

## ***How to be Good: The Possibility of Moral Enhancement*<sup>1</sup>**

**By: John Harris. Oxford: Oxford University Press, 2016, 195 pp.**

**Review by: Dr. Ayman Shabana**

Advances in neuroscience have raised the prospect of cognitive enhancement beyond mere physical enhancement in an effort to improve one's intellectual abilities and performance. Experiments on the human brain to influence a person's behavior, as well as underlying intentions and motivations, open the door for exploring the possibility of moral enhancement through biomedical means. Although the achievement of moral enhancement through biomedical intervention remains uncertain, if unlikely at least in the foreseeable future, the idea is quite intriguing. After all, the ability to influence the human brain to promote good choices and thwart bad ones has always been the ultimate goal of moral philosophy. In the age of biomedical revolution, moral (bio)enhancement can be seen as the most effective method to offset the untold dangers of scientific ambitions, especially ones associated with abuse of scientific research. However, the idea of moral bioenhancement raises important ethical concerns over its potential impact on fundamental human freedom to choose one's own actions and to bear responsibility for them.

In *How to be Good: The Possibility of Moral Enhancement*, John Harris argues that moral bioenhancement would conflict with the essence of moral action, which is inextricably tied to human freedom. The book is divided into twelve chapters. The first two introductory chapters open the discussion by highlighting the importance of morality and its myriad practical implications for contemporary neuroscience, social policy and criminal justice. The first chapter poses the question of how to be good and focuses on the limitations of technical biomedical means on engendering moral enhancement. For Harris, one of the main reasons why moral enhancement is problematic has to do with the difficulty in predicting the exact impact of modern technology on human nature or even on the future of mankind on earth. Harris predicts that humans in the future will have either disappeared (due to scientific miscalculations or natural forces) or undergone further degrees of evolution: "In the future there will be no more 'men' in Plato's sense, no more human beings therefore, and no more planet earth" (p. 2, 50). If the continuity of humanity (as we now know it) is in question, what then is the point of investigating the future of morality? The answer that Harris gives focuses on the difference made in the meantime either at the individual (being good) or collective level (doing good). Morality is conducive to the achievement of maximum levels of benefits as well as the safety of as many people as possible. The importance of morality is implicit in all social norms underscoring the value of human life. The importance of morality cannot, therefore, be overemphasized and moral enhancement, in the sense of enabling as many people as possible to make good/better moral choices, seems to be a worthwhile pursuit. What is debatable, however, is whether such enhancement can be better achieved by traditional methods such as education or through biomedical intervention. The second chapter explores some of the main features that Harris associates with morality and moral deliberation such as generality, objectivity, rationality, as well as their importance for human

---

<sup>1</sup> "This research was made possible by the NPRP grant "Indigenizing Genomics in the Gulf Region (IGGR): The Missing Islamic Bioethical Discourse", no. NPRP8-1620-6-057 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the author[s]."

welfare and flourishing. The account that Harris gives is mainly philosophical, but he also draws attention to its relevance to religiously-grounded theories of morality: “it is our ability to reason about the nature of the good independently ... that partially accounts for theology and indeed enables us to say non-vacuously that God is, or that the gods are good” (p. 31). This argument echoes extensive debates within the Islamic theological tradition, particularly among the rationalist Mu‘tazilis and traditionalist Ash‘aris over both the nature and sources of morality. Harris concludes the chapter by highlighting a major distinction in moral philosophy between judging characters (virtue ethics) and judging actions, as well as their outcomes (consequentialism or deontology). He suggests that one of the main limitations of moral bioenhancement would be that, if it materializes, it would apply more to virtue ethics. Evaluating actions and their consequences, on the other hand, would require rational deliberation. In the absence of free and independent rational deliberation, the morality of the action in question would be suspected.

If Harris opposes the idea of moral bioenhancement due to its impact on human freedom, this does not mean that he is against the notion of enhancement per se. By contrast, as his previous works make sufficiently clear, Harris celebrates scientific advances promising to improve human abilities. He welcomes the prospect that future generations (successors) will be able to enjoy such enhanced capacities and capabilities. The third chapter builds on this prospect and anticipates that successors to the current human species may not be entirely human. Therefore, he suggests major modifications to the current formulations of the concept of human rights to include partially human, or even non-human, rational persons/machines. For those who think that this is a far-fetched possibility, Harris argues that biomedical technology has already enabled “crossing of species boundaries” beyond classical fictional figures, such as centaurs and mermaids, whether in diet, being administered various types of medical treatment, and finally mixing of human and animal genes. Needless to say, Harris’s account is predicated on the Darwinian theory of evolution postulating a gradual and incremental “seamless transition between apes and humans.” Although this theory remains problematic, especially from a religious perspective, some scholars have suggested ways to reconcile evolution with a God-created universe.<sup>2</sup> Ultimately, Harris sees humanity as a long continuous process of evolution and therefore criticizes “human chauvinism”, which ends up holding back free scientific inquiry (p. 41). He, therefore, calls for embracing the limitless possibilities of scientific research even if this means the end of humanity as we know it in the hope that it would be replaced by a more enhanced type of human existence. For Harris, the most effective way to address ethical problems attending scientific inquiry is through “the best estimation of risk as against benefit” (55).

In the fourth and fifth chapters, Harris presents his main argument against moral bioenhancement. His critique focuses on the limits that such enhancement would place on fundamental human freedom and responsibility, including freedom to make morally questionable choices. He draws on John Milton’s account of the creation story in *Paradise Lost*, particularly on how man was created “sufficient to have stood, though free to fall.” In his criticism against the idea of moral bioenhancement, Harris points out

---

<sup>2</sup> See, for example, Nidhal Guessoum, *Islam’s Quantum Question: Reconciling Muslim Tradition and Modern Science* (London: I. B. Tauris, 2011), pp.271-324.

two main problems. First, the most morally questionable attitudes (e.g. racial discrimination) are based on cognitive assumptions and they, therefore, have to be corrected at that level. Second, negative emotions such as aversion are not entirely evil. Targeting such emotions medically risks compromising proper uses of such emotions. Harris, therefore, argues against the conflation of moral and medical inducements of “pro-social attitudes” (p. 78). He rejects moral bioenhancement as the safest measure to balance the perils of cognitive enhancement and argues that the cost of withholding cognitive enhancement until an effective method of moral bioenhancement becomes available could be much worse because many outstanding problems could be resolved by cognitive enhancement in the meantime. Moreover, until effective and reliable methods of moral bioenhancement are in place, cognitive enhancement remains the best alternative that humans have for moral enhancement (p.74, 84).

In chapter six, Harris continues his examination of the role of freedom in moral decision-making and the extent to which negation of freedom amounts to diminished autonomy, and hence responsibility for one’s actions. He criticizes propositions to separate thought from action to achieve behavioral control since this would negatively impact on the full moral agency. Even if it were possible to develop a “God Machine” that would, as Ingmar Persson and Julian Savulescu suggest, intervene to prevent major immoral acts, Harris argues that this technical device would fail due to the lack of sufficient cognitive ability to understand contextual details (e.g. difference between justified and unjustified use of violence). Another difficulty with devices such as the God-Machine would be the difficulty to assess potential implications on human conduct and even the ability to subsequently switch it off.

Chapters seven and eight discuss reductionist accounts of morality. In chapter seven, Harris focuses on the relationship between morality and emotions and criticizes approaches to moral enhancement that rely entirely on emotions, consequently supporting moral enhancement through medically induced emotions. For Harris, morality involves not only knowing, but also verifying as well as implementing the good. From this perspective, mere emotions or intuition, without rational examination, cannot serve as criteria to identify or verify moral positions. He reiterates his argument that in the absence of effective means for moral bioenhancement, cognitive enhancement can also serve to improve moral sensibility (p. 117). In chapter eight, Harris focuses on materialist accounts of morality, especially in light of theories of neuroscience explaining morality in terms of evolutionary origins and biological brain functions. Once more, Harris criticizes the emphasis placed on emotions, to the exclusion of rationality, especially when viewed as an expression of the evolutionary constraints that humans have to come to terms with. While Harris does not deny the evolutionary origins of morality, he argues against appealing to these origins to support the reduction of morality to emotions. If we ignore rational judgment, there will hardly be other reliable means to verify or evaluate emotions other than another experience of the same emotions: “If it is simply one feeling confirming another, then we really are in the situation Wittgenstein lampooned as buying a second copy of the same newspaper to confirm the truth of what we read in the first.”

In chapter nine, Harris turns to the issue of moral progress and the extent to which it can be facilitated by moral bioenhancement. He takes issue with one of the main arguments for moral bioenhancement in this

age of unprecedented technical advancement, including availability of weapons of mass destruction: capacity to do harm, even by a single individual or a few individuals, is much greater than either preventing harm or doing good. Considering the urgent need for moral bioenhancement to redress a risk of this magnitude, supporters argue that moral bioenhancement is not only required, but should even be imposed globally, without, however, indicating how this can be implemented. Although propositions to undertake moral bioenhancement tend to concentrate on solutions at the individual level, the problems for which moral bioenhancement is suggested require coordinated action at the collective and international levels. A strategy that Harris proposes, instead of medically induced moral bioenhancement, involves a plan to both raise living standards to achieve affluence and to raise people satisfaction to ensure empowerment and fight cynicism.

In chapter ten, Harris discusses the notion of mind reading as a tool of moral enhancement. Access to brain states has the potential of revealing important information that could be quite valuable in resolving complex issues of intent in criminal cases. In the past, there were efforts to access brain states through examination of external body parts such as one's face or eyes. Although, to date, no technology has been able to offer accurate mind reading through interpretation of physical brain states, Harris argues that the cloud (in the sense of any form of digitized data) has the potential of revealing a great deal about the mental states of users. However, as much as methods or techniques used to analyze brain states may offer important clues for mind-reading, they equally have the potential to result in mind-misreading. This is mainly due to one important factor that mind-reading techniques/mechanisms lack: accurate knowledge or interpretation of contextual details.

In chapter eleven, Harris places the discussion within a larger social context by pointing out both the collective dimensions of morality as well as the responsibility of the state for people's safety and liberty. As most articulations of the social contract theory illustrate, the relationship between the political sovereign and the people is governed by a contractual framework within which people obey political authority in return for guaranteeing their safety and freedom. Harris's discussion suggests that, to an extent, protecting people's safety and liberty is viewed as the primary responsibility of political authority; the latter should be concerned with any measure that risks diminishing the safety and freedom of the people, moral bioenhancement included.

In the final chapter, Harris addresses the potential role of moral bioenhancement in responding to the increasing dangers of modern technology, especially in light of trans-humanist tendencies anticipating the augmentation of human capabilities beyond biological limitations, including sickness and even death. Harris also addresses the interface between humans and artificial intelligence, which may occur in two main ways. The first involves a degree of interaction between humans or organic individuals and particular forms of artificial intelligence. This type of interface is already happening through the various smart applications on which people have increasingly become dependent upon. The second envisions the development of particular applications/forms of artificial intelligence that will become creatures in their own right, independently of humans. One important question that Harris raises is whether these forms of

artificial intelligence will remain subject to full human control. Another question involves the extent to which an artificially intelligent being could be considered a person. One answer that Harris suggests is if such a 'being' develops the capability to recognize the value of its existence as well as that of others. Another answer, suggested by Alan Turing, is if humans fail to distinguish the performance of an artificially intelligent being from the performance of a human. However, even if artificially intelligent beings are able to meet these criteria, important questions on how they will coexist with humans remain open. Ultimately, moral consideration of the relationship between human and trans-human persons or machines would require mutual understanding, which is difficult to obtain in the absence of actual concrete examples.

In this book, Harris draws on his earlier work to reiterate his emphasis on human freedom, including freedom to explore scientific methods to enhance human capabilities. He condemns biomedical means to achieve moral enhancement because of the constraints that such efforts would entail. The main argument that Harris makes in this book is, in the absence of an effective alternative, the best way to augment human moral sensibilities is through cognitive enhancement. Considering the difficulty, if not the futility or even undesirability, of forestalling scientific progress, moral reasoning, rather than inducement of moral emotions, should consist in the most accurate calculation of potential benefits against foreseeable risks. However, in light of the distinction that Harris makes between moral and cognitive enhancement, one may wonder whether, or to what extent, the reservations that Harris expresses regarding moral enhancement could also apply to the notion of enhancement itself. As noted above, his objection to medically-induced morality lies in the restrictions that it would place on human reasoning faculty, which is integral for fundamental human freedom. He extols the role of science as "our chief hope for the future of humankind" and calls for the embracement of free and limitless scientific inquiry even if it leads to the end of humanity (as we know it) or the end of human life on earth. From an evolutionary perspective, he deplores what he refers to as human chauvinism and questions unjustified defense of organic or even embodied forms of life. However, as his account on post-humanism and artificial intelligence makes clear, human freedom and autonomy, and indeed human existence, could equally be undermined by cognitive enhancement just as they could by moral bioenhancement.<sup>3</sup> At the theoretical level, concerns over the potential negative consequences of free scientific research could be addressed, as Harris and others do, by highlighting a distinction between freedom as independence and freedom as license (p 107). At the practical level, however, maintaining this distinction can become quite difficult. Yet, this remains the burden that humans have to recognize and bear as part of the price for full agency and responsibility for both moral and immoral actions. For the near future, moral bioenhancement is likely to remain as a philosophical proposition or theoretical possibility. More work is needed to clarify what it actually means and how it can be implemented beyond mere thought experiments. For example, in a recent issue of *Medicine Health Care and Philosophy*, the editors identify four main suppositions concerning moral bioenhancement, which require further investigation: biological foundations of morality, the extent to which these foundations could be known with sufficient details, possibility of examining this relationship and

---

<sup>3</sup> See, for example, Jurgen Habermas, *The Future of Human Nature* (Cambridge: Polity Press, 2003)

exploring potential side effects, and examining actual real-life contexts within which moral bioenhancement can be implemented.<sup>4</sup>

From an Islamic or religious perspective in general, the notion of enhancement elicits reflection on several theological as well as ethical legal concepts such as: creation and the extent to which it may be reconciled with evolution; religious freedom and determinism; notions of destiny and the scope of human agency in the face of sickness or disability; and finally conceptualizations of religious necessity as well as appeals to notions such as utility and personal interest. Within the Muslim context, normative bioethical discussions, coinciding with and responding to the various applications of biomedical technology, often make a distinction between necessary medical treatment and enhancement procedures. This distinction is reiterated in several resolutions as well as institutional and transnational fatwas in the past few decades.<sup>5</sup> While medical treatment is considered legitimate and could also be mandatory, enhancement procedures are perceived as more problematic. They should be subjected to careful scrutiny and evaluated individually on a case by case basis. One of the main guidelines often emphasized is that enhancement should not amount to, or result in, changing God's creation, which is associated in several scriptural references with the evil works of the devil.

---

<sup>4</sup> Bert Gordijn and Henk ten Have, "Bioenhancement of Morality," *Medicine Health Care and Philosophy* 20 (2017), 289-90.

<sup>5</sup> See, for example, Aḥmad Rajā'ī al-Jindī, *al-Nadawah al-`Ālamiyyah ḥawla al-Wirāthah wa al-Takāthur al-Basharī wa In 'ikāsātuhā, Ru'yat al-Adyān al-Samāwiyyah wa wihat naẓar al-'Almāniyyah*, 2 vols. (Kuwait: Islamic Organization for Medical Sciences, 2008).