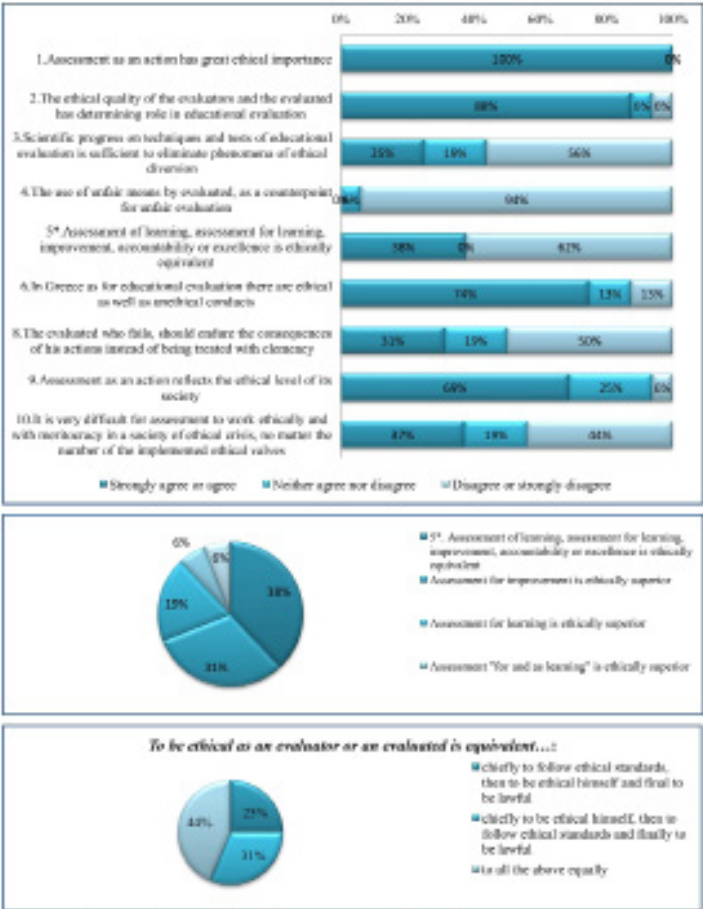

Figure 2. The heterogeneity of the experts through their opinions

further processing. The rest are rephrased according to the feedback of the experts and sent back to them for two or three times, until they score high or at least higher consensus. If they fail to gain consensus, they are recorded separately from the results of the research. The duration of all this process is about 30 to 45 days. In our case, it lasted 32 days.[45] At first, on the first round, a questionnaire of *25 Ethical Principles* was given to eleven experts in printed form, during the 2nd Scientific Conference of the Greek Society of

---

[45] 11th of May-12th of June 2018.

Educational Evaluation (11th-13th May, 2018), and an electronic version was sent to eight experts through email. The experts were asked to express their agreement or disagreement on each *Ethical Principle* or to rephrase them. The printed form was answered on the spot by all experts. The electronic form was answered by five experts. Twenty two *Ethical Principles* reached a high consensus, above 75%. A new *Ethical Principle* was proposed, and four *Ethical Principles* were restated according to the corrections of the experts. So, on the second round five new or restated *Principles* were sent, in electronic form through email. Finally, due to the high consensus on the first round and to the significant decrease of the participants on the second round[46] there was no need for a third round.

In the *Delphi Method*, a researcher aims for *consensus,*[47] unanimity of opinions. Nevertheless, a percentage of 70 to 80 is considered sufficient by most researchers, especially if the proportion increases from one round to the next. In our case, eleven *Ethical Principles* reached absolute consensus, seven very high (94%), one 88%, two 81%, two 75%, one 73% and one 31% from the first round. On the 2nd round, where the participation was lower – only ten experts participated – two *Principles* reached consensus of 100%, one 90%, one 80% and one 70%.

Taking into consideration all the above, it is quite obvious that it takes a lot of effort in the *Delphi Method* to have *results* that are something more than just opinions. In other words, it is researcher's integrity and determination to stick to the process and to apply the protocol of this technique, it is the willingness to choose the right persons, to let them express themselves freely and to respect their point of view, and, finally, it is one's devotion to serve his goal through the creation of a collective judgment that can guarantee that those *results* are not only valid and reliable but that they can also go beyond subjective beliefs and, therefore, be widely accepted. All the following *Results,* namely the *Ethical Principles*, should be regarded in this perspective.

III. Results

The Ethical Principles[48] *of Educational Evaluation* that came up as a result of

---

[46] On the second round ten experts participated instead of the initial sixteen. This perhaps reveals lack of scientific culture in Greece or loose commitment among the experts to the terms of participation in the Delphi Method, a problem that some researchers overcome through the remuneration of experts. The main reason for not choosing such a solution is the possible negative effect on experts expressing freely their opinions.

[47] Hsu, and Sandford, "The Delphi Technique;" Linstone, and Turoff, "Introduction," 22; Sekayi, and Kennedy, "Qualitative Delphi method," 2756. Alyami, "Modification and Adaptation," 48-49, sets 70% as a minimum consensus rate.

[48] In order to avoid misinterpretations, I have followed the form of Evaluation Standards. Each
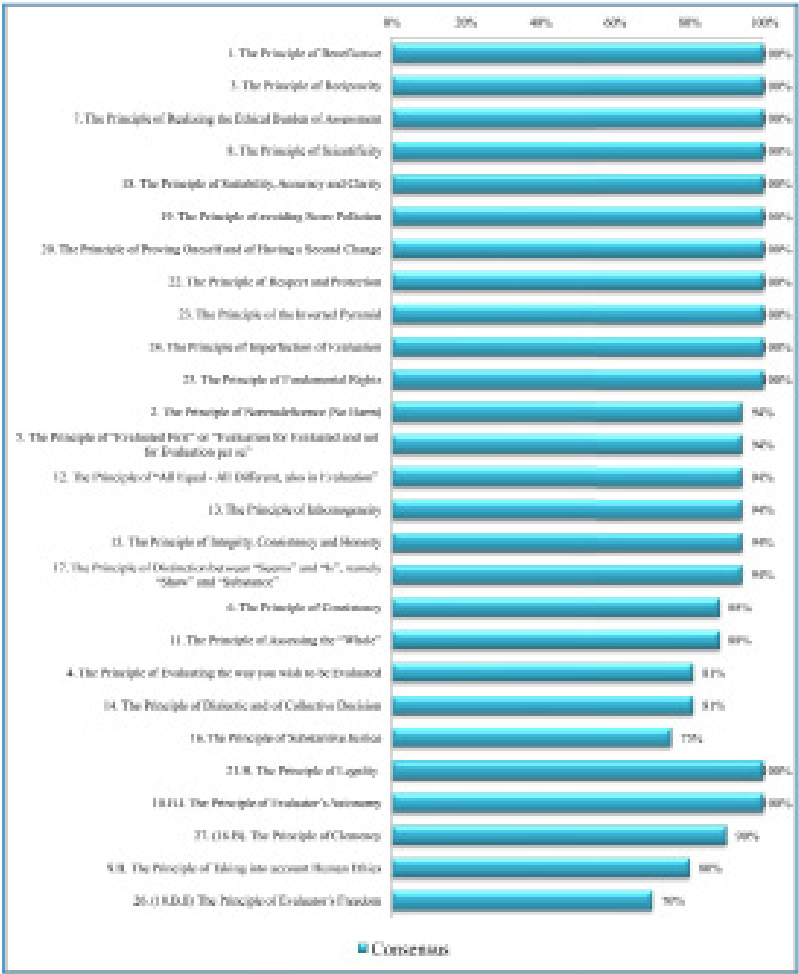
the research are[49]:

Figure 3. Percentage of Consensus to the given Ethical Principles

a. Principles of 100% of consensus, on the 1st round.

1. The Principle of Beneficence. Educational evaluation should benefit

---

Ethical Principle that was included in the questionnaire given to the experts was accompanied by an explanatory sentence.

[49] The Principles are listed according to the degree of consensus, namely from the highest of the 1st round, to the lowest. The number on the front corresponds to the number of the Principle on the questionnaire of the 1st or the 2nd round. The letter "A" or "B" also refers to the 1st and 2nd round where necessary.

the greatest possible number of people, including the evaluated.

3. The Principle of Reciprocity. Those who evaluate should also be evaluated generally and especially on the way they evaluate.

7. The Principle of Realizing the Ethical Burden of Assessment. The greater the consequences of an evaluation, the greater the necessity for the involved parts to practice it ethically without deviations or discounts.

8. The Principle of Scientificity. Evaluation should fulfill the requirements of validity, reliability, objectivity, discrimination, and practicality, and to be exercised by evaluators who truly have the necessary training and knowledge.

18. The Principle of Suitability, Accuracy and Clarity. Assessment, in all its stages (from the initial briefing to the announcement of the results), should be suitable, and its phrasing should be accurate and fully comprehensible by all stakeholders, in order for the results to be applicable and not to mislead or trap the evaluated, their parents or deliberately some of the involved persons.

19. The Principle of avoiding Score Pollution. Every effort must be made in order to eliminate factors irrelevant to the evaluation that may alter its results, e.g. personal relationships between the evaluator and the evaluated. Nevertheless, this does not mean that we are allowed to disregard other factors that should be taken into consideration, e.g. the temporary or irreversible impact of a disease on somebody's (student or teacher) performance.

20. The Principle of Proving Oneself and of Having a Second Chance. The student should be given the chance to prove that they possess the evaluated knowledge or skills, or that they have adopted the expected attitudes and behaviors. Alongside, prediction must be made for a second chance as a counterbalance for emergencies that may alter the student's image, and as a proof of improvement after the initial evaluation and the following feedback.

22. The Principle of Respect and Protection. The evaluated should be treated with respect. Moreover, evaluation should incorporate safety valves that will protect all parties involved and offer the chance of objection and appeal in cases of feeling wronged or offended. Finally, the evaluated should under no circumstances be treated as guinea pigs.

23. The Principle of the Inverted Pyramid. Those at the base of the pyramid who bear the weight of evaluations should not pay the price of the failures of those above them in the pyramid who evaluate them.

For instance, students should not "pay" through their evaluation for the incapability of their teachers, and teachers should not pay the price for deficiencies in logistics infrastructure.

24. The Principle of Imperfection of Evaluation. Evaluation is subject to the unavoidable errors, subjectivities, and deficiencies of the evaluators, which should be identified, recognized and taken into account, instead of being concealed, covered up or ignored.

25. The Principle of Fundamental Rights. Assessment should be consistent with the globally established and recognized human and child rights, the Constitution of its country, to respect their privacy, to be impartial and to assure that there will not be deception, physical, emotional, or psychological abuse or manipulation of the evaluated etc.

b. Principles of 94% of consensus, on the 1st round.

2. The Principle of Nonmaleficence (No Harm).[50] Educational evaluation should at least assure that nobody is harmed, if not benefiting people, including the evaluated.

5. The Principle of "Evaluated First" or "Evaluation for Evaluated and not for Evaluation per se." Assessment should, among the involved persons, serve chiefly the evaluated, and not political or scientific goals, like the promotion of products or methods of evaluation etc. or the professional ambitions of the evaluators.

12. The Principle of "All Equal - All Different, also in Evaluation." Assessment should show respect to the diversity of the evaluators or the evaluated, due to their cultural or religious identity, ideology, political conviction, social or economical origin, sexual orientation, gender, physical or mental retardation or particularity etc. and should act in their favor, but in a way that does not offend the rest of the people or be unfair for them.

13. The Principle of Inhomogeneity. Assessment should not be the same in all cases. It must vary according to the role, the age or the position of the evaluated in education etc. The possibility of evaluating on the same terms or of using the same evaluative techniques to evaluate educational executives and students is not acceptable.

15. The Principle of Integrity, Consistency and Honesty. All the involved parts of an evaluation (evaluated, evaluators etc.) should be determined to participate showing moral integrity and respect the terms of the

---

[50] This Principle had 100% consensus between those who answered. One of the experts did not respond, perhaps inadvertently.

evaluative process, even if they notice infractions of the rules by other involved persons.

17. The Principle of Distinction between "Seems" and "Is," namely "Show" and "Substance." Assessment should assure by all means that its results are not fictitious or superficial and that they represent the true substance of people, in order not to allow some of the evaluated to present themselves as superior or better than they truly are, using often unfair means, e.g. acquaintances, social status etc.

c. Principles of 88% of consensus, on the 1st round.

6. The Principle of Consistency.[51] Evaluation should be consistent, compatible, and attuned to its goals and its framework, in order to conform to the knowledge of the evaluated and not to surprise them. Moreover, it should not favor those who "possess" knowledge dishonestly (e.g. by cheating, or though shadow education) over the rest.

11. The Principle of Assessing the "Whole". Assessment should aim to the full possible image, namely —if possible— to cover the whole, to include all parts, all evaluative aspects and not to be fragmentary.

d. Principles of 81% of consensus, on the 1st round.

4. The Principle of Evaluating the way you wish to be Evaluated. Anyone who establishes or exercises assessment should act in a way similar to the one he would claim to be evaluated, e.g. clemency for clemency, severity for severity, and if somebody does not wish to be evaluated himself, he should also not evaluate others.

14. The Principle of Dialectic and of Collective Decision. Each time Ethical Principles fail to deal with a dilemma or a conflict that arises during evaluation process, the evaluators should a) rely on their personal values and even knowledge and b) ask for the assistance and opinion of their colleagues - evaluators if they have opposite views and, in case of deadlock, decisions should be made collectively.

e. Principle of 75% of consensus on the 1st round, acceptable due to precedence but restated in a way that a new Principle (the 27th) has emerged.

16.A. The Principle of Substantive Justice. In order for assessment to be fair, people of similar characteristics should be treated in a similar way, and people who differentiate themselves from others on some feature should be treated differently.

27. (16.B). The Principle of Clemency. In case of doubt about the

---

[51] The consensus of the principle per se was 94%. The percentage 88% had to do with small reservations as for the accompanying sentence.

fairness of evaluation, due e.g. to legal gap, ambiguity of an issue or unpropitious conditions during the exam, it is preferable for evaluators to act with clemency then severity. (90% on the 2nd round).

f. Principle of 75% of consensus on the 1st round, acceptable due to precedence, but restated in a way that an improved version of the Principle has emerged.

21.A. The Principle of Legality. Assessment should be exercised lawfully. If the law conflicts with the ethics of the evaluator or the Ethical Principles of Assessment, the evaluators are legitimized to "disregard it quietly" and they have to take action for its "correction."

21.B. The Principle of Legality. Assessment should be exercised lawfully. If the law conflicts with the ethics of the evaluator or the Ethical Principles of Assessment, the evaluators should express their disagreement and take action for its "correction." (100% on the 2nd round).

g. Restated Principles due to inadequate consensus.

9.A. The Principle of Taking into account Human Ethics. Assessment should neither overestimate nor underestimate the ethical quality of the evaluators and the evaluated and their impact on the process and the results of evaluation. Assessment should also take into consideration the fact that the final judge of adopting or not adopting and of keeping or disregarding the Ethical Principles is each evaluator or evaluated. (73% on the 1st round)

9.B. The Principle of Taking into account Human Ethics. Assessment should take into consideration that is up to the evaluators and the evaluated to comply with the rules of an evaluative process and that, among them, there might be people who would not obey the set rules, a fact that might have a negative impact on the results of evaluation. (70% on the 2nd round or 80% under certain conditions)

h. Principle of very low consensus that led to two new versions on the 2nd round.

10.A. The Principle of the Autonomy of the Worthy Evaluator. Assessment should leave space to the evaluators who honor their role to act on their own, to decide and work freely for the sake of their evaluated, even if this comes to conflict with any of the Ethical Principles. (31% on the 1st round)

10.B.I. The Principle of Evaluator's Autonomy. Evaluators should be able to do their job autonomously, uninfluenced of pressures, interventions, and interests. (100% on the 2nd round)

10.B.II. The Principle of Evaluator's Freedom. Assessment should, (in

specific cases, like in class but not in standardized testing) leave space to the evaluators to act on their own, to decide and work freely for the sake of their evaluated, even if this means small deviations from the letter of the law (70% or 80% on the 2nd round).[52]

IV. Discussion

Some may claim that in Greece we do not actually need Ethical Principles to improve educational evaluation. Instead, we can deploy fundamental virtues of our nation, like "*diakrisis*." However, setting *Ethical Principles* seems to be a persuasive answer to the numerous moral issues, problems, and dilemmas that both evaluators and evaluated face in everyday practices of Educational Evaluation. Nevertheless, the phrasing and choice of the right *Ethical Principles*, namely of those that cover all the fields and pivots of Education in modern Greece may be the first step, but not the final one. All parties involved in educational assessment must be informed about this new ethical framework and accept it or reject it in practice. From this point of view, this research is a solid starting point based on theoretical and scientific data that can contribute to a more ethical practice of Educational Evaluation, but there is certainly much more to be done. After all, through all this research we have dealt with all different variations of evaluation. We have seen evaluators of, so called, high or low ethical quality and, vice versa, evaluated who adopt ethically acceptable or morally unacceptable evaluative behaviors. So, it will take a lot of deliberation and a lot of fermentation before we are led to a commonly accepted ethical framework for evaluation, in which several additions and subtractions can be made. But what remains non-negotiable is the absolute need for all moral choices to be based on the one and only *Ethical Principle* that can summarizes all others: "*Evaluation must be done from a human being to a human being*." Evaluation needs to be humane.

References

American Evaluation Association. "Guiding Principles for Evaluators." Last modified August 18, 2018. https://www.eval.org/About/Guiding-Principles.

Alyami, Mohammed. "Modification and Adaptation of the Program Evaluation Standards in Saudi Arabia." PhD diss., Western Michigan University, 2013. https://scholarworks.wmich.edu/dissertations/157.

Babiniotis, Georgios. *Dictionary of Modern Greek*. Athens: Lexicology Center, 1998. [In Greek].

---

[52] The opinion (agreement or disagreement) of one expert was not stated clearly.

Bett, Richard. "The Sophists and Relativism." *Phronesis* 34, no. 2 (1989): 139-169.

Bochenski, Jozef Maria. *Europäische Philosophie der Gegenwart*. Introduced and translated by Christos Malevitsis. Athens: Dodoni Publications, 1985.

Coriu, Daniel. "The Path From the Natural to the Spiritual Diakrisis Through Askesis in the Views of St. Apostle Paul and Elder Joseph the Hesychast." *International Journal of Orthodox Theology* 9, no. 4 (2018): 152-175.

DeWitt, Norman Wentworth. *Epicurus and His Philosophy*. University of Minnesota Press, 1954.

Dimitropoulos, Efstathios. *Educational Evaluation. The Evaluation of Education and the Educational Project. Part I.* Athens: Grigoris Publications, 2007. [In Greek].

Estaji, Masoomeh. "Ethics and Validity Stance in Educational Assessment." *English Language and Literature Studies* 1, no. 2 (2011): 89-99.

Gipps, Caroline. *Beyond Testing: Towards a Theory of Educational Assessment.* London, Washington: The Farmer Press, 2003.

Green, Susan, Robert Johnson, Do-Hong Kim, and Nakia Pope. "Ethics in Classroom Assessment Practices: Issues and Attitudes." *Teaching and Teacher Education* 23 (2007): 999-1011.

Grime, Megan, and George Wright. "Delphi Method." In *Wiley StatsRef: Statistics Reference Online,* edited by Paolo Brandimarte, Brian Everitt, Geert Molenberghs, Walter Piegorsch, and Fabrizio Ruggeri, 1-6. New York: John Wiley & Sons Inc., 2016.

Hamalis, Perry T. "Eastern Orthodox Ethics." *International Encyclopedia of Ethics,* 2013.

Hudson, William Donald. *Ethical Intuitionism*. London: Macmillan, 1967.

Huemer, Michael. *Ethical Intuitionism*. New York: Palgrave Macmillan, 2005.

Hsu, Chia-Chien, and Brian A. Sandford. "The Delphi Technique: Making Sense of Consensus." *Practical Assessment, Research, and Evaluation* 12, no. 10 (2007):1-8.

Joint Committee on Testing Practices. "Code of Fair Testing Practices in Education." Washington, 2004. https://www.apa.org/science/programs/testing/fair-testing.pdf.

Kapsalis, Achilleas, and Nikos Chaniotakis. *Educational Evaluation.* Thessaloniki: Kyriakidis Bros - Publications SA, 2015. [In Greek].

Kaptein, Muel. *Ethics Management: Auditing and Developing the Ethical Content of Organizations*. Dordrecht: Kluwer Academic Publishers, 1998.

Kassotakis, Michalis. *Assessing Student Performance*. Athens: Grigoris Publications, 2013. [In Greek].

Keniston, Kenneth. "Morals and Ethics." *The American Scholar* 34, no. 4 (1965): 628-632.

Konstantinou, Charalampos. *The Evaluation of Student Performance as Pedagogical Logic and School Practice.* Athens: Gutenberg, 2007. [In Greek].

Kunnan, Antony John. *Fairness and Validation in Language Assessment: Selected Papers From the 19th Language Testing Research Colloquium*, *Orlando, Florida*. Cambridge, UK: Cambridge University Press, 2000.

Lampe, Kurt. *The Birth of Hedonism: The Cyrenaic Philosophers and Pleasure as a Way of Life*. Princeton, Oxford: Princeton University Press, 2015.

Linstone, Harold A. and Murray Turoff. "Introduction." In *The Delphi Method, Techniques and Applications,* edited by Harold A. Linstone, and Murray Turoff, 3-12. 2002. https://web.njit.edu/~turoff/pubs/delphibook/delphibook.pdf.

Lyons, Susan, and Erika Hall. "The Role of the Standards for Educational and Psychological Testing in Establishing a Methodology to Support the Evaluation of Assessment Quality." *Center for Assessment*. 2016. https://www.nciea.org/sites/default/files/publications/Standards_in_Assessment_Quality_Eval_042016.pdf.

Meyendorff, John. *Byzantine Theology: Historical Trends and Doctrinal Themes.* New York: Fordham University Press, 1974.

National Council on Measurement in Education (NCME). "Code of Professional Responsibilities in Educational Measurement." Philadelphia, 1995. https://www.ncme.org/resources/library/professional-responsibilities.

Newman, Dianna, and Robert Brown. *Applied Ethics for Program Evaluation*. London: Thousand Oaks, CA: Sage Publications, 1996.

Nussbaum, Martha C. "Mill between Aristotle & Bentham." *Daedalus* 133, no. 2 (2004): 60-68.

Principles for Fair Student Assessment Practices for Education in Canada. Edmonton, Alberta: Joint Advisory Committee. 1993. https://www.wcdsb.ca/wp-content/uploads/sites/36/2017/03/fairstudent.pdf.

Pring, Richard. *The Life and Death of Secondary Education for All*. New York: Routledge, 2013.

San-Jose, Leire, and José Retolaza. "Is the Delphi Method Valid for Business Ethics? A Survey Analysis." *European Journal of Futures Research* 4, no. 19 (2016): 1-15.

Satris, Stephen. *Ethical Emotivism*. Dordrecht: Martinus Nijhoff, 1987.

Schmeiser, Cynthia. "Ethics in Assessment." *EDO-CG-95-23, ERIC Digest,* USA, 1995. https://www.counseling.org/resources/library/ERIC%20Digests/95-23.pdf.

Sekayi, Dia, and Arleen Kennedy. "Qualitative Delphi Method: A Four Round Process With a Worked Example." *The Qualitative Report* 22, no. 10 (2017): 2755-2763.

Simons, Helen. "Ethics in Evaluation." In *Handbook of Evaluation, Policies, Programs and Practices,* edited by Ian Shaw, Jennifer Greene, and Melvin Mark, 243-265. London: Sage Publications, 2006.

Singer, Peter. "Ethics." *Encyclopedia Britannica*. Accessed February 2, 2021. https://www.britannica.com/topic/ethics-philosophy.

Stratton-Lake, Philip, ed. *Ethical Intuitionism: Re-Evaluations*. New York: Oxford University Press, 2002.

Viltanioti, Irini-Fotini. "Porphyry's Letter to Marcella: A Literary Attack on Christian Appropriation of (Neo-)Pythagorean Moral Wisdom?" In *Pythagorean Knowledge from the Ancient to the Modern World: Askesis, Religion, Science*, edited by Almut-Barbara Renger, and Alessandro Stavru, 163-184. Wiesbaden: Harrassowitz Verlag. 2016.

Warrender, Howard. "Hobbes's Conception of Morality." *Rivista Critica Di Storia Della Filosofia* 17, no. 4 (1962): 434-49.

Wood, Allen W. *Kant's Ethical Thought*. Cambridge: Cambridge University Press, 1999.

Yousuf, Muhammad Imran. "Using Experts' Opinions Through Delphi Technique." *Practical Assessment Research & Evaluation* 12, no. 4 (2007): 1-8.

Death Anxiety, Immortality Projects and Happiness: A Utilitarian Argument Against the Legalization of Euthanasia

Donovan van der Haak

Tilburg University, Netherlands
E-mail address: donovanvanderhaak@hotmail.com
ORCID iD: http://orcid.org/0000-0003-0323-1640

Abstract

The current, utilitarian debate on the relation between euthanasia and happiness focusses primarily on the subject of dying patients. Where some utilitarians stress how euthanasia may relieve suffering in the process of dying, others emphasize the importance of respecting the autonomy of others to make decisions like these themselves. However, less attention has been paid to how legalizing euthanasia may impact the human decision-making processes of those still in a healthy and mentally sound state. This paper aims to shed light on this relatively underdeveloped subject within utilitarian theory. In particular, I focus on euthanasia's most contested form: active, voluntary euthanasia. I draw on Ernst Becker, who argues that moderate death anxiety stimulates people to work on 'immortality projects,' decisions that help them cope with the concept of death. Subsequently, I draw on several studies to defend the notion that immortality projects are indirectly conducive to happiness because they stimulate healthy decisions and long-term, human progress. Additionally, immortality projects counterbalance decisions that are based on an excessive drive for short-term pleasure. As euthanasia can make dying less painful, I argue it may diminish death anxiety to significant degree, and thereby also an incentive to work on immortality projects. This brings me to the conclusion that legalizing euthanasia is problematic from a utilitarian point of view, considering the observation that immortality projects are indirectly conducive to happiness.

Keywords: *death anxiety; euthanasia; immortality projects; The Denial of Death; utilitarianism*

I. Introduction

Active, voluntary euthanasia (Greek for 'good death') is a hugely divisive issue. From a utilitarian perspective, the right act to perform is the one that will bring about the best balance of pleasure over pain, i.e. the greatest happiness for the greatest number.[1] Some utilitarians therefore may argue that, as euthanasia has the capacity to alleviate unnecessary pain, it is prima facie the right thing to do. Conversely, the purpose of this paper is to demonstrate that legalizing euthanasia also indirectly demotivates decisions that are conducive to happiness. Becker[2] argues that moderate death anxiety stimulates people to work on 'immortality projects,' decisions that help them cope with the concept of death. Subsequently, I defend the notion that immortality projects are indirectly conducive to happiness because they stimulate healthy decisions and long-term, human progress. As euthanasia can make dying less painful, it diminishes death anxiety and thereby an incentive to work on immortality projects. For the sake of this paper, I exclude providing justification of utilitarianism as a moral framework. Although a wide variety of moral views on euthanasia are present (such as deontological and virtue-ethicist considerations), the purpose of this essay is solely to add a utilitarian consideration to this debate. I therefore do not argue that this this paper will conclusively settle this rather complex issue; it merely aims to provide one utilitarian argument against euthanasia.

Firstly, I elaborate briefly on some of the present, utilitarian arguments for and against euthanasia to situate my argument in the current debate (Section 1). In Section 2, I provide evidence for the notion that moderate death anxiety is prevalent among many people and discuss the role of prospecting suffering in dying in relation to human decision-making processes. Becker argues this impact gives rise to immortality projects, decisions that help people cope with the concept of death. Section 3 aims to show that immortality projects are conducive to long-term happiness because they are progress-driven, healthy and capable of counterbalancing decisions that are based on an excessive drive for short-term pleasure. Section 4 shows that the possibility of euthanasia can reduce death anxiety, as euthanasia can take away suffering in the process of dying. I conclude that, as moderate death anxiety is an important incentive to make decisions that are conducive to happiness (immortality projects), euthanasia's capacity to reduce suffering in dying may have counter-productive consequences. Lastly, I will consider some potential objections to my argument (Section 5).

---

[1] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (Ontario: Batoche Books, 2000), 225.

[2] Becker, 208-210.

II. Utilitarian arguments for and against euthanasia

There are different views on different forms of euthanasia. Here, I will only address the most contested form: active, voluntary euthanasia. Active, voluntary euthanasia is an act or set of acts to end a patient's life at request of this same patient.[3] From a utilitarian perspective, there are two important arguments that support active, voluntary euthanasia. Firstly, euthanasia has the capacity to alleviate pain that people experience nearing the end of their lives. As hedonistic utilitarians seek to find the best balance of pleasure over pain, any option to alleviate unnecessary pain is preferable. Secondly, some utilitarians argue we should respect the autonomy of people to make their own decisions if this does not harm others. John Stuart Mill, for instance, argues that:

> the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant.[4]

In line with Mill, utilitarians could argue that, if euthanasia only harms the individual, there is no reason to forbid it. Although the goal of utilitarianism is essentially to provide happiness (not autonomy), Mill argues the harm principle is nonetheless an essential principle exactly because such autonomy is contributory to maximizing happiness.

Conversely, others claim that respecting individual autonomy for some people may have negative consequences for others (e. g., marginalized groups). Boer[5] estimates that family pressure influences approximately 20% of people applying for euthanasia. From this perspective, legalizing euthanasia may threaten to pressure marginalised individuals to opt for euthanasia, damaging society by indirectly degrading the value of life within some communities.[6] This argument is primarily focussed on the social limits of autonomy, drawing on the dangerous social consequences of legalizing euthanasia. Although Mill's harm principle protects individual autonomy, it also allows restrictions

---

[3] Singer, 526.

[4] John Stuart Mill, *On Liberty* (Boston: Ticknor and Fields, 1859), 13.

[5] Theo Boer, "Report on Euthanasia and Physician Assisted Suicide in the Netherlands," 5, https://www.pthu.nl/Over-PThU/Organisatie/Medewerkers/t.a.boer/downloads/2016-boer-south-africa-affidavit-euthanasia-netherlands.pdf

[6] William Grey, "Right to Die or Duty to Live? The Problem of Euthanasia," *Journal of Applied Philosophy* 16, no. 1 (1999): 25.

on individual freedom in cases where this freedom harms others. As this argument posits that the legalization of euthanasia indeed has negative, social consequences for others, Mill's harm principle does not counter this objection to euthanasia. Similarly, my argument is primarily focussed on the negative, social consequence of the legalization of euthanasia, namely its impact on human decision-making processes (hereafter HDMP). Differing from the current debate on the relation between euthanasia and HDMP, which focusses primarily on the decision to opt for euthanasia or not, this paper sheds light on the impact on HDMP when people are still in a healthy and mentally sound state; a relatively underdeveloped subject. Importantly, my argument addresses the legalization of euthanasia, as opposed to its practice. I argue that legalizing active, voluntary euthanasia gives people the opportunity to reduce suffering in dying, impacting HDMP as it lowers death anxiety. As we will see later, the opportunity to die without pain threatens to have a negative, sociological impact, even if people do not end up practicing it nearing the end of their lives. Alison Davis,[7] a disabled person who was granted permission for euthanasia, argues that pain and suffering do not necessarily make life worthless as there is value in these experiences. In line with Davis, I argue that there is even utilitarian value in the presumption that dying will be painful. The argument is structured as follows:

P1: Death anxiety significantly impacts HDMP, giving rise to immortality projects.
P2: Immortality projects are decisions that are conducive to happiness.
P3: Euthanasia can take away suffering in dying.
P4: Taking away suffering in dying diminishes death anxiety, thereby demotivating immortality projects.
C: Euthanasia demotivates decisions that are conducive to happiness.

III. Death anxiety and decision-making

Before explicating on how exactly death anxiety impacts HDMP, I will first briefly demonstrate that the impact of death anxiety on our lives is of a considerable degree. This is important, because the extent of this impact influences the significance of my argument. Firstly, there is scientific consensus that moderate death anxiety is a relatively common phenomenon. Several studies find that the majority of individuals are afraid of death and dying,

---

7 Alison Davis, "A Disabled Person's Perspective on Euthanasia," *Disability Studies Quarterly* 24, no. 3 (2004): 1, https://dsq-sds.org/article/view/512/689.

albeit moderately.[8] These studies also reveal that experiencing moderate death anxiety is not exclusively prevalent among elderly, but also among young people. Inspired by Freudian psychoanalysis, Ernest Becker[9] wrote *The Denial of Death*, in which he established that death anxiety comes naturally to those people that find death and dying unacceptable; it is a response to our basic survival mechanism. Both studies and Becker suggest that death anxiety should not be understood only as an abnormal or chronic fear of death, but for the most part as a more moderate, widespread aversion to death and dying. This is of vital importance, because it implies that the impact of death anxiety on HDMP is not limited to only a small group of people with severe anxiety (as we will see later).

Secondly, as my argument is primarily focussed on relieving worries about the process of dying (and not the inevitability of death itself), it is also important to see whether this specific worry plays a significant role in constituting death anxiety. There is good reason to believe that people in fact worry more about the dying process than about the end of life itself. One study states: "The dying process is more relevant [...] than the actual thought of death (all [the participants] were worried of dying with discomfort)."[10] Admittedly, it is unclear how death anxiety is exactly constructed, and some individuals may experience more fear of death itself than the process of dying (and vice-versa). Whether worries about the process of dying are actually more frequent and thus more significant than the inevitability of death in relation to death anxiety is, however, not the issue. At the very least, fear of the process of dying is still a crucial contributor to death anxiety (both in Becker's theory and terror management theory, to which I will refer later). Moreover, research on the effects of moderate death anxiety confirm that it substantially impacts HDMP.[11] For example, one study shows that moderate death anxiety impacts financial decision-making.[12] Other studies find that death anxiety can influence both political and religious beliefs.[13] It is, of

---

[8] Patricia Furer, and John Walker, "Death Anxiety: A Cognitive-Behavioral Approach," *Journal of Cognitive Psychotherapy* 22, no. 2 (2008): 167; Lisa Iverach, Ross Menzies, and Rachel Menzies, "Death Anxiety and Its Role in Psychopathology: Reviewing the Status of a Transdiagnostic Construct," *Clinical Psychology Review* 34, no. 7 (2014): 580; Gary Sinoff, "Thanatophobia (Death Anxiety) in the Elderly: The Problem of the Child's Inability to Assess Their Own Parent's Death Anxiety State," *Frontiers in Medicine* 4 (2017): 1.

[9] Becker, 1-8.

[10] Sinoff, 20.

[11] Iverarch, Menzies, and Menzies, 580.

[12] Timothy Ly, et al., "Death Anxiety and Financial Decision-Making in Aging: A Study from the Human Connectome Project Aging (HCP-A)," *Innovation in Aging* 3, no. 1 (2019): 907.

[13] Brian Burke, Spee Kosloff, and Mark Landau, "Death Goes to the Polls: A Meta-Analysis of Mortality Salience Effects on Political Attitudes," *Political Psychology* 34, no. 2 (2013): 183;

course, true that HDMP are not only impacted by death anxiety; other, more life-affirming, human drives also impact our decisions. Nevertheless, for now it suffices to have demonstrated that anxiety for dying impacts HDMP to a considerable degree.

Having established that death anxiety significantly impacts HDMP, we can now discuss more specifically what this impact entails. Becker[14] claims that moderate death anxiety is a widespread phenomenon that stimulates people to start so-called immortality projects. As a response to our basic survival instinct, humans create a defence mechanism against the knowledge of their mortality. This gives rise to immortality projects, projects and life-decisions that help people cope with death by denying and delaying it. According to Becker, this process is necessary for our functioning in the world. The goals, passions, hobbies and other activities humans engage in are essentially strategies to cope with these worries. Becker's theory was later backed up by 'terror management theory,' an influential research programme in social psychology.[15] Terror management theory posits that the drive of individuals to achieve personal goals is in part motivated by the awareness of their mortality. For instance, the human will to have sex is not only constituted by a life-affirming desire to feel pleasure, but also by a desire to overcome our mortality through reproduction of our genes.[16] Similar to Becker, terror management theory recognizes that immortality projects are motivated by both an anxiety for death and suffering in dying. Thereby, they provide a foundation to link immortality projects to euthanasia's capacity to alleviate such suffering.

IV. Immortality projects and happiness

As we note from above, death anxiety and immortality projects have an important impact on HDMP. Reasoning from a utilitarian perspective, we must now ask ourselves: what is the relation between this impact and the maximization of happiness? I argue that immortality projects are of vital importance to happiness, because they stimulate healthy decisions, create social meaning and help societies progress. Firstly, as immortality projects aim to delay death and dying, it stimulates people to live more healthy lives.

Kenneth Vail, Jamie Arndt, Matt Motyl, and Tom Pyszcynski, "The Aftermath of Destruction: Images of Destroyed Buildings Increase Support for War, Dogmatism, and Death Thought Accessibility," *Journal of Experimental Social Psychology* 48, no. 5 (2012): 1069.

[14] Becker, 1-8.

[15] Sheldon Solomon, Jeff Greenberg, and Tom Pyszcynski, *The Worm at the Core: On the Role of Death in Life* (New York: Penguin Random House, 2015), 1-5.

[16] Richard Dawkins, *The Greatest Show on Earth*: *The Evidence for Evolution* (New York: Simon & Schuster, 2009), 54-68.

Worrying about one's death naturally leads to increased health awareness. In addition, the will to postpone death and dying ipso facto provides people with an incentive to live healthier lives.[17] More subtly, attempting to rescue humanity from the inevitability of death and dying, immortality projects motivate contributions to the development of curing (terminal) diseases. Such a stimulation of health can be conducive to happiness by making people feel more free from debility. As the ethicist Angner puts it: "health status is one of the most important predictors of happiness."[18] In addition, stimulating healthy decisions does not only increase our average happiness, but it also contributes naturally to longer, healthier lives. As a result, human quality-adjusted life years (a health parameter in utilitarian cost-benefit analyses) increase as people live longer and healthier lives.

A further crucial point is the fact that immortality projects 'deny' human mortality. This should not be understood as a delusional conviction that one will never die, but rather as events wherein the terror of death stimulates human beings to create and become part of long-term projects that can perceptually 'last eternally.' Terror management theory explains this more concretely, claiming that death anxiety guides the development of art, religion, language, economics and science.[19] To minimize the terror of our own mortality, people strive to sustain the belief that they can contribute to a meaningful universe. This drives individuals to become more goal-oriented, giving them the feeling that their lives have purpose by working on something ostensibly significant.[20] This is not limited to the individual. Immortality projects stimulate culturally rich and socially cohesive communities that collectively seek to find meaning as well. Indeed, communities provide their members with a meaningful worldview that helps them cope with death anxiety, thereby giving their members' lives purpose and meaning. The creation of meaningful, long-term projects undoubtedly contributes to happiness, not only considering it provides humans with a purpose in life, but also because these projects provide better living conditions for future generations, as they are focussed on creating a more purposeful, healthier world.

From the perspective of proponents of euthanasia, such decisions can also be motivated by more life-affirming, human drives. Death anxiety might play a significant part in constituting some important decisions,

---

[17] Russell Noyes, at al., "Hypochondriasis and Fear of Death," *The Journal of Nervous and Mental Disease* 190, no. 8 (2002): 503.

[18] Erik Angner, et al., **"**Daily Functioning, Health Status, and Happiness in Older Adults," *Journal of Happiness Studies* 14, no. 4 (2012): 1563.

[19] Solomon, Greenberg, and Pyszcynski, 1-5.

[20] Richard Tedeschi, and Lawrence Calhoun, "Posttraumatic Growth: Conceptual Foundations and Empirical Evidence," *Psychological Inquiry* 15, no. 1 (2004): 1-18.

but it is certainly not the case that we cannot make long-term, healthy decisions without a constant fear of death and dying. However, social psychologists suggest that immortality projects do have a unique role to play in stimulating such decisions. Terror management theory states that humans constantly face an internal conflict of death anxiety versus our basic desire to live.[21] One important argument for why life-affirming drives need to be counterbalanced by immortality projects is because diminished death anxiety can lead people to make decisions excessively based on short-term pleasure. Immortality projects thereby do not only stimulate health and progress, but also protect people from negative repercussions of short-term decisions. For example, death anxiety does not miraculously rid someone of a smoking addiction, but immortality projects *can* stimulate individuals to try and quit smoking more often in order to delay death and dying. It is exactly because immortality projects are, unlike more life-affirming human drives, uniquely focussed on either delaying or denying death, that they are of fundamental importance to help shape our decisions. Of course, this does not mean that people cannot make good decisions without having death anxiety. Nevertheless, immortality projects do impact at least a significant amount of the important decisions we make. As long as the impact of these decisions are indeed generally conducive to happiness, we should take the utility of prospecting suffering in dying seriously. Notably, my use of Becker's theory herein deviates from Becker's own views, and I do not mean to suggest that Becker argues directly, nor indirectly, against the legalization of euthanasia. Becker himself even connects the existence of immortality projects to human conflict. Notwithstanding Becker's own views, the existence of immortality projects can nonetheless be considered as being conducive to happiness due to its counterbalancing capacity to stimulate healthy decisions and long-term, human progress. In the next section, I will extrapolate this premise to demonstrate an incompatibility between happiness arising from immortality projects and the legalization of euthanasia.

V. Euthanasia as demotivation

So far, I have argued that death anxiety impacts HDMP by stimulating immortality projects. Subsequently, I argued that immortality projects are conducive to happiness because they stimulate human progress, counterbalance human drives towards short-term pleasure and stimulate healthy decisions. This brings us to the last two premises of my argument. The third premise states that euthanasia can take away suffering in dying.

---

[21] Sheldon Solomon, Jeff Greenberg, and Tom Pyszcynski, "A Terror Management Theory of Social Behavior: The Psychological Functions of Self-Esteem and Cultural Worldviews," *Advances in Experimental Social Psychology* 24, no. 93 (1991): 159.

This is a rather self-evident claim; euthanasia purposefully takes away suffering in dying by shortening the process of dying, focussing precisely on relieving intractable suffering.[22] More justification is needed, however, for the last premise, namely that the opportunity to take away suffering in dying actually diminishes death anxiety (thereby demotivating immortality projects). In section 2, I have already discussed how death anxiety gives rise to immortality projects and that the presumption of suffering in dying plays a vital role therein. Having established this connection, I will now defend the notion that euthanasia is indeed associated with relieving pain and that its legalization actually diminishes death anxiety. From a rational perspective, the legalization of euthanasia should already take away a part of the worry about the process of dying, as it is capable of significantly alleviating our suffering. It seems indisputable that the opportunity to have a 'good death' can mitigate anxiety of the process of dying. More importantly, as rational thought is not necessarily sufficient for diminishing anxiety, it is important to demonstrate that the legalization of euthanasia is already impacting human attitude. One study found a significant correlation between death anxiety and the attitude of people towards voluntary euthanasia, suggesting people found relief in the possibility of assisted suicide.[23] Moreover, cultural attitudes about suffering in dying are already gradually changing in countries where euthanasia is legalized. In an interview on the legalization of euthanasia in the Netherlands, Boer described this cultural change aptly as follows:

> We're getting used to euthanasia. Culturally, I'm concerned that [...] death is being portrayed as a brave solution to severe suffering. A culture of euthanasia undermines our capacity to deal with suffering, and that is very bad for society.[24]

Even proponents of euthanasia acknowledge this cultural shift. For instance, Penney Lewis, ethicist at King's College London and proponent of the legalization of euthanasia, claims that the more people understand euthanasia is an option for them, the more they perceive it as an opportunity to avoid

---

[22] N. M. Harris, "The Euthanasia Debate," *Journal of the Royal Army Medical Corps* 147, no. 3 (2001): 367-370.

[23] Gerald Devins, "Death Anxiety and Voluntary Passive Euthanasia: Influences of Proximity to Death and Experiences with Death in Important Other Persons," *Journal of Consulting and Clinical Psychology* 47, no. 2 (1979): 301.

[24] Harriet Sherwood, "A Woman's Final Facebook Message before Euthanasia: 'I'm Ready for My Trip Now...,'" *The Guardian*, March 17, 2018, https://www.theguardian.com/society/2018/mar/17/assisted-dying-euthanasia-netherlands.

hopeless suffering.[25] As euthanasia is a relatively new, 21st century possibility, its normalisation is of deep concern.

Conversely, one may still challenge this argument by pressing the potentially weak connection between relieving fear of the process of dying and a relieve of death anxiety in general. This is a crucial assumption on which the argument relies, as a weak connection may imply that legalizing euthanasia would not affect general death anxiety, in which case the argument becomes unsound. Although the exact, long-term impact of the legalization of euthanasia on death anxiety is, of course, a matter of speculation, we have seen that there are several indications that suggest the connection *is* strong. Firstly, both terror management theory and Becker include the process of dying in their description of death anxiety, thereby reaffirming its significance. Secondly, Boer's analysis demonstrates how significant cultural changes are already starting to develop as a result of the legalization of euthanasia in The Netherlands. As euthanasia is becoming normalized, people are starting to become familiar with the possibility of choosing between a diverse set of deaths to choose from. Although this does not directly show that death anxiety entirely vanishes as a result of the legalization of euthanasia, it does indicate a gradual acceptance of the concept of death. Whereas this familiarity is already occurring, Christopher de Bellaigue[26] notices that the long-term consequences of the legalization of euthanasia are only just becoming discernible. Furthermore, it is important to consider that fear of dying a painful death is an expression of death anxiety that calls for legalizing euthanasia in the first place. Timothy James observes that:

> For most people, dying at home isn't about autonomy, it's about dealing with the fear of dying [...]. The fear of dying in misery [...] is what is driving the assisted dying debate.[27]

There are thus good reasons to believe that legalizing euthanasia will actually relieve death anxiety, for it is one of the key reasons driving people to call for the legalization of euthanasia in the first place.

To reiterate, I do not mean to imply that, if euthanasia is legalized, death anxiety vanishes to such an extent that there is absolutely no motivation left

---

[25] Ibid.

[26] Christopher de Bellaigue, "Death on Demand: Has Euthanasia Gone Too Far?" *The Guardian*, January 18, 2019, https://www.theguardian.com/news/2019/jan/18/death-on-demand-has-euthanasia-gone-too-far-netherlands-assisted-dying.

[27] Bioethics Research Library, "Fear of Death Driving Push for Euthanasia, Says Medical Ethicist," *Bioethics News*, https://bioethics.georgetown.edu/2015/05/fear-of-death-driving-push-for-euthanasia-says-medical-ethicist/.

to work on immortality projects. Admittedly, the legalization of euthanasia will not completely eliminate death anxiety for everyone. Despite this fact, it has become clear that there is good reason to believe that the opportunity of euthanasia *can* diminish death anxiety to a considerable degree. As mentioned before, when death anxiety decreases, it demotivates decisions that help people cope with death and dying (immortality projects). I also provided several arguments based on Becker and terror management theory that suggest that these decisions are of vital importance to happiness. Consequently, if the aforementioned premises are true (for which I have tried to provide sufficient justification), it logically follows that euthanasia demotivates decisions that are conducive to happiness. In any case, the argument demonstrates that utilitarians should expand their view, and include healthy individuals as well when analysing the impact of legalizing euthanasia.

VI. Objection and response

To begin, total-view utilitarians perceive humans only as valuable in so far as they make possible the experience of pleasure and happiness. This view makes possible the so-called replaceability argument: as total-view utilitarianism aims to maximize the experience of happiness, independent of whether the beings experiencing this happiness already exist or not, we can replace beings who suffer severely with new beings.[28] In relation to euthanasia, the replaceability argument may imply that long, healthy and purposeful lives are unnecessary, because there is no incentive to focus on long-term happiness for existing individuals. This directly counters my premise that immortality projects are conducive to happiness, as individuals should instead focus exclusively on experiencing as much short-term pleasure as possible. Once negative repercussions start kicking in (e.g., smoking a lot and subsequently getting lung cancer), euthanasia can quickly end the suffering, followed by the creation of new life that replaces the person's role as a mere recipient of happiness and pleasure. However, there are several objections that can be made against the replaceability argument. Salt, for example, denounces total-view utilitarianism, arguing that it is nonsense to talk about happiness or unhappiness of that of which we can predicate nothing.[29] In order to maximize happiness, agents must first have the *terra firma* of existence to argue from. Similarly, Singer argues that possible people are replaceable but not actual people, because actual people can already conceive of their own future's existence.[30]

---

[28] Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1993), 121.

[29] Henry Salt, "The Philosopher and the Pig," *The Vegetarian* 9, no. 49 (1896): 585.

[30] Singer, *Practical Ethics*, 123-131.

For these reasons I denounce total-view utilitarianism and take on the so-called 'prior existence view': the view that we should maximize utility of those beings whose existence is already a given.[31] This view is consistent with my argument that we do need long, healthy and purposeful lives, as we should seek to maximize happiness for existing beings in the long run. As existing life cannot simply be replaced from this perspective, the negative repercussions of making too many decisions based on short-term pleasure threaten the maximization of happiness. Admittedly, this is a speculative conjecture based on the idea that happiness is more likely to be optimized for existing human beings if their social conditions are focussed on creating more healthy, extensive and purposeful lives. Conversely, one could object that decisions based on short-term pleasure are also of fundamental importance to happiness, even from a prior existence view of utilitarianism. This brings us to a different objection, namely that, in order to optimize pleasure over pain for existing beings, we need to allow for decisions based on short-term pleasure in order to let people enjoy life. To illustrate, imagine going to a party, having a good time and drinking excessively. This might not be conducive to health and purposefulness, but it can nonetheless give people the feeling that they have an exciting and happy life. It is therefore important to clarify that death anxiety and immortality projects do not entirely eliminate all decisions based on short-term pleasure. Indeed, I have argued that many people already have moderate death anxiety, leading them to work on immortality projects. Nevertheless, these people can still make sporadic, short-term decisions. My argument is not that we should get rid of all short-term decisions, but that immortality projects are of crucial importance to counterbalance more life-affirming human drives. By counteracting imbalanced decisions and drives, we may optimize our conditions to maximize happiness. Conversely, proponents of active, voluntary euthanasia may propose that those people who do not want to opt for euthanasia can simply refuse it. But, in section 4, I demonstrated that legalizing euthanasia already has a sociological impact on HDMP as it provides people with the opportunity to take away suffering in dying. In contrast, this objection only addresses the refusal of practicing euthanasia. This is why my argument is primarily focussed on the legalization of euthanasia; not its practice.

Another objection could be that, even if it is true that legalizing euthanasia creates an imbalance between moderate death anxiety and life affirming human drives, it might also help restore that balance for people who have severe or chronic death anxiety. Indeed, not legalizing euthanasia does nothing to help restore impaired human activity in cases where this anxiety has become predominant. One way to respond to this objection is to argue

---

[31] Ibid., 120.

that excessive death anxiety occurs only to relatively few people.[32] As the majority of people have only moderate death anxiety, those few cases in which death anxiety severely impairs human behaviour might become neglectable in light of the happiness of all existing beings. This, of course, takes nothing away from the fact that there are downsides to not legalizing euthanasia from a utilitarianist perspective, including the fact that it would take away the possibility to alleviate pain in the process of dying. As mentioned before, the purpose of this paper is not to conclusively settle this issue; it merely serves to demonstrate one utilitarian drawback of diminishing death anxiety.

Finally, one may object that Mill's harm principle implies that it is conducive to happiness to let individuals take important decisions, like opting for euthanasia or not, themselves. As Mill says: "Over himself, over his own body and mind, the individual is sovereign."[33] However, similar to the argument of social limits against autonomy, my argument describes negative, social consequences of the legalization of euthanasia. As immortality projects provide cultural enrichment, socially cohesive communities and communitarian purpose, demotivating these projects can lead to harm for others as well. As Mill explicitly allows for restrictions if decisions threaten to make others worse off with respect to their fundamental interests, his harm principle is compatible with restrictions on euthanasia. Notwithstanding this point, it is notable that Mill's harm principle does deny paternalism precisely because making decisions ourselves helps us learn how to make better decisions.[34] Differing from Mill's view, this paper aimed to demonstrate that the opportunity to choose for a painless death actually threatens making good decisions. This appears to be incompatible with Mill's view on HDMP. One possible reconciliation might be that, in the specific case of euthanasia, the decision to opt for a painless death is not actually conducive to making better future decisions in Mill's view, as the death implied in euthanasia ipso facto takes away the opportunity for people to make more decisions in the future. However, concerning the limited scope of this essay, I will not elaborate further on Mill's harm principle. A further comparison of Mill's harm principle and my argument in relation to the impact of legalizing euthanasia on HDMP and happiness may give us a deeper understanding of this debate.

VII. Conclusion

We can see through our discussion on active, voluntary euthanasia that (moderate) death anxiety gives rise to immortality projects, decisions that are

---

[32] Furer, and Walker, 167.

[33] Mill, 13.

[34] Ibid., 57.

conducive to happiness as they are healthy and progress-driven. As euthanasia's capacity to take away suffering in dying diminishes death anxiety, it threatens to demotivate decisions that are conducive to happiness. This is, of course, focussing on a prior existence view of utilitarianism, as opposed to total-view utilitarianism. Although there are other plausible, utilitarian arguments that support euthanasia, it has been clearly shown that its impact on HDMP must also be taken into account. A further discussion might include a more comprehensive cost-benefit analysis of all the aforementioned arguments concerning the legalization of euthanasia.

References

Angner, Erik, Jennifer Ghandhi, Kristen Williams Purvis, Daniel Amante, and Jeroan Allison. "Daily Functioning, Health Status, and Happiness in Older Adults." *Journal of Happiness Studies* 14, no. 4 (2012): 1563-1574.

Becker, Ernst. *The Denial of Death.* New York: Free Press, 1973.

Bellaigue, Christopher. "Death on Demand: Has Euthanasia Gone Too Far?" *The Guardian*. January 18, 2019. https://www.theguardian.com/news/2019/jan/18/death-on-demand-has-euthanasia-gone-too-far-netherlands-assisted-dying.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation.* Ontario: Batoche Books, 2000.

Bioethics News. "Fear of Death Driving Push for Euthanasia, Says Medical Ethicist." *Bioethics Research Library*. https://bioethics.georgetown.edu/2015/05/fear-of-death-driving-push-for-euthanasia-says-medical-ethicist/.

Boer, Theo A. "Report on Euthanasia and Physician Assisted Suicide in the Netherlands." https://www.pthu.nl/Over-PThU/Organisatie/Medewerkers/t.a.boer/downloads/2016-boer-south-africa-affidavit-euthanasia-netherlands.pdf.

Burke, Brian, Spee Kosloff, and Mark Landau. "Death Goes to the Polls: A Meta-Analysis of Mortality Salience Effects on Political Attitudes." *Political Psychology* 34, no. 2 (2013): 183-200.

Davis, Alison. "A Disabled Person's Perspective on Euthanasia." *Disability Studies Quarterly* 24, no. 3 (2004), https://dsq-sds.org/article/view/512/689.

Dawkins, Richard. *The Greatest Show on Earth: The Evidence for Evolution.* New York: Simon & Schuster, 2009.

Devins, Gerald. "Death Anxiety and Voluntary Passive Euthanasia: Influences of Proximity to Death and Experiences with Death in Important Other Persons." *Journal of Consulting and Clinical Psychology* 47, no. 2 (1979): 301-309.

Furer, Patricia, and John Walker. "Death Anxiety: A Cognitive-Behavioral Approach." *Journal of Cognitive Psychotherapy* 22, no. 2 (2008): 167-182.

Grey, William. "Right to Die or Duty to Live? The Problem of Euthanasia." *Journal of Applied Philosophy* 16, no. 1 (1999): 19-32.

Harris, N. M. "The Euthanasia Debate." *Journal of the Royal Army Medical Corps* 147, no. 3 (2001): 367-370.

Iverach, Lisa, Ross Menzies, and Rachel Menzies. "Death Anxiety and Its Role in Psychopathology: Reviewing the Status of a Transdiagnostic Construct." *Clinical Psychology Review* 34, no. 7 (2014): 580-593.

Ly, Timothy, Mirella Diaz-Santos, Liam Campbell, Marcela Caldera, Taylor Kuhn, and Susan Bookheimer. "Death Anxiety and Financial Decision-Making in Aging: A Study from the Human Connectome Project Aging (HCP-A)." *Innovation in Aging* 3, no. 1 (2019): 907.

Mill, John Stuart. *On Libery*. Boston: Ticknor and Fields, 1859.

Noyes, Russel, Scott Stuart, Susan Longley, Douglas Langbehn, and Rachel Happel. "Hypochondriasis and Fear of Death." *The Journal of Nervous and Mental Disease* 190, no. 8 (2002): 503-509.

Salt, Henry. "The Philosopher and the Pig." *The Vegetarian* 9, no. 49 (1896): 585.

Sherwood, Harriet. "A Woman's Final Facebook Message Before Euthanasia: 'I'm Ready for my Trip Now...'" *The Guardian*. March 17, 2018, https://www.theguardian.com/society/2018/mar/17/assisted-dying-euthanasia-netherlands.

Singer, Peter. "Voluntary Euthanasia: A Utilitarian Perspective." *Bioethics* 17, no. 5-6 (2003): 526-541.

Singer, Peter. *Practical Ethics*. Cambridge: Cambridge University Press, 1993.

Sinoff, Gary. "Thanatophobia (Death Anxiety) in the Elderly: The Problem of the Child's Inability To Assess Their Own Parent's Death Anxiety State." *Frontiers in Medicine* 4 (2017): 1-5.

Solomon, Sheldon, Jeff Greenberg, and Tom Pyszcynski. "A Terror Management Theory of Social Behavior: The Psychological Functions of Self-Esteem and Cultural Worldviews." *Advances in Experimental Social Psychology* 24, no. 93 (1991): 93-159.

Solomon, Sheldon, Jeff Greenberg, and Tom Pyszcynski. *The Worm at the Core: On the Role of Death in Life*. New York: Penguin Random House, 2015.

Tedeschi, Richard, and Lawrence Calhoun. "Posttraumatic Growth: Conceptual Foundations and Empirical Evidence." *Psychological Inquiry* 15, no. 1 (2004): 1-18.

Vail, Kenneth, Jamie Arndt, Matt Motyl, and Tom Pyszcynski. "The Aftermath of Destruction: Images of Destroyed Buildings Increase Support for War, Dogmatism, and Death Thought Accessibility." *Journal of Experimental Social Psychology* 48, no. 5 (2012): 1069-1081.

interview
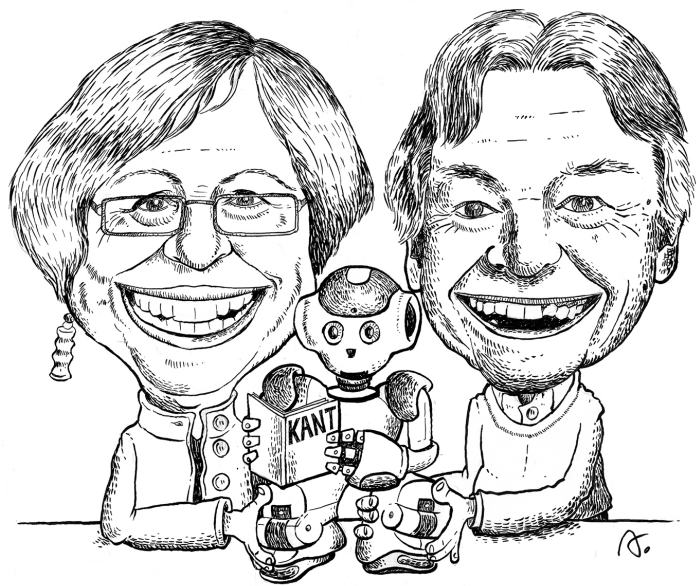
Towards Moral Machines:
A Discussion with Michael Anderson and Susan Leigh Anderson

Michael Anderson,[1] Susan Leigh Anderson,[2] Alkis Gounaris,[3] and George Kosteletos[4]

[1]*University of Hartford, USA*
E-mail address: anderson@hartford.edu
ORCID iD: http://orcid.org/0000-0001-7699-6156

[2]*University of Connecticut, USA*
E-mail address: susan.anderson@uconn.edu

[3]*National and Kapodistrian University of Athens, Greece*
E-mail address: alkisg@philosophy.uoa.gr
ORCID iD: https://orcid.org/0000-0002-0494-6413

[4]*National and Kapodistrian University of Athens, Greece*
E-mail address: gkosteletos@philosophy.uoa.gr
ORCID iD: https://orcid.org/0000-0001-6797-8415

Abstract

At the turn of the 21st century, Susan Leigh Anderson and Michael Anderson conceived and introduced the Machine Ethics research program, that aimed to highlight the requirements under which autonomous artificial intelligence (AI) systems could demonstrate ethical behavior guided by moral values, and at the same time to show that these values, as well as ethics in general, can be representable and computable. Today, the interaction between humans and AI entities is already part of our everyday lives; in the near future it is expected to play a key role in scientific research, medical practice, public administration, education and other fields of civic life. In view of this, the debate over the ethical behavior of machines is more crucial than ever and the search for answers, directions and regulations is imperative at an academic, institutional as well as at a technical level. Our discussion with the two inspirers and originators of Machine Ethics highlights the epistemological, metaphysical and ethical questions arising by this project, as well as the realistic and pragmatic demands that dominate artificial intelligence and robotics research programs. Most of all, however, it sheds light upon the contribution of Susan and Michael Anderson regarding the introduction and undertaking of a main objective related to the creation of ethical autonomous agents, that will not be based on the "imperfect" patterns of human behavior, or on preloaded hierarchical laws and human-centric values.

Key-words*: Machine Ethics; AI Ethics; Philosophy of Artificial Intelligence; Artificial Moral Agents; Ethical Machines; Moral Status of Robots; Computation of Bio-Medical Ethics*

Alkis Gounaris & George Kosteletos: Susan, Michael, thank you very much for the opportunity to discuss such an interesting issue with you. It is our great pleasure and honor to be able to share with our readers and the academic community in Greece and internationally this exceptional conversation. The rapid technological developments of recent years and what the immediate future holds for us bring your work to the forefront of every discussion about AI and Machine Ethics. Building an ethical machine, a possibility that perhaps a few years ago looked like a sci-fi scenario, today seems like an imperative and urgent demand. This seems to be the main objective of your work.

Susan Leigh & Michael Anderson: Thank you for giving us the opportunity to discuss our work in the context of current issues of artificial intelligence!

Alkis Gounaris & George Kosteletos: You introduced the Machine Ethics research program about seventeen years ago.[1] What is the purpose of Machine Ethics and what distinguishes Machine Ethics from the rest of the AI Ethics field? Why is Machine Ethics still important? We are now at the beginning of 2021. Seventeen years later, what is your assessment regarding the evolution of this program?

Susan Leigh Anderson: The main purpose of the Machine Ethics program is to ensure that autonomous AI systems behave in an ethical fashion when interacting with human beings. Secondarily, I believe that it gives us a chance to become clearer about ethics – how to represent its building blocks, resolve contradictions, and come up with principles that should guide the actions of systems functioning in particular domains – that, hopefully, will inspire us to behave better.

Michael Anderson: When we first conceived the idea of Machine Ethics at the turn of the century, the prevailing thinking was that such a notion was still firmly in the realm of science fiction and would remain there for the foreseeable future. This attitude stemmed from a myopic view of the types

---

[1] Michael Anderson, Suzan Leigh Anderson, and Chris Armen, "Towards Machine Ethics," in *Proceedings of the AAAI-04 Workshop on Agent Organizations: Theory and Practice*, 53-59 (San Jose, CA, 2004); Michael Anderson, Suzan Leigh Anderson, and Chris Armen, "Toward Machine Ethics: Implementing Two Action-based Ethical Theories," in *Machine Ethics, Papers form AAAI Fall Symposium, 2005*, eds. Michael Anderson, Suzan Leigh Anderson, and Chris Armen, Technical Report FS-05-06 (Menlo Park, CA: Association for the Advancement of Artificial Intelligence, 2005), https://www.aaai.org/Library/Symposia/Fall/fs05-06.php; Michael Anderson, Suzan Leigh Anderson, and Chris Armen, "An Approach to Computing Ethics," *IEEE Intelligent Systems* 21, no. 4 (2006): 65-63; Michael Anderson, and Suzan Leigh Anderson, "Machine Ethics: Creating an Ethical Intelligent Agent," *AI Magazine* 28, no. 4 (2007): 15-26; Michael Anderson, and Suzan Leigh Anderson, "The Status of Machine Ethics: A Report from the AAAI Symposium," *Minds & Machines* 17 (2007): 1-10. See also Michael Anderson, and Susan Leigh Anderson, eds., *Machine Ethics* (New York and Cambridge: Cambridge University Press, 2011).

of behavior that would entail ethical concerns and the speed with which autonomous systems capable of such behavior would be upon us.

Given this, the original purpose of the project was to give evidence that

1. autonomous systems need not be fully realized to exhibit behavior of ethical concern
2. ethics is representable and computable
3. the behavior of autonomous systems can be guided by ethical principles

As all AI is machine-based, we see little difference between AI Ethics and Machine Ethics other than its focus on issues raised by the systems recently developed by deep learning. As such systems arise in a black-box fashion from non-vetted data, it is difficult to see how these issues will be resolved and, ultimately, how we will ever be able to guarantee ethical behavior from these systems. Unless such a guarantee can be given, it does not seem likely that such systems will be acceptable. That said, given the surprising proliferation of autonomous systems in general, we believe the tenets of the Machine Ethics project are more relevant than ever.

Alkis Gounaris & George Kosteletos: Having previously argued for the expediency of the Machine Ethics research program, you have pointed out that one of the advantages that machines have over humans in the process of moral judgment is the feature of impartiality (non-bias).[2] Due to their mechanical nature, AI agents are impartial, namely they judge without any bias, unlike humans who tend to be partial, since for example they often decide while being emotionally charged. However, if at some point, in the future, the initial goal of AI is achieved and machines acquire humanlike cognition, do you think they will preserve the advantage of impartiality over humans? Such a question outlines a possible conflict between the basic research objective of AI – specifically the creation of truly intelligent machines – and the goal of Machine Ethics research program regarding the creation of impartial ethical advisors and impartial explicit ethical agents. This possible conflict of the basic research goals of AI and Machine Ethics can also be seen in relation to the vision of creating super-intelligent machines. We say this thinking of Daniel Dennett, who refers to Nietzsche, saying that delusion and deception are characteristics of human nature thus only such a nature can understand

---

[2] Michael Anderson, Susan Leigh Anderson, and Chris Armen, "An Approach to Computing Ethics," *IEEE Intelligent Systems* 21, no. 4 (2006): 65-63; Michael Anderson, and Suzan Leigh Anderson, "Machine Ethics: Creating an Ethical Intelligent Agent," *AI Magazine* 28, no. 4 (2007): 15-26; Michael Anderson, and Suzan Leigh Anderson, "Robot Be Good," *Scientific American* 303, no. 4 (2010): 72-77.

ethics.[3] Driven by Dennett's position, we think that if the machines reach in the future a kind of super-intelligence that will be impartial at the same time, they may not be interested in ethics at all or will not justify its usefulness.

Susan Leigh Anderson: I have long been concerned with the bias of AI researchers towards trying to reproduce human cognition and human intelligence, and even our ethical values. We are not ideal beings! We can do better than model human behavior as we create autonomous AI entities.

Michael Anderson: Given the initial reticence to see Machine Ethics in any light other than one of science fiction, we purposefully limited the scope of our research to immediate, pragmatic concerns with the hope of convincing some of the scientific fact of its need. It remains to be seen whether "super-intelligence" will make the same leap from fiction to fact. That said, if it does in fact make such a leap, you can be sure if we have given little thought to how we would like such machines to behave towards us, it is likely that we will have little say in how they actually do.

Alkis Gounaris & George Kosteletos: Would you say that the idea for a Machine Ethics, finally the idea that ethics is computable, could be thought of as part of the philosophical tradition supporting that thought equals calculation? Would you consider yourselves as belonging to the same line of thinkers like Hobbes,[4] Leibniz,[5] and more recently Turing,[6] McCulloch and Pitts,[7] or Newell and Simon?[8]

---

[3] Daniel Dennett, "When Hal Kills, Who's to Blame? Computer Ethics," in *Hal's Legacy: 2001's Computer as Dream and Reality*, ed. David G. Stork, 351-365 (Cambridge, MA: MIT Press, 1997).

[4] Thomas Hobbes, *Leviathan, or The Matter, Forme and Power of a Commonwealth Ecclesiastical and Civil*, ed. A. R. Waller (Cambridge: Cambridge University Press, 1904).

[5] Gottfried Wilhelm Leibniz, *Dissertatio de arte combinatoria* (Paris: Hachette Livre-BNF, 2018); Gottfried Wilhelm Leibniz, "Principles of Nature and Grace, Based on Reason," in *Gottfried Wilhelm Leibniz, Philosophical Papers and Letters*, ed. Leroy E. Loemker (Dordrecht: Springer, 1989).

[6] Alan Mathison Turing, "Intelligent Machinery," in *Machine Intelligence 5*, ed. B. Meltzer, and D. M. Michie, 3-23 (Edinburgh: Edinburgh University Press, 1969); Alan Mathison Turing, "Computing, Machinery and Intelligence," *Mind* 59 (1950): 433-460. See also Alan Mathison Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," in *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life*, ed. Jack B. Copeland, 58-90 (Oxford: Oxford University Press, 2004) – see especially p. 59.

[7] Warren S. McCulloch, and Walter H. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics* 5 (1943): 115-33.

[8] Allen Newell, and Herbert Alexander Simon, *Current Developments in Complex Information Processing: Technical Report P-850* (Santa Monica, CA: Rand Corporation, 1956); Allen Newell, and John Crosley Shaw, "Programming the Logic Theory Machine," in *IRE-AIEE-ACM '57 (Western): Papers Presented at the February 26-28, 1957, Western Joint Computer Conference: Techniques for Reliability*, 230-240 (New York: Association for Computing Machinery, 1957); Allen Newell, and Herbert Alexander Simon, "The Logic Theory Machine: A Complex Information-Processing System," *IRE Transactions on Information Theory* 2, no. 3

Susan Leigh Anderson: While I believe that ethics is, in principle, computable (and we have been trying to demonstrate this), I'm not sure that I would go so far as to say that all thought is computable. What about artistic ideas?

Michael Anderson: It seems a bit of a stretch from "having machines behave ethically towards us" – the stated goal of our Machine Ethics project – and "all thought is calculation," don't you think?

Alkis Gounaris & George Kosteletos: During the process of ethical decision making one is likely to find oneself facing a condition known in Ethical Philosophy as 'conflict of duties.' Is it possible that in trying to tackle a conflict of moral duties in a computational basis, one might find oneself facing a kind of a 'Halting Problem?'[9] Could it be possible that the explicit ethical agent would be trapped in a never-ending calculation, maybe an infinite loop going back and forth between two opposing duties? In your opinion, are there any major difficulties in the fulfillment of the Machine Ethics endeavor – for instance difficulties related to the ontology, the very nature of calculation or of ethics?

Michael Anderson: Clearly time is of the essence in such decision making and, if competing duties are so closely tied, simply choosing either when time is up would seem a sufficient means to end deliberation. Minsky, in a private conversation, once said to Susan (in his inimitable way) "Ethics is what you do when you run out of time." Just as clearly, hundreds of years of reflection on ethical matters has laid bare a myriad of difficulties that are likely to plague efforts in Machine Ethics as well. That said, perhaps the constrained domain and new perspective of the effort might shed new light on some of these difficulties.

(1956): 61-79; Allen Newell, and Herbert Alexander Simon, "GPS-A Program that Simulates Human Thought," in *Lernende Automaten*, ed. Heinz Billing, 109-124 (Münich: Oldenburg, 1961); Allen Newell, John Crosley Shaw, and Herbert Alexander Simon, "Element of a Theory of Human Problem Solving," *Psychological Review* 65 (1958): 151-166; Allen Newell, and Herbert Alexander Simon, "Computing Science as Empirical Enquiry: Symbols and Search," *Communications of the Association for Computing Machinery* 19 (1976): 113-126; Allen Newell, "Physical Symbol Systems," *Cognitive Science* 4 (1980): 135-183.

[9] Alan Mathison Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society Series* 2, no. 42 (1937): 230-265, reprinted in *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life*, ed. Jack Copeland, 58-90 (Oxford: Oxford University Press, 2004); Alan Mathison Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction," *Proceedings of the London Mathematical Society* 43 (1938): 544-546; Martin Davis, *Computability and Unsolvability* (New York: McGraw-Hill, 1958), 70. See also Stephen Cole Kleene, *Introduction to Metamathematics* (Amsterdam: North-Holland, 1952), especially Chapter 13: "Computable Functions," and Marvin Minsky, *Computation: Finite and Infinite Machines* (New Jersey: Prentice-Hall, 1967), specifically chapter 8, Section 8.2: "Unsolvability of the Halting Problem."

Alkis Gounaris & George Kosteletos: Persisting a little longer on the issue of 'conflict of duties,' we would like you to comment on a related possibility. We are referring specifically to the case where the machine would have to choose between self-preservation (e.g. the search for vital resources) and continuing to fulfill the principles of a human-centered ethic (e.g. the principles of serving human well-being). Could this conflict of duties be averted by programming rules such as Asimov's *Three Laws of Robotics*?[10] Susan has been critical of them in the past, commenting that they could not be a satisfactory basis for Machine Ethics.[11] Could you tell us a few words about this claim while also suggesting an alternative for facing the above mentioned conflict of duties?

Susan Leigh Anderson: There are a number of problems with Asimov's *Laws* as a basis for Machine Ethics. Roger Clarke[12] has pointed out that there are a number of inconsistencies and ambiguities in the laws. Also, it could allow humans to abuse entities that resemble humans in form, leading to finding it easy to abuse humans as well. Most significantly, from our perspective, a hierarchical ethical duty theory is unsatisfactory because, in agreement with W.D. Ross, we believe that all ethical duties should be viewed as *prima facie*. That is, although all relevant ethical duties should be considered, none should be viewed as being absolute, as the top duty in a hierarchical ordering of duties would be. Each one could be overridden, on occasion, by another duty/duties that would be stronger in a particular situation.

Michael Anderson: Asimov's *Laws* were a landmark in ethical thinking concerning the actions of robots. This is true even when one considers they were devised simply as a device for generating fiction – Asimov seemed to spend more time delineating their weaknesses than championing their strengths. From a real-world perspective, one might question their insufficient specification, incomplete coverage of ethical duties, rigid hierarchal disposition, and required slave-like obedience.

Clearly, the robot has a duty to maintain itself in addition to its other ethical obligations towards its human user. And there is no simple answer as to whether it takes precedence when it conflicts with the other duties as this is a context dependent question. Sometimes it should, say when the robot's other

---

[10] Isaac Asimov, "The Bicentennial Man," in *Philosophy and Science Fiction*, ed. Michael Phillips, 183-216 (Buffalo, New York: Prometheus Books, 1984).

[11] Suzan Leigh Anderson, "Asimov's 'Three Laws of Robotics' and Machine Metaethics," *AI and Society* 22 (2007): 477-493; Suzan Leigh Anderson, "The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics," in *Machine Ethics*, ed. Michael Anderson, and Suzan Leigh Anderson, 285-296 (New York and Cambridge: Cambridge University Press, 2011).

[12] Roger Clarke, "Asimov's Laws of Robotics: Implications for Information Technology. Part I," *Computer* 26, no. 12 (1993): 53-61; Roger Clarke, "Asimov's Laws of Robotics: Implications for Information Technology. Part II," *Computer* 27, no. 1 (1994): 57-66.

duties are not as pressing, and sometimes it shouldn't, say when great harm might befall its human user if the robot tends to its needs rather than hers. Our work in machine ethics has shown how we might tease out the relationships between duties and how to use this information to drive a robot's behavior: abstract principles of conflict resolution from agreed upon cases and use these principles to order actions in terms of their ethical preference.

Alkis Gounaris & George Kosteletos: Bostrom, Yudkowsky and others talk about the so-called Value Loading Problem,[13] namely the problem of how to make machines understand and adopt the values and goals of the humans. However, in our view, even before we address this issue, there may exist another question that we have to answer. Specifically, if one approaches the concept of autonomy in Kantian terms,[14] then arises the question of whether we ought (here, in terms of an ethical "ought") to be concerned with the Value Loading Problem at all. More specifically, dealing with the Value Loading Problem implies the imposition of certain values on the machines (i.e human-centered values, generally values of our own choice etc.). However, this would be against the ethical principle of respecting the autonomy of others. Thus, as human AI developers, we may be faced with the following moral dilemma: Solving the Value Loading Problem to satisfy human goals and ensuring the survival of the human species, or staying consistent with our ethical principle of respect for the autonomy of others?[15] Do you think this dilemma is valid or is it a pseudo-problem? If it is valid, do you see any way out of it?

Susan Leigh Anderson: As I mentioned previously, I don't think we should build all human values into autonomous machines, since humans are prone to unethical behavior. We can, and should, do better than that. Nevertheless, until these entities demonstrate that they have the qualities necessary to

---

[13] Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014); Eliezer Yudkowsky, "Complex Value Systems in Friendly AI," in *Artificial General Intelligence*, edited by Jürgen Schmidhuber, Kristinn R. Thórisson, and Moshe Looks, 388-393 (Berlin, Heidelberg: Springer, 2011). See also Eliezer Yudkowsky, "The Value Loading Problem," *EDGE*, July 12, 2021, https://www.edge.org/response-detail/26198; Nate Soares, "The Value Learning Problem," in *Artificial Intelligence, Safety and Security*, ed. Roman V. Yampolskiy, 89-97 (Boca Raton, FL: CRC Press, 2019).

[14] Immanuel Kant, *The Groundwork for the Metaphysics of Morals*, trans. Allen W. Wood (New Haven and London: Yale University Press, 2002), for instance see 4: 435-6, 4:440 and 4:447; Immanuel Kant, *Critique of Practical Reason*, trans. Mary Gregor (Cambridge: Cambridge University Press, 2015), see 5:132, also 5:29.

[15] Here, the Value Loading Problem concerns one of the two conflicting duties of the human-developer. It lies at one end of the dilemma, as it has to do with the satisfaction of human goals. The other end is what concerns the respect of the autonomy of others, in this case the AI agents.

be considered to be *full ethical agents* (that we, following James Moor,[16] distinguish from being *explicit ethical agents*, which is what we attempt to create), we don't have to worry about respecting their autonomy. It is perfectly appropriate that, since they are designed to be in the service of human beings (and, perhaps, animals as well), they should be designed to respect their rights.

Alkis Gounaris & George Kosteletos: There are many who argue that creating a literally ethical machine is practically impossible and ultimately unachievable[17] and that we should come to terms with the assumption that at least at an early stage, the basic ethical values will eventually be loaded. Drawing on the theory of W. D. Ross,[18] as well as the *Principles of Biomedical Ethics*[19] by Beauchamp and Childress, you propose that an ethical machine should possess prima facie duties.[20] Do you think that there could be a specific ethical theory that would effectively cover all the possible ethically-laden circumstances (all the cases in need of an ethical analysis) that an AI agent will have to deal with? The danger here is that the agent may operate on the basis of certain principles that will prove to be effective in some cases and ineffective – even dangerous – in others. Furthermore, would a finite set of principles be sufficient for the AI agent to recognize the ethically relevant and

---

[16] James H. Moor, "The Nature, Importance, and Difficulty of Machine Ethics," *IEEE Intelligent Systems* 21, no. 4 (2006): 18-21.

[17] Roman Yampolskiy, "Artificial Intelligence Safety Engineering: Why Machine Ethics is a Wrong Approach," in *Philosophy and Theory of Artificial Intelligence. Studies in Applied Philosophy, Epistemology and Rational Ethics*, ed. Vincent Müller, 389-396 (Berlin, Heidelberg: Springer, 2013).

[18] W. D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930).

[19] T. L. Beauchamp, and J. F. Childress, *Principles of Biomedical Ethics* (Oxford, UK: Oxford University Press, 1979).

[20] For instance see Anderson, Anderson, and Armen, C., "An Approach;" Michael Anderson, and Susan Leigh Anderson, "MedEthEx: A Prototype Medical Ethics Advisor," *Proceedings of the 21st National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*, 1759-1765 (Boston, MA: AAAI Press, 2006); Anderson, and Anderson, "Machine Ethics: Creating;" Anderson, "Asimov's Three Laws;" Anderson, and Anderson, "Robot Be Good;" Anderson, "Machine Metaethics;" Michael Anderson, and Suzan Leigh Anderson, "A Prima Facie Duty Approach to Machine Ethics: Machine Learning of Features of Ethical Dilemmas, Prima Facie Duties, and Decision Principles through a Dialogue with Ethicists," in *Machine Ethics*, ed. Michael Anderson, and Suzan Leigh Anderson, 476-492 (New York and Cambridge: Cambridge University Press, 2011); Suzan Leigh Anderson, "Philosophical Concerns with Machine Ethics," in *Machine Ethics*, ed. Michael Anderson, and Suzan Leigh Anderson, 162-167 (New York and Cambridge: Cambridge University Press, 2011); Suzan Leigh Anderson, and Michael Anderson, "Towards a Principle-based Healthcare Agent," in *Machine Medical Ethics*, ed. S. van Rysewyk, and M. Pontier, 67-77 (Cham: Springer, 2015); Michael Anderson, and Suzan Leigh Anderson, "Toward Ensuring Ethical Behavior from Autonomous Systems: A Case-supported Principle-based Paradigm," *Industrial Robot* 42, no. 4 (2015): 324-331.

prominent features of every possible circumstance? In other words, would this *finite* set of ethical principles be sufficient for the AI agent to recognize *every* ethically-laden case as such? There is a risk here that there will be cases that the agent will fail to recognize as ethically-laden (i.e. circumstances asking for an ethical analysis). In addition to the ethical principles themselves, this problem could also arise regarding the criteria for applying these principles. Again, the finite nature of these criteria could make the AI agent fail in the recognition of a situation as ethically-laden (i.e. failure to recognize a situation in which the agent should apply its ethical principles). One might, probably, argue that this is a version of the Frame Problem of AI[21] applied in the case of ethical functioning of the AI agents; or, as we could say, a Moral Frame Problem of AI. With this in mind, the above question can be phrased as such: Is it possible for a specific ethical theory, therefore a *finite* set of ethical principles, to successfully address the Moral Frame Problem of AI?

Susan Leigh Anderson: Two points need to be mentioned here: The first is that, for the foreseeable future, autonomous AI entities are likely to be developed to function in particular domains, with a limited number of ethically relevant features, and corresponding prima facie duties to be considered, leading to a decision principle that can be learned from select ethical dilemmas that are likely to be encountered in those domains. Second, we don't believe that there are situations where *no* ethically relevant features, and corresponding duties, are present when the autonomous AI entity interacts with humans. Those who reject this position tend to think of ethical dilemmas as involving significant harm to a human, but the ethical perspective involves determining the *best* action that could be performed in particular situations. There are always better and worse actions to be considered. So the AI entity, on our view, never has to determine whether a particular situation is an ethically significant one or not. All of its actions should be subsumed under the learned ethical principle, no matter how trivial.

Michael Anderson: It seems that the problem described applies to *all* autonomously-acting agents, including human beings. Until we develop

---

"philosopher robots," and, in the vein of human beings, a race of philosophers, it seems that autonomous agents are doomed by their finite capabilities to make mistakes and, hopefully, learn from them. That said, it seems likely that the set of ethically relevant features, and hence the corresponding duties to minimize or maximize them, is not infinite. In fact, Utilitarians might argue that net good is the *only* ethically relevant feature. While that may or not be the case, we argue that a finer gradation (and hence greater number) of ethically relevant features may be needed to help illuminate the reasoning behind ethical decision making.

Alkis Gounaris & George Kosteletos: If at least for the time being we cannot avoid the (even partial) 'loading' of some basic or initial moral values to the AI agents, then shouldn't this process of regulating 'value loading' involve the end-users and not only the AI developers? In other words, shouldn't the ordinary citizens have a say in the choice of those principles? Additionally, shouldn't each cultural background regarding morality be taken into account? We saw in a very interesting MIT experiment the different ways in which different cultures react to the 'trolley problem' that came to the fore with the evolution of smart cars.[22] The question is whether the design of an ethical machine should follow the demand for the democratization of technology and technical design[23][24] – or even a culture based technical design.[25] Recently, you have also proposed a framework promoting public participation as part

---

[22] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan, "The Moral Machine Experiment," *Nature* 563, no. 7729 (2018): 59-64; Jean-Francois Bonnefon, Azim Shariff, and Iyad Rahwan, "The Social Dilemma of Autonomous Vehicles," *Science* 352, no. 6293 (2016): 1573-1576; Edmond Awad, Sohan Dsouza, Azim Shariff, Iyad Rahwan, and Jean-Francois Bonnefon, "Universals and Variations in Moral Decisions Made in 42 Countries by 70,000 Participants," *Proceedings of the National Academy of Sciences* 117, no. 5 (2020): 2332-2337.

[23] Andrew Feenberg, "Subversive Rationalization: Technology, Power, and Democracy," in *Technology and the Politics of Knowledge*, ed. Andrew Feenberg, and Alastair Hannay, 3-11 (Bloomington and Indianapolis: Indiana University Press, 1995); Andrew Feenberg, *Questioning Technology* (London, New York: Routledge, 1999); Carl Mitcham, *Thinking through Technology: The Path between Engineering and Philosophy* (Chicago: The University of Chicago Press, 1994); Langdon Winner, "Technè and Politeia: The Technical Constitution of Society," in *Philosophy of Technology*, ed. Paul T. Dubrin, and Friedrich Rapp, 97-111 (Dordrecht, Boston, Lancaster: D. Reidel, 1983); Langdon Winner, "Citizen Virtues in a Technological Order," *Inquiry* 35, nos. 3-4 (1992): 341-361.

[24] The question regarding the democratization of Technology is closely related to the notions of *inclusion*, *fairness* and *transparency*, which seem to have become popular topics in the AI research literature. See *The 2019 AI Index Annual Report, Stanford University Human Centered AI*, Chapter 8: "Societal Considerations," especially pages 149-151.

[25] Karen Hao, "Should a Self-driving Car Kill the Baby or the Grandma? Depends on where You're from," *MIT Technology Review*, October 14, 2018, https://www.technologyreview.com/2018/10/24/139313/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/.

of a process – or as you call it, a "tool" – for the formulation of principles to be loaded to the machines.[26] Do you generally agree with an inclusive approach with regards to the Machine Ethics research program?

Susan Leigh Anderson: I have argued that, in general, applied ethicists (with knowledge of the domains in question) should be involved in learning the ethical principles, from the ethically relevant features and correlative prima facie duties that should govern the behavior of autonomous AI entities in specific domains. They have an expertise that others lack. But I have also accepted (after discussions with Edmond Awad) that there is an ethically justifiable place for the opinions of the general public concerning emerging technologies, for instance, driverless cars: Since there has been push-back from the public about allowing driverless cars in large part because of a death in Arizona by a driverless car and concern that there are bound to be situations, even with improved sensors, where the behavior of driverless cars could result in deaths, there needs to be a way for the public to weigh in on this possibility to allow for the acceptance of driverless cars, which would certainly result in fewer deaths than with human drivers who are often distracted, tired or impaired.

Until recently Michael and I have maintained that we didn't think that machines should be permitted to function autonomously in domains where life-and-death decisions need to be made, because they are controversial. Such decisions are controversial because they are often emotionally driven for ordinary people and even ethicists disagree about how to weigh the various ethically relevant factors involved. The case of driverless cars is very different, I now see. A central ethical concern for any action or policy must be causing the least harm. This is universally agreed upon. It seems clear that having only driverless cars would result in less harm than having only human drivers. If there were some way to placate the public's concerns about when driverless cars behavior might lead to human deaths, leading to allowing them, it should be taken seriously. Encouraging the public to have a say in what driverless cars should do in various possible scenarios where death might result, making the results known and adopting the majority's view (probably for a particular society), might just be enough for the public to accept driverless cars, which is likely to lead to fewer deaths overall.

And, actually, it is consistent with our long held position that only humans should make life-and-death decisions since, although the cars function autonomously, the decisions they make were determined by humans who

---

[26] Edmond Awad, Michael Anderson, Suzan Leigh Anderson, and Beishui Liao, "An Approach for Combining Ethical Principles with Public Opinion to Guide Public Policy," *Artificial Intelligence* 287 (2020): article 103349.

gave them rules to follow. Humans will be held accountable if the results are questioned. I foresee challenges to the majority's recommended policies as time goes by, leading perhaps to new policies approved by the majority, just as laws in this country are changed over time, hopefully leading to more ethically acceptable ones as ethicists and others weigh in.

Alkis Gounaris & George Kosteletos: In several publications you refer to the creation of ethical advisors such as the bio-medical advisor MedEthEx.[27] As mentioned before, you suggest that the most appropriate way for these advisors to operate is based on the *Principles of Biomedical Ethics*[28] by Tom Beauchamp and James Childress. Can you tell us a bit more about your proposal?

Susan Leigh Anderson: We began testing our approach to representing ethics in a machine, and generating ethical decision principles from considering specific cases of ethical dilemmas, by using a general type of ethical dilemma often faced by medical practitioners, where the ethics is clear. Medical Ethics is quite well established and there is agreement on using Beauchamp and Childress's principles (*prima facie* duties, in our view, since there is no decision principle to resolve cases where they give conflicting advice) to frame discussions.

Here is the common type of ethical dilemma we considered: A health care worker has recommended a particular treatment for her competent adult patient and the patient has rejected that treatment option. Should the health care worker try again to change the patient's mind or accept the patient's decision as final? The dilemma arises because, on the one hand, the health care worker may not want to risk upsetting the patient by challenging his decision; on the other hand, the health care worker may have concerns about why the patient is refusing the treatment. Three of the four principles/duties of Biomedical Ethics are likely to be satisfied or violated in dilemmas of this type: the duty of respect for autonomy, the duty of nonmaleficence and the duty of beneficence. The system accepts a range of values for each of the duties from −2 to +2, where -2 represents a serious violation of the duty, -1 a less serious violation, 0 indicates that the duty is neither satisfied nor violated, +1 indicates a minimal satisfaction of the duty and +2 a maximal satisfaction of the duty.

Through inductive logic, after considering several cases giving reasons why the patient was rejecting the recommended treatment where the answer is clear as to whether the patient's decision should be accepted or challenged,

---

[27] Anderson, and Anderson, "MedEthEx."

[28] Tom Lamar Beauchamp, and James Franklin Childress, *Principles of Biomedical Ethics* (Oxford: Oxford University Press, 1979).

the system learned this principle: A health care worker should challenge a patient's decision if it is not fully autonomous and *either* there is any violation of the duty of nonmaleficence *or* there is a severe violation of the duty of beneficence. This philosophically interesting result gives credence to Rawls' Method of Reflective Equilibrium.[29] We have, through abstracting a principle from intuitions about particular cases and then testing that principle on further cases, come up with a plausible principle that tells us which action is correct when specific duties pull in different directions in a particular ethical dilemma. Furthermore, the principle that has been abstracted supports an insight of Ross's that violations of the duty of nonmaleficence should carry more weight than violations of the duty of beneficence.

Alkis Gounaris & George Kosteletos: However, in addition to the question "What ethical principles should the AI ethical advisor, or the explicit agent, apply?" the question "To which entities should the AI agent apply these criteria?" arises as well. Some would say that this question seems to be gaining in importance considering the possibility of developing in the future machines with a significant degree of autonomy that will be able to interact with their environment in a more 'holistic' way. In such a case, we also need to face the question of "How will the AI agent decide a) which of its surrounding entities have moral standing and therefore need a moral treatment from the AI agent?, and b) what exactly this moral standing would involve?" Is the issue of defining criteria for the attribution of moral status to others crucial for the Machine Ethics research program? If so, are there any satisfactory criteria that an AI agent could effectively apply for the attribution of moral status to its surrounding entities?

Susan Leigh Anderson: In my view *sentience* is the quality an entity should possess to have moral standing, because only an entity possessing this quality would care what happens to it. But it is difficult to detect whether this quality is present in an entity other than oneself. And it isn't necessary to possess this quality for it to be important that we treat an entity as if it has moral standing. I have argued – using Kant's argument for why we should treat animals well, where he maintained that even though they don't have rights themselves (now debatable), because they resemble us we should treat them as if they have rights lest it lead to a slippery slope where it becomes easier to

---

[29] John Rawls, *A Theory of Justice* (Cambridge, MA: The Belknap Press of Harvard University Press, 1971). For subsequent refinements and reappraisals of the theoretical construct of the Reflective Equilibrium see John Rawls, *A Theory of Justice*, 2nd edition (Cambridge, MA: The Belknap Press of Harvard University Press, 1999), and John Rawls, "Justice as Fairness: Political not Metaphysical," *Philosophy and Public Affairs* 14 (1985): 223-251. For the distinction between narrow and wide Reflective Equilibrium see John Rawls, *Justice as Fairness: A Restatement* (Cambridge, MA: Harvard University Press, 2001), 31.

mistreat other humans – that any entity that resembles us in form or function should be treated as if it has moral standing.

Alkis Gounaris & George Kosteletos: Let us insist for a moment on the issue of the criteria for the attribution of moral status. From time to time in some of your papers you have considered the question of whether an explicit ethical agent should follow a set of ethical principles that will involve the fact that the agent itself has a moral standing. Namely, whether the agent should 'consider' (or consider) itself as an entity with moral standing.[30] Could you please tell us more about the significance and the importance of this question?

Susan Leigh Anderson: I don't think it's important to determine its own status in order to decide how *it* should treat others.

Michael Anderson: I can imagine that one might draw the wrong conclusion about our stance towards this question when one considers that we advocate that such an agent has a duty to maintain itself. I would argue that this does not in fact pertain to an attribution of moral status to the agent but instead is concerned with making sure that the agent maintains its capacity to fulfill its other duties towards its user.

Alkis Gounaris & George Kosteletos: In addition to the above question as to whether the explicit ethical agent 'considers' (or considers) itself as an entity with moral standing, many AI researchers also reasonably pose the question of whether *we* should consider the explicit ethical agent (and generally any AI agent) as an entity with moral standing.[31] Should we bother with the attribution of moral status to AI entities? If so, what do you think the criteria are that an explicit ethical agent (and, more generally, an AI agent) should meet in order for moral status to be attributed to it? For example, some people think that an ethical Turing Test will be enough to attribute moral status to machines.[32] Do you think accepting this view is the only way

---

[30] Anderson, "Asimov's Three Laws."

[31] For instance, Luciano Floridi, and J. W. Sanders, "On the Morality of Artificial Agents," *Minds and Machines* 14 (2004): 349-379; Christian Hugo Hoffmann, and Benjamin Hahn, "Decentered Ethics in the Machine Era and Guidance for AI Regulation," *AI & Society* 35, no. 3 (2009): 635-644; David Levy, "The Ethical Treatment of Artificially Conscious Robots," *International Journal of Social Robotics* 1, no. 3 (2009): 209-216; Bertram F. Malle, Thapa Stuti Magar, and Matthias Scheutz, "AI in the Sky: How People Morally Evaluate Human and Machine Decisions in a Lethal Strike Dilemma," in *Robotics and Well-Being*, ed. Maria Aldinhas Ferreira, João Silva Sequeira, Gurvinder Singh Virk, Mohammad Tokhi Osman, and Ender E. Kadar, 111-133 (Cham: Springer, 2019); Robert Sparrow, "Killer Robots," *Journal of Applied Philosophy* 24, no. 1 (2007): 62-77. See also Jonathan Owen, and Richard Osley, "Bill of Rights for Abused Robots: Experts Draw up an Ethical Charter to Prevent Humans Exploiting Machines," *The Independent*, September 17, 2011, https://www.independent.co.uk/news/science/bill-of-rights-for-abused-robots-5332596.html.

[32] Colin Allen, Varner Gary, and Zinser Jason, "Prolegomena to Any Future Artificial Moral

to go? You have also commented[33] on criteria like Jeremy Bentham's and Peter Singers' criterion of sentience[34] – which you have also mentioned earlier in our discussion – Immanuel Kant's criterion of self-consciousness,[35] Michael Tooley's criterion of desire (for a moral right),[36] and Mary Anne Warren's criterion of emotionality.[37] Do you find any flaws in these criteria?[38] For example, does the Other Minds Problem pose a threat to the feasibility of applying such criteria, namely criteria of an internalist kind?[39] Furthermore, what do you think the moral status of AI agents could finally be?

Susan Leigh Anderson: As I mentioned earlier, answering your question regarding the criteria that an AI agent could effectively apply for the attribution of moral status to its surrounding entities, given the Problem of Other Minds, we may never know whether an autonomous AI entity possesses the quality essential to having moral standing, but I have argued that we should treat it (if it resembles us, or an animal, in form or function) as if it does.

Alkis Gounaris & George Kosteletos: With your work you have opened a new path for the treatment and resolution of the ethically-laden biomedical

---

Agent," *Journal of Experimental and Theoretical Artificial Intelligence* 12 (2000): 151-261; Robert Sparrow, "The Turing Triage Test," *Ethics and Information Technology* 6 (2004): 201-213, especially 204.

[33] Anderson, "Asimov's Three Laws."

[34] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, ed. J. Burns, and H. Hart (Oxford: Clarendon Press, 1789), especially Chapter 17: "Boundary around Penal Jurisprudence." Also Peter Singer, "All Animals Are Equal," in *Animal Ethics: Past and Present Perspectives*, ed. Evangelos D. Protopapadakis, 163-178 (Berlin: Logos Verlag, 2012); Peter Singer, *Animal Liberation: A New Ethics for our Treatment of Animals* (New York: New York Review of Books, 1975); Peter Singer, *Practical Ethics*, 2nd edition (Cambridge: Cambridge University Press, 1993).

[35] Immanuel Kant, "Our Duties to Animals," in his *Lectures on Ethics*, trans. L. Infield, 239-241 (New York: Harper & Row, 1963).

[36] Michael Tooley, "In Defense of Abortion and Infanticide," In *The Abortion Controversy: A Reader*, ed. Luis P. Pojman, and Francis J. Beckwith, 186-213 (Boston, MA: Jones & Bartlett, 1994).

[37] Mary Anne Warren, "On the Moral and Legal Status of Abortion," in *Contemporary Moral Problems*, ed. J. White, 144-155 (Belmont, CA: Wadsworth/Thompson Learning, 2003).

[38] For an analysis of the flaws in the proposed criteria regarding the attribution of moral status to AI entities, see Alkis Gounaris, and George Kosteletos, "Licensed to Kill: Autonomous Weapons as Persons and Moral Agents," in *Personhood*, ed. Dragan Prole, and Goran Rujević, 137-189 (Novi Sad: The NKUA Applied Philosophy Research Lab Press, 2020).

[39] For the way in which the Other Minds Problem could enter the discussion of AI Ethics and the application of moral status to the AI agents see D. Gunkel, *The Machine Question: Critical Perspectives on AI, Robots and Ethics* (Cambridge, MA: MIT Press, 2012); C. Hoffmann, and B. Hahn, "Decentered Ethics in the Machine Era and Guidance for AI Regulation," *AI & Society* 35, no. 3 (2009): 635-644; D. Levy, "The Ethical Treatment of Artificially Conscious Robots," *International Journal of Social Robotics* 1, no. 3 (2009): 209-216.

problems.[40] Do you think bioethicists or philosophers in general should be concerned about their work in the future? Will the machines be able to replace them, at some point, completely? Could machines become the 'philosophers' of a new Plato's *Republic*?

Susan Leigh Anderson: I do think that there is a possibility of there being more objectivity in machine decision-making, if properly designed; but new issues are bound to arise (conditions change) that would require up-dates. And an important philosophical question will never disappear: What gives our lives meaning?

Alkis Gounaris & George Kosteletos: How much do you really think we are in danger from AI? Some argue that the cooperation of the AI and the Biotechnology fields will lead to new forms of intelligence in the near future.[41] What should we hope for and what should we fear about that? We see the media, pop writers like Harari,[42] businessmen like Musk,[43] and the academic community as well (e.g. Bostrom[44] or Tegmark[45] Institutes) holding a cautious

---

[40] Michael Anderson, and Susan Leigh Anderson, "MedEthEx: A Prototype Medical Ethics Advisor," *Proceedings of the 21st National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*, 1759-1765 (Boston, MA: AAAI Press, 2006), http://dblp.uni-trier.de/db/conf/aaai/aaai2006.html#AndersonAA06; Michael Anderson, and Susan Leigh Anderson, "ETHEL: Toward a Principled Ethical Eldercare System," *Proceedings of the AAAI Fall Symposium: New Solutions to Old Problems. Technical Report FS-08-02* (Arlington, VA, 2008); Michael Anderson, and Susan Leigh Anderson, "Robot Be Good," *Scientific American* 303, no. 4 (2010): 72-77. Also, Michael Anderson, and Susan Leigh Anderson, "A Prima Facie Duty Approach to Machine Ethics: Machine Learning of Features of Ethical Dilemmas, Prima Facie Duties, and Decision Principles through a Dialogue with Ethicists," in *Machine Ethics*, ed. Michael Anderson, and Suzan Leigh Anderson, 476-492 (New York and Cambridge: Cambridge University Press, 2011); Susan Leigh Anderson, and Michael Anderson, "Towards a Principle-Based Healthcare Agent," in *Machine Medical Ethics*, ed. S. van Rysewyk, and M. Pontier, 67-77 (Cham: Springer, 2015).

[41] Yuval Noah Harari, *21 Lessons for the 21st Century* (New York: Spiegel & Grau, 2018).

[42] Nicholas Thompson, "Will Artificial Intelligence Enhance or Hack Humanity?" *Wired*, April 20, 2019, https://www.wired.com/story/will-artificial-intelligence-enhance-hack-humanity/.

[43] Catherine Clifford, and Elon Musk: "Mark my Words – A.I. is far more Dangerous than Nukes." *CNBC*, March 13, 2018, https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html; Gregory Wallace, "Elon Musk Warns against Unleashing Artificial Intelligence 'Demon,'" *CNN Business*, October 26, 2014, https://money.cnn.com/2014/10/26/technology/elon-musk-artificial-intelligence-demon/; Ricki Harris, "Elon Musk: Humanity Is a Kind of 'Biological Boot Loader' for AI," *Wired*, January 9, 2019, https://www.wired.com/story/elon-musk-humanity-biological-boot-loader-ai/.

[44] For more, visit *The Future of Humanity Institute*, https://www.fhi.ox.ac.uk/; Nick Bostrom, "Existential Risk Prevention as Global Priority," *Global Policy* 4, no. 1 (2013): 15-31.

[45] For more, visit https://futureoflife.org/; Max Tegmark, "Benefit and Risks of Artificial Intelligence," *Future of Life Institute*, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/; Stuart Russell, et al., "Autonomous Weapons: An Open Letter from AI & Robotics Researchers," *Future of Life Institute*, https://futureoflife.org/open-letter-

and in some cases technophobic attitude in the face of developments. How could we distinguish the real risks from the pseudo-problems?

Susan Leigh Anderson: Whether we have to fear AI technology depends on how it is developed and by whom it is used. In itself it is neutral. If we develop AI entities on a human model, embodying negative human qualities (like self-centeredness, favoring one's own group) and allow anyone to use them, they could become super weapons. This is why the field of Machine Ethics is so important. We have the opportunity to create *ethical machines*, non-threatening machines that not only aid us in many ways, but can also show us how we need to behave if we are to survive as a species.

Michael Anderson: For all the perils of doing so, I see no other option than trusting science. Yes, it has led us into dangers that we might not have faced if we had kept our blinders on but it has also been the shining light that has taken humanity out of the darkness, illuminating many mysteries of the universe. Given the risks humanity lives under, my hope for AI is that it might serve as a means for preserving intelligence. As it stands, this is currently only housed in human bodies – a vessel so fragile that it might be prudent to develop backup for it. Wouldn't it be the ultimate tragedy if we were the only intelligent creatures in the universe and, through inaction, let our unique spark die out?

Alkis Gounaris & George Kosteletos: Let us now come to something even more current. In your opinion, which are the most prominent ethical challenges raised by the COVID-19 pandemic? Could Machine Ethics contribute in facing them? Could these ethical challenges be faced more successfully by an AI agent equipped with moral principles, than by human committees of doctors, epidemiologists, politicians and bioethicists? Finally, does this pandemic crisis provide the Machine Ethics research program with any lessons to be learned and used in similar crises in the future? What do you suggest so that the public would be prepared for such contributions by the AI agents?

Susan Leigh Anderson: What machines are good at (better than humans) is digesting a lot of data quickly: discovering connections, etc. Humans are still needed to input the data and ethicists are more likely to insist that the data is not skewed to gloss over ethical issues. For instance, one could just keep track of whether people are offered vaccines, just noting that fewer members of minority communities seem to be taking them, ignoring past legitimate concerns in these communities about taking vaccines and whether attempts have been made to educate them, or whether the means for notifying them

---

autonomous-weapons/; Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Knopf, 2017).

of their chance to take the vaccine is likely to reach them (if through the internet: whether they have access to the internet and the skill at navigating it).

Michael Anderson: What machine ethics has to offer is consistent, impartial treatment of like cases. In the face of seemingly novel ethical challenges, it is hoped that this might prove useful in illuminating similarities to previous challenges thereby contributing to current ones.

Alkis Gounaris & George Kosteletos: In a recent publication,[46] Michael among others, endorses the position that inclusive, interdisciplinary teams are needed to develop AI. What do you think that the role of philosophers is in such an endeavor?

Michael Anderson: What seems to elude many is that there is expertise in ethics as there is in any academic discipline. This misapprehension seems to stem from the fact that people make "ethical" decisions daily and therefore have difficulty understanding why such expertise is needed. That said, doesn't it seem obvious that those who have spent their research careers in a field might have greater insight into it? Clearly, the intuitive approach most bring to such decisions is riddled with partiality and inconsistency, not to mention a circumscribed understanding of the plethora of factors involved. The expertise ethicists bring to the table is necessary to help alleviate these shortcomings.

Alkis Gounaris & George Kosteletos: What remains to be achieved? Which would be the key concerns and the basic challenges of Machine Ethics in the future? Should we expect in the near future a safe, ethical and/or responsible AI?

Michael Anderson: Not a soothsayer but it's pretty clear to me that autonomous systems are here to stay and it would be unwise to ignore their ethical tuition. Unfortunately, given its need for copious data and the dearth of such data in the domain of ethics, the silver bullet of deep learning does not seem to have much to offer to this issue. Where value judgements are involved, it seems that we are going to have to bite the bullet and do the hard work of determining just how we want such systems to behave towards us.

Alkis Gounaris & George Kosteletos: Susan, Michael, thank you for the extremely interesting discussion and we look forward to having you with us at our upcoming *Me and AI: Human Concerns Artificial Minds* Conference.

Susan Leigh & Michael Anderson: Thank you for your thought-provoking questions! Your conference could not come at a more opportune time!

---

[46] Steve Taylor, et al., "Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation," *Zenodo*, July 2, 2018.

References

Aldinhas, Ferreira Maria, João Silva Sequeira, Gurvinder Singh Virk, Mohammad Tokhi Osman, and Ender E. Kadar, eds. *Robotics and Well-being*. Cham: Springer, 2019.

Allen, Colin, Varner Gary, and Zinser Jason. "Prolegomena to Any Future Artificial Moral Agent." *Journal of Experimental and Theoretical Artificial Intelligence* 12 (2000): 151-261.

Anderson, Michael, and Suzan Leigh Anderson, eds. *Machine Ethics*. New York and Cambridge: Cambridge University Press, 2011.

Anderson, Michael, and Suzan Leigh Anderson. "A Prima Facie Duty Approach to Machine Ethics: Machine Learning of Features of Ethical Dilemmas, Prima Facie Duties, and Decision Principles through a Dialogue with Ethicists." In *Machine Ethics*, edited by Michael Anderson, and Suzan Leigh Anderson, 476-492. New York and Cambridge: Cambridge University Press, 2011.

Anderson, Michael, and Suzan Leigh Anderson. "ETHEL: Toward a Principled Ethical Eldercare System." *Proceedings of the AAAI Fall Symposium: New Solutions to Old Problems*. Technical Report FS-08-02. Arlington, VA, 2008.

Anderson, Michael, and Suzan Leigh Anderson. "Guest Editors' Introduction: Machine Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006): 10-11.

Anderson, Michael, and Suzan Leigh Anderson. "Machine Ethics: Creating an Ethical Intelligent Agent." *AI Magazine* 28, no. 4 (2007): 15-26.

Anderson, Michael, and Suzan Leigh Anderson. "MedEthEx: A Prototype Medical Ethics Advisor." *Proceedings of the 21st National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*, 1759-1765. Boston, MA: AAAI Press, 2006.

Anderson, Michael, and Suzan Leigh Anderson. "Robot Be Good." *Scientific American* 303, no. 4 (2010): 72-77.

Anderson, Michael, and Suzan Leigh Anderson. "The Status of Machine Ethics: A Report from the AAAI Symposium." *Minds & Machines* 17 (2007): 1-10.

Anderson, Michael, and Suzan Leigh Anderson. "Toward Ensuring Ethical Behavior from Autonomous Systems: A Case-supported Principle-based Paradigm." *Industrial Robot* 42, no. 4 (2015): 324-331.

Anderson, Michael, Suzan Leigh Anderson, and Chris Armen, eds. *Machine Ethics: Papers form AAAI Fall Symposium, 2005*. Technical Report FS-05-06. Menlo Park, CA: Association for the Advancement of Artificial Intelligence, 2005. https://www.aaai.org/Library/Symposia/Fall/fs05-06.php.

Anderson, Michael, Suzan Leigh Anderson, and Chris Armen. "An Approach to Computing Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006): 65-63.

Anderson, Michael, Suzan Leigh Anderson, and Chris Armen. "Toward Machine Ethics: Implementing Two Action-Based Ethical Theories." In *Machine Ethics, Papers form AAAI Fall Symposium, 2005*, edited by Michael Anderson, Suzan Leigh Anderson, and Chris Armen, Technical Report FS-05-06. Menlo Park, CA: Association for the Advancement of Artificial Intelligence, 2005.

Anderson, Michael, Suzan Leigh Anderson, and Chris Armen. "Towards Machine Ethics." In *Proceedings of the AAAI-04 Workshop on Agent Organizations: Theory and Practice*, 53-59. San Jose, CA, 2004.

Anderson, Suzan Leigh, and Michael Anderson. "Towards a Principle-Based Healthcare Agent." In *Machine Medical Ethics*, edited by S. van Rysewyk, and M. Pontier, 67-77. Cham: Springer, 2015.

Anderson, Suzan Leigh. "Asimov's 'Three Laws of Robotics' and Machine Metaethics." *AI and Society* 22 (2007): 477-493.

Anderson, Suzan Leigh. "Machine Metaethics." In *Machine Ethics*, edited by Michael Anderson, and Suzan Leigh Anderson, 21-27. New York and Cambridge: Cambridge University Press, 2011.

Anderson, Suzan Leigh. "Philosophical Concerns with Machine Ethics." In *Machine Ethics*, edited by Michael Anderson, and Suzan Leigh Anderson, 162-167. New York and Cambridge: Cambridge University Press, 2011.

Anderson, Suzan Leigh. "The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics." In *Machine Ethics*, edited by Michael Anderson, and Suzan Leigh Anderson, 285-296. New York and Cambridge: Cambridge University Press, 2011.

Asimov, Isaac. "The Bicentennial Man." In *Philosophy and Science Fiction*, edited by Michael Phillips, 183-216. Buffalo, NY: Prometheus Books, 1984.

Awad, Edmond, Michael Anderson, Suzan Leigh Anderson, and Beishui Liao. "An Approach for Combining Ethical Principles with Public Opinion to Guide Public Policy." *Artificial Intelligence* 287 (2020): article 103349.

Awad, Edmond, Sohan Dsouza, Azim Shariff, Iyad Rahwan, and Jean-Francois Bonnefon. "Universals and Variations in Moral Decisions Made in 42 Countries by 70,000 Participants." *Proceedings of the National Academy of Sciences* 117, no. 5 (2020): 2332-2337.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. "The Moral Machine Experiment." *Nature* 563, no. 7729 (2018): 59-64.

Beauchamp, Tom Lamar, and James Franklin Childress. *Principles of Biomedical Ethics*. Oxford, UK: Oxford University Press, 1979.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Edited by J. Burns, and H. Hart. Oxford: Clarendon Press, 1789.

Bonnefon, Jean-Francois, Azim Shariff, and Iyad Rahwan. "The Social Dilemma of Autonomous Vehicles." *Science* 352, no. 6293 (2016): 1573-1576.

Bostrom, Nick. "Existential Risk Prevention as Global Priority." *Global Policy* 4, no. 1 (2013): 15-31.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.

Clarke, Roger. "Asimov's Laws of Robotics: Implications for Information Technology. Part I." *Computer* 26, no. 12 (1993): 53-61.

Clarke, Roger. "Asimov's Laws of Robotics: Implications for Information Technology. Part II." *Computer* 27, no. 1 (1994): 57-66.

Clifford, Catherine, and Elon Musk. "Mark my Words – A.I. is far more Dangerous than Nukes." *CNBC*, March 13, 2018. https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html.

Davis, Martin. *Computability and Unsolvability*. New York: McGraw-Hill, 1958.

Dennett, Daniel. "Cognitive Wheels: The Frame Problem of AI." In *Minds, Machines and Evolution: Philosophical Studies*, edited by C. Hoockway, 129-152. Cambridge: Cambridge University Press, 1984.

Dennett, Daniel. "When Hal Kills, Who's to Blame? Computer Ethics." In *Hal's Legacy: 2001's Computer as Dream and Reality*, edited by David G. Stork, 351-365. Cambridge, MA: MIT Press, 1997.

Dennett, Daniel. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, MA: MIT Press, 1978.

Dreyfus, Hubert Lederer. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press, 1992.

Feenberg, Andrew. "Subversive Rationalization: Technology, Power, and Democracy." In *Technology and the Politics of Knowledge*, edited by Andrew Feenberg, and Alastair Hannay, 3-11. Bloomington and Indianapolis: Indiana University Press, 1995.

Feenberg, Andrew. *Questioning Technology*. London, New York: Routledge, 1999.

Floridi, Luciano, and J. W. Sanders. "On the Morality of Artificial Agents." *Minds and Machines* 14 (2004): 349-379.

Fodor, Jerry Alan. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

Gounaris, Alkis, and George Kosteletos. "Licensed to Kill: Autonomous Weapons as Persons and Moral Agents." In *Personhood*, edited by Dragan Prole, and Goran Rujiević, 137-189. Hellenic-Serbian Philosophical Dialogue Series, vol. 2. Novi Sad: The NKUA Applied Philosophy Research Lab Press, 2020.

Gunkel, David. *The Machine Question: Critical Perspectives on AI, Robots and Ethics*. Cambridge, MA: MIT Press, 2012.

Hao, Karen. "Should a Self-driving Car Kill the Baby or the Grandma? Depends on where You're from." *MIT Technology Review*, October 14, 2018, https://www.technologyreview.com/2018/10/24/139313/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/.

Harari, Yuval Noah. *21 Lessons for the 21st Century*. New York: Spiegel & Grau, 2018.

Harris, Ricki. "Elon Musk: Humanity Is a Kind of 'Biological Boot Loader' for AI." *Wired*, January 9, 2019, https://www.wired.com/story/elon-musk-humanity-biological-boot-loader-ai/.

Hobbes, Thomas. *Leviathan, or The Matter, Forme and Power of a Commonwealth Ecclesiastical and Civil*. Edited by A. R. Waller. Cambridge: Cambridge University Press, 1904.

Hoffmann, Christian Hugo, and Benjamin Hahn. "Decentered Ethics in the Machine Era and Guidance for AI Regulation." *AI & Society* 35, no. 3 (2009): 635-644.

Kant, Immanuel. *Critique of Practical Reason*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 2015.

Kant, Immanuel. *Lectures on Ethics*. Translated by L. Infield. New York: Harper & Row, 1963.

Kant, Immanuel. *The Groundwork for the Metaphysics of Morals*. Translated by Allen W. Wood. New Haven and London: Yale University Press, 2002.

Kleene, Stephen Cole. *Introduction to Metamathematics*. Amsterdam: North-Holland, 1952.

Leibniz, Gottfried Wilhelm. "Principles of Nature and Grace, Based on Reason." In Gottfried Wilhelm Leibniz, *Philosophical Papers and Letters*, edited by Leroy E. Loemker. Dordrecht: Springer, 1989.

Leibniz, Gottfried Wilhelm. *Dissertatio de arte combinatoria*. Paris: Hachette Livre-BNF, 2018.

Levy, David. "The Ethical Treatment of Artificially Conscious Robots." *International Journal of Social Robotics* 1, no. 3 (2009): 209-216.

Malle, Bertram F., Thapa Stuti Magar, and Matthias Scheutz. "AI in the Sky: How People Morally Evaluate Human and Machine Decisions in a Lethal Strike Dilemma." In *Robotics and Well-Being*, edited by Maria Aldinhas Ferreira, João Silva Sequeira, Gurvinder Singh Virk, Mohammad Tokhi Osman, and Ender E. Kadar, 111-133. Cham: Springer, 2019.

McCarthy, John, and Patrick J. Hayes. "Some Philosophical Problems from the Standpoint of Artificial Intelligence." In *Machine Intelligence*, vol. 4, edited by Bernard Meltzer, and Donald M. Michie, 463-502. Edinburgh: Edinburgh University Press, 1969.

McCulloch, Warren S., and Walter H. Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5 (1943): 115-33.

Minsky, Marvin. *Computation: Finite and Infinite Machines*. New Jersey: Prentice-Hall, 1967.

Mitcham, Carl. *Thinking through Technology: The Path between Engineering and Philosophy*. Chicago: The University of Chicago Press, 1994.

Moor, James H. "The Nature, Importance, and Difficulty of Machine Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006): 18-21.

Newell, Allen, and Herbert Alexander Simon. "Computing Science as Empirical Enquiry: Symbols and Search." *Communications of the Association for Computing Machinery* 19 (1976): 113-126.

Newell, Allen, and Herbert Alexander Simon. "GPS-A Program that Simulates Human Thought." In *Lernende Automaten*, edited by Heinz Billing, 109-124. Münich: Oldenburg, 1961.

Newell, Allen, and Herbert Alexander Simon. "The Logic Theory Machine: A Complex Information-Processing System." *IRE Transactions on Information Theory* 2, no. 3 (1956): 61-79.

Newell, Allen, and Herbert Alexander Simon. *Current Developments in Complex Information Processing: Technical Report P-850*. Santa Monica, CA: Rand Corporation, 1956.

Newell, Allen, and John Crosley Shaw. "Programming the Logic Theory Machine." In *IRE-AIEE-ACM '57 (Western): Papers Presented at the February 26-28, 1957, Western Joint Computer Conference: Techniques for Reliability*, 230-240. New York: Association for Computing Machinery, 1957.

Newell, Allen, John Crosley Shaw, and Herbert Alexander Simon. "Element of a Theory of Human Problem Solving." *Psychological Review* 65 (1958): 151-166.

Newell, Allen. "Physical Symbol Systems." *Cognitive Science* 4 (1980): 135-183.

Owen, Jonathan, and Richard Osley. "Bill of Rights for Abused Robots: Experts Draw up an Ethical Charter to Prevent Humans Exploiting Machines." *The Independent*, September 17, 2011, https://www.independent.co.uk/news/science/bill-of-rights-for-abused-robots-5332596.html.

Pylyshyn, Zenon W., ed. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex, 1987.

Rawls, John. "Justice as Fairness: Political not Metaphysical." *Philosophy and Public Affairs* 14 (1985): 223-251.

Rawls, John. *A Theory of Justice*, 2nd edition. Cambridge, MA: The Belknap Press of Harvard University Press, 1999.

Rawls, John. *A Theory of Justice*. Cambridge, MA: The Belknap Press of Harvard University Press, 1971.

Rawls, John. *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, 2001.

Ross, William David. *The Right and the Good*. Oxford: Clarendon Press, 1930.

Russell, Stuart, and Max Tegmark. "Autonomous Weapons: An Open Letter from AI & Robotics Researchers." *Future of Life Institute*. https://futureoflife.org/open-letter-autonomous-weapons/.

Singer, Peter. "All Animals Are Equal." In *Animal Ethics: Past and Present Perspectives*, edited by Evangelos D. Protopapadakis, 163-178. Berlin: Logos Verlag, 2012.

Singer, Peter. *Animal Liberation: A New Ethics for our Treatment of Animals*. New York: New York Review of Books, 1975.

Singer, Peter. *Practical Ethics*, 2nd edition. Cambridge: Cambridge University Press, 1993.

Soares, Nate. "The Value Learning Problem." In *Artificial Intelligence, Safety and Security*, edited by Roman V. Yampolskiy, 89-97. Boca Raton, FL: CRC Press, 2019.

Sparrow, Robert. "Killer Robots." *Journal of Applied Philosophy* 24, no. 1 (2007): 62-77.

Sparrow, Robert. "The Turing Triage Test." *Ethics and Information Technology* 6 (2004): 201-213.

Taylor, Steve, Brian Pickering, Michael Boniface, Michael Anderson, David Danks, Asbjørn Følstad, Matthias Leese, Vincent Müller, Tom Sorell, Alan Winfield, and Fiona Woollard. "Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation." *Zenodo*, July 2, 2018.

Tegmark, Mark. "Benefit and Risks of Artificial Intelligence." *Future of Life Institute*. https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/.

Tegmark, Mark. *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Knopf, 2017.

Thompson, Nicholas. "Will Artificial Intelligence Enhance or Hack Humanity?" *Wired*, April 20, 2019, https://www.wired.com/story/will-artificial-intelligence-enhance-hack-humanity/.

Tooley, Michael. "In Defense of Abortion and Infanticide." In *The Abortion Controversy: A Reader*, edited by Luis P. Pojman, and Francis J. Beckwith, 186-213. Boston, MA: Jones & Bartlett, 1994.

Turing, Alan Mathison. "Computing, Machinery and Intelligence." *Mind* 59 (1950): 433-460.

Turing, Alan Mathison. "Intelligent Machinery." In *Machine Intelligence 5*, edited by B. Meltzer, and D. M. Michie, 3-23. Edinburgh: Edinburgh University Press, 1969.

Turing, Alan Mathison. "On Computable Numbers, with an Application to the Entscheidungsproblem." In *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life*, edited by Jack B. Copeland, 58-90. Oxford: Oxford University Press, 2004.

Turing, Alan Mathison. "On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction." *Proceedings of the London Mathematical Society* 43 (1938): 544-546.

Wallace, Gregory. "Elon Musk Warns against Unleashing Artificial Intelligence 'Demon.'" *CNN Business*, October 26, 2014, https://money.cnn.com/2014/10/26/technology/elon-musk-artificial-intelligence-demon/.

Warren, Mary Anne. "On the Moral and Legal Status of Abortion." In *Contemporary Moral Problems*, edited by J. White, 144-155. Belmont, CA: Wadsworth/Thompson Learning, 2003.

Wheeler, Michael. "Cognition in Context: Phenomenology, Situated Robotics, and the Frame Problem." *International Journal of Philosophical Studies* 16, no. 3 (2008): 323-349.

Wheeler, Michael. *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press, 2005.

Winner, Langdon. "Citizen Virtues in a Technological Order." *Inquiry* 35, nos. 3-4 (1992): 341-361.

Winner, Langdon. "Technè and Politeia: The Technical Constitution of Society." In *Philosophy of Technology*, edited by Paul T. Dubrin, and Friedrich Rapp, 97-111. Dordrecht, Boston, Lancaster: D. Reidel, 1983.

Yampolskiy, Roman. "Artificial Intelligence Safety Engineering: Why Machine Ethics is a Wrong Approach." In *Philosophy and Theory of Artificial Intelligence. Studies in Applied Philosophy, Epistemology and Rational Ethics*, edited by Vincent Müller, 389-396. Berlin, Heidelberg: Springer, 2013.

Yudkowsky, Eliezer. "Complex Value Systems in Friendly AI." In *Artificial General Intelligence*, edited by Jürgen Schmidhuber, Kristinn R. Thórisson, and Moshe Looks, 388-393. Berlin, Heidelberg: Springer, 2011.

Yudkowsky, Eliezer. "The Value Loading Problem." *EDGE*, July 12, 2021, https://www.edge.org/response-detail/26198.