

# A Path to the Metaphysics of Moral\*\*\*

Won, Suck Young\*\*\*

**Subjects** Metaphysics, Morals, The process of perfection of human beings

**Key words** Kant, the metaphysics of morals, Hobbes, Dawkins, Frankl, Bostrom

## Abstract

The title of this paper might suggest to some readers Kant's *Grundlegung zur Metaphysik der Sitten*. This is intended. Correspondingly, what I will attempt here is to present a reflective version of Kant's path which led him to the idea and need of the metaphysics of morals in the *Groundwork of The Metaphysics of morals*. The following will be somewhat free and speculative, but entirely loyal to Kant's spirit. In doing this I will use help from Thomas Hobbes, Richard Dawkins, and Victor Frankl. I believe that such a reflection on Kant's path can have a meaning of its own. Furthermore, it can have important implications for the current debate about the artificial process of making human beings entirely homogeneous and perfect with respect to their bodies and

---

\* This work has been supported by a National Research Foundation of Korea Grant funded by the Korean Government (NRF-2015S1A5B5A07043437). Many thanks to three anonymous referees.

\*\* Part of my discussion about Hobbes, Dawkins, and Frankl was presented in Korean on November 23<sup>rd</sup>, 2015 in my talk entitled "Death and a small reflection on life" in Medical Humanities, Good Doctor Institution, Anam Forum.

The first incomplete draft of this paper was presented at the Korean Society of Modern Philosophy on October 21<sup>st</sup>, 2017.

\*\*\* SungKyunkwan University

intelligence, that is, for the so-called transhumanism or posthumanism debate. As far as I can discern, the ultimate reason for Nick Bostrom, one of the most representative figures of the transhumanism movement, for pursuing the process without limits and so in the end producing an all-mighty superintelligence called “posthuman” is its inevitability due to the following foreseeable danger and subsequent fear: that otherwise, it would be possible for some evil minds to produce it ahead of us and then enslave us. I think that, if this is really the situation in which we find ourselves now, we must be said to be really in a very miserable state. If so, then it is worth thinking back and trying to get a better understanding of why we have started pursuing the process of perfection, what in fact we are doing in doing so, and what basic problems this might have. A reflection on Kant’s path to the idea and need of the metaphysics of morals can help us. For this purpose, my reflective version will be attuned to this aim.

The title of this paper might suggest to some readers Kant's *Grundlegung zur Metaphysik der Sitten*. This is intended. Correspondingly, what I will attempt here is to present a reflective version of Kant's path which led him to the idea and need of the metaphysics of morals in the *Groundwork of The Metaphysics of morals* (transl. Mary J. Gregor here and throughout). The following will be somewhat free and speculative, but entirely loyal to Kant's spirit.<sup>1)</sup> In doing this I will use help from Thomas Hobbes, Richard Dawkins, and Victor Frankl.

I believe that such a reflection on Kant's path can have a meaning of its own. Furthermore, it can have important implications for the current debate about the artificial process of making human beings entirely homogeneous and perfect with respect to their bodies and intelligence, that is, for the so-called transhumanism or posthumanism debate.

---

1) Kant showed the great insight that there must be something in what we are thinking of and reflecting on while deciding how to act or appraising actions performed by others, something deep in the human mind or reason which leads us to the idea of the metaphysics of morals and the need for it (Cf. *Groundwork of The Metaphysics of morals*, 107). His practical philosophy in general can thus be understood as his laborious effort to give adequate and systematic expression to it. The expression of the main task of the *Groundwork of The Metaphysics of morals*, namely that of "the search for and establishment of the *supreme principle of morality*" can be taken as its first incorporation (47) which is undoubtedly meant to be metaphysical. However, as the title of this paper also suggests, it is motivated not so much by Kant's main task, but rather by Kant's insight itself along with his thoughts which led him to the idea and need of the metaphysics of morals. So, what I will focus on here is not his main task, but rather his subsidiary explanations and remarks added to it since it is the latter which give us more useful information about it.

As far as I can discern, the ultimate reason for Nick Bostrom, one of the most representative figures of the transhumanism movement, for pursuing the process without limits and so in the end producing an all-mighty superintelligence called “posthuman” is its inevitability due to the following foreseeable danger and subsequent fear: that otherwise, it would be possible for some evil minds to produce it ahead of us and then enslave us (Bostrom, *The Transhumanist FAQ*, 23-24 and 33-34). It is worth noticing that this foreseeable danger and fear is also the ultimate reason for Kass that the transhumanist movement must be stopped or banned.<sup>2)</sup> Apart from this, Bostrom seems to think that not only he himself, but also his transhumanist fellows (including investors and politicians) are of good minds. How and why should we believe this? I will put this serious question aside here despite its enormous importance. Instead I will, for the sake of argument, suppose that they are of good minds. Now, their ultimate reason seems to me to show that they themselves are having some suspicion about what they are doing: They themselves are not so certain of what in fact they are doing in pursuing the process of perfection without limits; they unconsciously have it in mind that the process should originally not be pursued without limits, but rather stopped at a certain point. The problem is, however, that it seems to be impossible for them to determine such a point and justify it. For they have, from the outset, taken their most decisive motive and drive for pursuing the process

---

2) For a discussion about this, see Leon R. Kass, “Ageless Bodies, Happy Souls: Biotechnology and the Pursuit of Perfection”; Nick Bostrom, “In defense of Posthuman Dignity”; Dónald P. O’mathúna, “Human Dignity and the Ethics of Human Enhancement”.

from renouncing the distinction between what is given by nature and what is not, namely the distinction which has been accepted and held to be important in our traditional conception of human beings. It is the same with bad minds. So what remains for Bostrom and his good fellows is to encourage themselves by chanting this motto: We got on the train running on the track, which cannot be stopped; Accordingly, what we can and have to do is, even though we want to get out of it, only to try to run faster and faster in order to arrive at the destination before other evil minds arrive at it first.

I think that, if this is really the situation in which we find ourselves now, we must be said to be really in a very miserable state. If so, then it is worth thinking back and trying to get a better understanding of why we have started pursuing the process of perfection, what in fact we are doing in doing so, and what basic problems it might have. A reflection on Kant's path to the idea and need of the metaphysics of morals can help us. For this purpose, my reflective version will be attuned to this aim. Let's start first with Hobbes!<sup>3)</sup>

---

3) So the order of my presentation in this paper deviates from the order Kant follows in the *Groundwork*. Kant there first reflects on and analyzes a certain type of cases of our everyday use of the concept of duty to show that in "estimating the total worth of our actions" we make use of the concept of an absolute good will which accordingly constitutes the condition of all the rest. This work results in finding and formulating the supreme principle of morality which lies in the concept of good will: namely "*I ought never to act except in such a way that I could also will that my maxim should become a universal law*" (57). Kant originally thinks that what he has done and called philosophical is enough. But he knows that our "innocence" is not able to protect itself and "is easily seduced" by desires and inclinations. This leads us to find a resort in popular moral philosophy not based on the supreme principle of morality. For this reason, he felt the need of the metaphysics of

## I. Hobbes's conception of the nature of human beings

Death or dying someday is one of those things of which we are most certain despite the fact that our knowledge that someday we will die is based on indirect evidence only. Can I know that I am dead when I die? I don't know whether I can. However, this much can at least be said: The verdict on whether or not I am dead is reached not by me, but by others' biological or medical diagnoses. Therefore, if death is a problem of every human being, it can be a problem insofar as we are living while knowing that we will die someday. The problem of death can thus be expressed only from within our life. Too many expressions of the problem of death from within life have been made and are known. So it is needless to go into the details here. Nevertheless, it seems that we have to start from what is best known about it.

---

morals which can help protect itself, and so first felt the need of laying its ground which is pursued in the *Groundwork of the Metaphysics of Morals* (58-65). In pursuing this job Kant's strategy is as follows. As we can see soon in Hobbes, what he as a representative of such popular moral philosophy tries to show is that what we in everyday life hold to be moral imperatives or duties are nothing other than hypothetical imperatives in Kant's terms: imperative skill or precepts of prudence (68-69). What does this show? According to Kant, this shows that they can at least understand that if something is a moral imperative, it must be a categorical imperative whose first formulation is the supreme principle of morality introduced above. This shows then that popular moral philosophy does not deny that we (including them) have the concept of an absolute good will and understand it (61). This is, according to Kant, enough to justify his metaphysical work on morals by analyzing the concept of a categorical imperative and further investigating the possibility of a categorical imperative (68). As said, I will follow a reverse order: first Hobbes as a representative popular moral philosopher and then the concept of an absolute good will by considering how we use it in our everyday moral judgment of the concept.

The primary expression of death from within life is “fear”. Fear of death is usually said to be an instinct. What do we, however, mean to say by “instinct” here? Fear of death is thought to be directly and inseparably related to the desire for self-preservation, and the desire for self-preservation is again held to be directly and inseparably related to life. Here, life is taken as what is just given by nature to an individual and so is assumed to be what cannot be explained any further. Isn’t this the reason why fear of death and the desire for self-preservation, two sides of one and the same coin, are so strong and so intense?<sup>4)</sup> This can be confirmed by the fact that the fear of death along with the desire for self-preservation are what we, human beings, have in common with all other animals. If so, then it is very easy to come to the idea of, so to speak, a “natural right” in the sense in which Thomas Hobbes used it. If life is what is just given by nature to an individual and the individual is thereby determined to strive to preserve his own life and further care about his own welfare<sup>5)</sup>, then isn’t it

---

4) We can imagine how strong and intense they might be, along with Tolstoy in his description of the scene of Anna Karenina’s second attempt of suicide which is made immediately after her first attempt, and which, this second time, at last succeeds in the following way: “But she [Anna Karenina] did not take her eyes off the wheels of the approaching second truck, and at the very moment when the midway point between the wheels drew level, she threw away her red bag, and drawing her head down between her shoulders threw herself forward on her hands under the truck, and with a light movement as if preparing to rise again, immediately dropped on her knees. And at the same moment she was horror-struck at what she was doing. ‘Where am I? What am I doing? Why? **She wished to rise, to throw herself back**, but something huge and relentless struck her on the head and dragged her down.” (*Anna Karenina*, 904, my emphasis)

5) Hobbes defines happiness in the following way: “.....the felicity of this life,

reasonable to also think that the same nature must at the same time have permitted him to do everything necessary for his survival and his own welfare? Hobbes called such permission given by nature to an individual a “natural right”, and he mentioned three principal natural rights as we know (*Leviathan*, Chapter XIII and XIV). According to Hobbes, the ego-centric point of view seems to originate at the same time as the start of life itself and the aim and meaning of life of human beings and other animals are initially and primarily determined by their biological conditions.

It is, however, really an irony or, according to Hobbes, a tragedy that the exercise by each of us of the natural rights, far from ensuring self-preservation and contributing to one’s own welfare, gets us into the state of war in which everyone becomes an enemy of everyone else, and which for that very reason endangers everyone’s own life and welfare the most. It is needless to say that in such a state of war we must, whenever and wherever we may be, be seized by the fear of death. It is, however, also the right place and time for fear of death along with the instinctive desires of self-preservation and one’s own welfare to engage in thinking about a way out of this situation:

“The passions that incline men to peace, are fear of death; desire

---

consisteth not in the repose of a mind satisfied. For there is no such finis ultimus, (utmost aim), nor summum bonum, (greatest good) as is spoken of in the book of the old moral philosophers. Nor can a man any more live, whose desires are at an end, than he, whose senses and imaginations are at a stand. Felicity is a continual progress of the desire, from one object to another; the attaining of the former, being still but the way to the latter.” (*Leviathan*, Part I, Chapter XI, 1)

of such things as are necessary to commodious living; and a hope by their industry to obtain them. And reason suggesteth convenient articles of peace, upon which men may be drawn to agreement.”  
(*Leviathan*, Part I, Chapter XIII, 14)

Reason first proposes step by step a long list of strategies which Hobbes calls “natural laws”, and which we have to practice if we are to live in peace with one another. This is how Hobbes replaces the traditional metaphysical moral philosophy by the science of natural laws understood in the above way (*Leviathan*, Part I, Chapter XV, 40 and 41). Apart from this, after finding human beings’ inability to spontaneously keep promises and contracts along with the natural laws, reason advises each of us to construct a Commonwealth consisting of various subordinate systems which are given full powers necessary for making us keep contracts and promises along with the natural laws. So, all social, economic, political, cultural systems, and the like can be said to be there, in principle, for the sake of everyone’s self-preservation and self-welfare. Accordingly, there is also a sense in which the aim and meaning of human life in such a Commonwealth can in general be said to be determined by their biological, social, economic, political, cultural conditions, and the like. On the other hand, it is important to stress that the state of nature along with the nature of human beings is not removed by introducing the artificial state, namely by constructing a Commonwealth. This is no more than a security net system laid upon the state of nature. Accordingly, there is always the danger that it partially doesn’t work or will even be wholly removed. In the latter case the state of nature comes back.

As we know, Hobbes declares that animals are nothing other than *automata* (*Leviathan*, Introduction, 1). Reason is nothing other than an instrumental calculating system whose only mission is to find out the best ways of satisfying our desires in general, and free will means nothing other than the overall desire in which the competition of various different desires results (*Leviathan*, Part I, Chapter 6, 53). Accordingly, there is in principle no qualitative difference between human beings on the one hand and other animals on the other hand in their calculating systems and choosing mechanisms between various different desires. There is a difference in degrees only.

## II. Some reflections on possible motives for Hobbes's conception of the nature of human beings

What might have led Hobbes to such an extremely biological and ego-centric conception of the nature of human beings? There might be many reasons, but let me speculate just about one reason on the basis of a story about Hobbes.

As a friend saw Hobbes give money to a beggar, he noted that he had thereby done something against his own theory. Hobbes replied to his friend that he had done it not for the beggar, but rather for himself: It makes him happy to see the pleasure that the beggar obtained from the gift. Peter Singer interprets this story in this way:<sup>6)</sup>

---

6) This is a completely different and revised version of my discussion of Singer in "Hobbes' Moral Philosophy and Materialism"(written in Korean; 256-259).

“Both Hobbes and Alexander’s views of human motivation are, in the end, entirely compatible with the existence of all the altruism (in the ordinary sense of the term) that anyone would ever want to argue for. After all, who cares what the ‘real significance’ of this kind of altruism might be, when we are interested in understanding how people can be motivated to act ethically.” (*How are we to live?*, 105)

According to Singer, the story above shows that Hobbes makes a distinction between the ‘real significance’ of altruism or the concept of an absolute good will in Kant’s terms and the ordinary sense of altruism.<sup>7)</sup> This is true. And it is because of the application of the concept of an absolute good will to his judgement that Hobbes can say that his act is a selfish one whereas it is because his friend makes use of the ordinary sense of altruism that he judges Hobbes’s act to be altruistic. This is also true. Now, the question is how Singer gets from this to the conclusion that Hobbes’s theory is compatible with the existence of altruism in the ordinary sense. I do not believe that Singer presupposes here a certain kind of contextualism to the extent that whether an act is altruistic or not depends on the contexts in which it is judged. If I am right, then, the only ground on which Singer comes to that conclusion is the following claim in the quotation: “After all, who cares what the ‘real significance’ of this kind of altruism might be, when we are interested in understanding how people can be

---

7) I think that it is safe to take the ‘real significance’ of altruism in Singer’s terms as a version of the concept of “an absolute good will” in Kant’s terms, and therefore I will from now on use, for the sake of convenience, the concept of an absolute good will instead of the ‘real significance’ of altruism.

motivated to act ethically.” I will not here go into the question whether this claim can justify Singer’s inference. What I am interested in is only whether Singer’s claim is true.

Let me first stress that in order for Hobbes’s friend to judge Hobbes’s act in the way he did, he does not need to be a philosopher or presuppose any philosophical theory of altruism and selfishness. Singer also thinks so. This is why Singer takes him as a representative of making use of the ordinary sense of altruism in judging whether an act is altruistic or not. How about Hobbes’s judgment of his own act? The first sentence of the quotation shows that Singer takes Hobbes in the story as a philosopher, as a representative of making use of the concept of an absolute good will. Let me stress that Singer does not say that ordinary people do not make the distinction in question, but only that they are not interested in the concept of an absolute good will to the effect that they do not make use of it in judging whether or not an act is altruistic. Whether or not the latter is true is what matters here. I think that Singer is wrong, and that things are very much more complicated than Singer thinks.<sup>8)</sup>

In everyday life there are many cases in which we meet ordinary people who show the same kind of reaction as Hobbes’ after doing acts which we might without further ado hold to be altruistic: Hobbes is not an exception at all. Then, there is also no need for us ordinary people to be philosophers or presuppose any philosophical theory of

---

8) According to Singer, it is only some philosophers, not everyday people, who are interested in the ‘real significance’ of altruism and thereby driven to exclusively concentrate on the internal problems with the ego (Cf. chapters 8-10). And the end result is that “many of them must be in *daily* psychoanalysis” (206).

altruism and selfishness in order to show such a reaction. I think that this alone might show that Singer is wrong in making the claim above. The interesting and important question is why they show such a reaction when they do. To answer this question, let me develop the story a little further by asking the question: How does Hobbes's friend react to Hobbes when both of them are now taken not as philosophers, but as ordinary people? Hobbes's friend would not say anything like what Singer says in the quotation since what Singer says there is what Singer says as a philosopher, not as an ordinary person. Instead, Hobbes's friend would think like us in the following way. We can understand why Hobbes shows such a reaction even though he did an act which is worthy of being taken as altruistic in the ordinary sense of altruism. It is that Hobbes used the concept of an absolute good will as a criterion for assessing his act. We know that in this case there arise inner conflicts. And we know further that, once those conflicts arise, other people's judgements according to the ordinary sense of altruism cannot count at all. Furthermore, suppose that there have not been actions at all which fulfilled the concept of an absolute good will. Even though this might console us, it cannot help us in Hobbes's present situation since whether there have been such actions or not cannot matter at all when we decide to make use of the concept of an absolute good will in judging our actions, as Kant claims (*Groundwork*, 62). This means at the same time that we also know that it is nearly impossible to solve such inner conflicts in a positive way since the concept of will in Kant's terms is an ideal which it is nearly impossible to fulfill. Then, how can we make ourselves free from the struggle with those inner conflicts and find some peace of mind? There

are many different ways, and Hobbes's reaction is one of them: admitting that he has not done the act in conformity with an absolute good will.

How would we or Hobbes's friend then characterize Hobbes's reaction? We would say that he is "humble", in its genuine sense in contrast to a hypocritical one. I think this shows what we mean when we say of someone that he is humble in the respect in question. It is that he tries to judge his acts according to the concept of an absolute will and so is thereby prepared to suffer the inner conflicts due to it and to struggle with them, and, without flattering "by falsely attributing to" himself "a nobler motive" (*Groundwork*, 61), confesses to himself that he is vanquished in the struggle in which it is nearly impossible to be a winner.<sup>9)</sup>

The considerations so far are, I think, enough to show that we know and recognize that there is a certain kind of *worth* in it. This is why we usually show respect to people who we hold to be humble. It is, to speak with Kant, a moral worth. At the same time, they also show that Singer's claim above is not true, and Kant is entirely right in thinking that we have the concept of an absolute good will we think of and reflect on while deciding how to act or appraising actions performed by others, and that we have inner conflicts due to our having it. So far, too optimistic, however. Now is the time for pessimistic things to come

---

9) Another interesting expression of this would be what Captain Louis Renault says in the movie "Casablanca" to Rick Blaine who, whenever he does an altruistic act, always tells its beneficiaries who want to say "many thanks" to him that he has done it only for himself, not for them. What he says is that Rick seems to try to make others believe what he himself does not believe.

into the picture.

As said above, Hobbes's reaction is just one way of solving the inner conflicts due to having the concept of an absolute good will. There are other possible ways. One of them is already suggested by Singer's claim: pressing our interest in the concept of an absolute good will and trying not to make use of it. Another one which can be called "pessimistic" would be to negate its possibility entirely and so to take it as an illusion. For example, Edward Wilson says of Mother Teresa's deeds that they cannot be taken to be altruistic but should be seen as selfish since she must be motivated by her expectation of reward in heaven.<sup>10</sup> I think that some of us would assent to it but some of us would not. This shows that it is also not necessary for someone to be a philosopher or presuppose any philosophical theory in order to assent to Wilson or not. It is possible to make such a judgment as Wilson's from our own pessimistic experiences with ourselves and others only. We do not negate thereby that we have the concept of an absolute good will. The difference arises only from the different attitudes to it: as illusion or ideal.

If so, Hobbes's theory can be said to be too radical since he seems to try to eradicate the concept of an absolute good will itself by reducing the traditional metaphysical moral terms and their worth to the science of natural laws and a qualitatively different sorts of worth. I am very skeptical whether this is really possible, and I think that Hobbes himself is not convinced of its possibility. Accordingly, I think that Kant's diagnosis of some philosophers who go in that direction is

---

10) This example is taken also from Singer (86).

correct:

“Hence there have at all times been philosophers who have absolutely denied the reality of this disposition [to act from pure duty] in human actions and ascribed everything to more or less refined self-love. They did not, on account of this, call into doubt the concept of morality but rather spoke with deep regret of the frailty and impurity of human nature, which is indeed noble enough to take as its precepts an idea so worthy of respect but at the same time is too weak to follow it, and uses reason, which should serve it for giving law, only to look after the interests of the inclinations, whether singly or, at most, in their greatest compatibility with one another” (*Groundwork*, 61; cf. 59-60 and 78)

Apart from it, if we consider Hobbes’s conception of the nature of human beings and of the artificial state with written laws which are expected to incorporate natural laws, then it is, I think, easy to come to the transhumanist aspiration. Let me illustrate how it happens in the next section by considering 5 cases.

### III. The path from Hobbes’s conception to the idea of transhumanism or posthumanism

**Case 1:** Most of us are very much fascinated by an athlete when he sets a new world record in the Olympics and appreciate his deed very highly. Why? For instance, why are we so fascinated by Usain Bolt as he set a new world record for the 100-meter race in the Beijing Olympics 2008? There might, of course, be various different reasons.

Let me here just focus on the following one. It can be said that there is a traditional conception of human beings we consciously or unconsciously share. It naturally consists of various different sorts of aspects: physical, biological, social, political, cultural, moral, and so on. This is, of course, not a water-tight claim, nor is it a rigid all-or-nothing definition. Take our example once again. When human beings' running ability first appeared on the list of the aspects of the traditional conception, it might have been described just in such a way that human beings cannot run faster than a normal horse. It is true that human beings cannot run faster than a normal horse, and this question can thus be said to be closed in that respect. This limit is, however, too vague. So one reason that we are so fascinated by an athlete who sets a new world-record and appraise him highly might be because we think of him, a representative of human beings, as showing how close human beings can come to the limit: Even though we recognize that there are differences in our natural talents or capacities, we identify ourselves with him as members of one and the same species of nature.

However, things might appear different if we take Hobbes's ego-centric theory. For example, I would think that I would be very happy and fascinated if it were me who set the world record. But it is not me, but the athlete who did in fact set the world record. So there is no reason for me to be fascinated. Naturally, this thought can be applied to all sorts of things given by nature to us.

**Case 2:** Suppose a very excellent secret agent, 007 for example, unfortunately lost his two legs in performing his duty. Since he is not replaceable at all, the chief of the Intelligence Agency decides to let

him receive artificial legs. One of its results is that with his new legs he can now run a 100 meter distance in about 3 seconds. Suppose he intends to participate in the 100-meter race in the Olympics. If he were admitted, there is no doubt that he would win the race with a new world record. Would or should he be admitted? I am not sure whether he would or should be. However, I can bet everything that the IOC would not allow it. In making a public announcement of their decision the IOC would give an explanation which is based on the traditional conception: “Even though we have looked into the traditional conception of human beings thoroughly, we could not find anywhere the expression “born with artificial legs.””

Apart from this, how would we think about the very enhancement made to the agent outside the context of the Olympics? Answers to this question differ depending on contexts. Some of us would say “O.K.” while some of us would not. Suppose that the agent received just normal prosthetic legs. Then, there would be nothing to discuss. If so, those who say “O.K.” have given more weight to the fact that the agent lost his legs and that he must receive some replacement whatever it might be. On the contrary, those who say “No” have given more weight to the fact that the enhancement is made beyond the limit conditioned by the biological nature of human beings: It thereby breaks the rules which our society has made for ensuring fair competition in the society as the rules of the Olympics for the same purpose.

But how would we think about it if we were the agent? I think that we would say that it is phantastic. Remember that we are still discussing a case in which the agent by accident lost the legs he originally had.

Things can change, however, in all possible ways when competitions become fierce and, correspondingly, the difference between the winner and loser in their rewards becomes bigger. Sure, at the beginning, present rules and laws would resist the change. Then, even the rules and laws would be changed. Let's look at further cases which go in that direction.

**Case 3:** Suppose I get plastic surgery since I don't like my face. Great and hot debates would arise amongst my friends, family, and so on. In the case of eastern societies, there is a contrary argument which is typically offered in such debates. Let me paraphrase it in the same way as in the cases above: "We have looked into the traditional conception thoroughly and found there the item that we should keep our body as it was originally given from our parents."<sup>11)</sup>

We must say that this argument has already lost its power. Today, plastic surgery is not anymore considered as a taboo.

**Case 4:** I remember that, when the compact calculator was first invented, it very quickly spread among people, but students were not allowed to use it, for instance, in math exams. Why not? As in case 2, the reason which might intuitively be given is this: "We have looked into the traditional conception thoroughly and we could not find anywhere an expression like "born with a compact calculator", but only "born with a brain".

We now know how short a time it took for it was allowed over the

---

11) This is known as a dictum by Confucius.

world to use compact calculators.

**Case 5:** Suppose that the compact calculator had been prohibited very strictly until now. And suppose further that, since the job which I wanted required a very high calculating ability, I decided to let a microchip be installed in my brain which makes my calculating ability immensely better and even my memory as well. But, it would not help me at all since this would be detected by the electronic detector at the gate of the company where I wanted the job. But there will soon be a way out of it thanks to recent breakthroughs in genetic engineering and advanced technologies of gene manipulation, it would be soon possible, to enhance my calculating ability by manipulating brain cells responsible for it.

The last three cases illustrate conflicts between roughly two kinds of people today, namely between those who think that the traditional conception of human beings must be defended and preserved on the one hand, and those who think that it can be changed on the other hand.<sup>12)</sup> The position of the defending party is primarily based on the

---

12) One referee made some important and helpful comments for which I want to say special thanks. I will consider them when I will have a chance to write a subsequent paper. For now, let me just reply to his comment that I should make reference to the discussion between Sandel and Buchanan in Sandel's *The Case against Perfection* when I engage in my speculative reflection on these matters and cases. The reason why I did not deal with this is simply that nowadays everybody is inclined to deal with such cases; I only intended to have them **illustrate** how easily we ordinary people might nowadays be tempted to ignore and dismiss the distinction between what is given by nature and what is not given by nature. Correspondingly, Confucius also is just an example for the many different reasons that one could have for arguing against the process.

idea that what is given by nature must be kept and preserved as it is given. This idea presupposes that there is a relatively clear distinction between what is given by nature and what is not given by nature, and that we can know what is what. Things, however, seem to have become entirely different today. We know that what makes evolution possible is ultimately the occurrence of “errors or mistakes in the copying process of genes,”<sup>13)</sup> and it seems to be in vain to try to give substance to the distinction between natural errors and artificial errors after all. Furthermore, due to recent breakthroughs in molecular nanotechnology and super-computing, it seems that it’s getting harder and harder to sustain such a distinction and it might soon become an idle one. If so, all the problems or conflicts of the kind we saw in the above cases will also disappear. Consequently, we can easily be encouraged, led, and even driven to the idea of making ourselves into what we want. All this is incorporated in the basic idea of transhumanism:

“[Transhumanism] is the intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to **eliminate aging** and to greatly enhance human intellectual, physical, and psychological capacities....

Just as we use rational means [like wheels] to improve the human condition and the external world, we can also use such means **to improve ourselves, the human organism**. In doing so, we are not limited to traditional humanistic methods, such as education and

---

13) See *The Selfish Gene*, Chapter II.

cultural development. We can also use technological means [molecular nanotechnology and super computer] that will eventually enable us **to move beyond what some would think of as “human”** [making ourselves posthumans, namely all mighty superintelligences]. It is **not our own shape or the details of our current human biology that define what is valuable about us**, but rather our aspirations and ideals, our experiences, and the kinds of lives we lead.” (*The Transhumanist FAQ*, 4; my own emphasis)

What in fact are we doing in pursuing this process and what are its problems if any? Michael Sandel’s answers to these questions are the following:

“According to this view, enhancement threatens our humanity by eroding human agency. Its ultimate expression is a wholly mechanistic understanding of human action at odds with human freedom and moral responsibility.

Though there is much to be said for this account, I do not think that the main problem with enhancement and genetic engineering is that they undermine effort and erode human agency. The deeper danger is that they represent a kind of hyper agency, a Promethean aspiration to remake nature, including human nature, to serve our purposes and satisfy our desires. The problem is not the drift to mechanism but the drive to mastery. And what the drive to mastery misses, and may even destroy, is an appreciation of the gifted character of human powers and achievements.” (The case against perfection, 26-27; my own emphasis)

Even though I do in general agree with this line of criticism, I think that we can take one step more and go deeper, as follows. As we can see in the quotation from Bostrom, the basic idea of transhumanism is

to improve all the biological human conditions which set limits to our abilities. All this, but, for what purpose? The answer is to maximally fulfill the aim and meaning of life which are originally determined by our biological conditions (as representatives are given, for example, eliminating aging and eternal life). On the other hand, if we try to overcome limits set by our biological conditions for achieving and realizing the aim and meaning of life which are determined by our biological conditions, then this means that we thereby subject ourselves to the principles which are also originally set by our biological conditions for achieving and realizing the aim and meaning of life. If so, I think that there is some paradox or irony here. For, according to my feeling, it is hard to think of pursuing and achieving a certain aim according to a certain principle as a process of “overcoming biologically conditioned limits” if both the aim and the principle are conditioned by our biological nature.

Accordingly, in the next section I will try to give an illustration of what in fact we are doing in pursuing the process of perfection. This is nothing other than what I have explained above by appealing to Dawkins. I will make some reflective remarks on it by relying on Kant.

#### IV. Dawkins's conception of human beings in *The Selfish Gene*

Let me first briefly explain what Dawkins is trying to do and how. There are 4 theses which he is trying to argue for:<sup>14)</sup>

- D1) The fundamental units of natural selection are genes.
- D2) Genes are selfish.
- D3) Living beings are survival machines which genes made in order to survive and pass their copies on to next generations.
- D4) Living beings are selfish.

The way in which Dawkins argues for these theses is in examining whether, given (D1) and (D2), (D3) and (D4) are to be accepted. Let us confine our discussion to human beings and other animals as in the case of Hobbes. The strategy Dawkins chose consists in seeing whether alleged representative types of counterexamples to (D3) and (D4) can on the basis of (D1) and (D2) be explained away.

Let's first ask the question of how we think about (D3) and (D4). It is safe to say that most of us would feel very uneasy about both (D3) and (D4). Why? Even provided that (D1) and (D2) are true, it does not seem to directly follow from them that we are just machines made by genes. I am, so to speak, an individual who is conscious of himself

---

14) See *The Selfish Gene*, Chapter 1.

as an individual as a whole and as having a free will to do what I as an individual want to do. And this is also what is taken to explain how it is possible to try to live as a “being dominated by culture” (“culture” in its broadest sense of everything that is not given by nature), “by influences learned and handed down,” not just by following the rule of the ruthless selfishness of genes (*The Selfish Gene*, 3). If so, we have to acknowledge that both “free will” and “culture” are used by us in a somewhat “snobbish” sense (189): they are what distinguishes us from other animals. Let’s then return to our original problem.

If it is true that we live according to our free will and under the guidance of our culture, then it has to be said that it is nothing other than our free will and culture which lead us to the process of the perfection of human beings. How could it be otherwise? To make the point clear, let me first dramatise the final result of the process of perfection. Let’s see one of many dramatic descriptions of it given in many science fictions of different types. Remember any fiction in which colonizing another planet happens. To directly send human beings in order to colonize another planet is too risky, however. So instead we produce robots which are perfectly and highly specialized in certain activities necessary for, so to speak, survival and performing their missions. They naturally look all alike in every respect. I think that this is precisely the same as what human beings would become when the process of perfection would be completed. This means that we are on the way to make ourselves most efficient and perfect survival machines for genes under the guidance and regulation of what we claim makes us independent of our creators, namely genes, and so

different from other animals.<sup>15)</sup> Isn't it thus really an irony that it is ultimately genes which will be the winners? How would this be possible? Let me reconstruct a possible explanation relying on Dawkins.

Dawkins says:

(DQ1) “Genes are the primary policy-makers; brains are the executives. But as brains became more highly developed, they took over more and more of the actual policy decisions, using tricks like learning and simulation in doing so. The logical conclusion to this trend, not yet reached in any species, would be for the genes to give the survival machine a single overall policy instruction: do whatever you think best to keep us alive.”(60)

I do not believe that Dawkins is sincere in saying “not yet reached in any species,” and I am convinced that I am right as it will be shown soon. For the moment, let me just interpret the quotation freely, namely in the following way. Genes produced human beings in which a program-making-program is built, namely one master program which is totally open with respect to its potential power, and then the genes gave them just a single overall policy instruction described in the quotation. So, human beings feel thoroughly free to live as they want, but with the only strong exception for the instinct to survive. Now, what human beings do with their free will is to produce culture in its broadest sense including all artificial systems along with all possible

---

15) A similar point at the level of the relation between AI and human beings is made and stressed by Prof. Lee, Jong-Kwan. “Future of Work and its temporality in the Age of 4<sup>th</sup>IndustrialRevolution”,3-4.

ideologies, religions and the like, and this is taken to be a powerful and “snobbish” expression of their being free. Their feeling of being free can be thought even to go beyond the limit posed by the single overall policy instruction described in the quotation since, to take just one extreme example, they see people who commit suicide. But as we saw, the result which it finally leads us to does not seem to be worthwhile taken in that way in reality: It is nothing other than us, ourselves, who in the place of genes make ourselves most efficient and perfect survival machines for genes. Then it must be admitted that genes seem anyway so smart.<sup>16)</sup>

How could this happen? According to Dawkins, the right explanation must be sought in the essential feature or character of culture in general which human beings created: cultures which human beings created with the high pride that they are what characterizes human beings as unique, having inherited the essential character of biological genes, namely individual and group selfishness (Chapter 11 and 220-233; especially 6-9).<sup>17)</sup>

I think these considerations are enough to make my point illustrate that the process of perfection is no more than a process loyal to the aim and the principle both of which are originally set by our biological conditions. It does not help to say that in the future our biological

---

16) I think, even though it might sound curious, that if this conspiracy of genes were detected by us, we would find it problematic and even try to resist it. However, how can we then detect it if we can at all? I think that we can detect it only if we can find it problematic

17) This is a conclusion which I arrived at together with Prof. Ahn, Sung-Doo in our study of *The Selfish Gene*. See “A Buddhist Interpretation of the Evolutionism” (written in Korean), 429-430.

genes themselves will also be replaced by new types of genes which are made by ourselves through nanotechnology. What matters is just the aim and principle.

On the other hand, Dawkins does not think of such a process as a genuine overcoming of our biological conditions. Even though Dawkins declares “I am not advocating a morality based on evolution. I am saying how things have evolved. I am not saying how we humans morally ought to behave,” he, nevertheless, thereby suggests that there might be another kind of morality which is not based on evolution and different from Thomas Hobbes’s version of morality. This point becomes clear when we see Dawkins’s warning to some possible readers who expect “a morality based on evolution” from his book:

(DQ2) “Be warned that if you wish, as I do, to build a society in which individuals cooperate generously and unselfishly towards a common good, you can expect little help from biological nature. Let us try to teach generosity and altruism, because we are born selfish. Let us understand what our own selfish genes are up to, because we may then at least have the chance to upset their designs, something that no other species has ever aspired to.”(3)

What Dawkins means here by “generosity and altruism” is not simply something strategic which is useful to some degree for our own survival and welfare in some social relations. He has more in mind as he says:

“A question that sociologists and psychologists sometimes ask is why blood donors (in countries, such as Britain, where they are not paid) give blood. I find it hard to believe that the answer lies

in reciprocity or disguised selfishness in any simple sense. It is not as though regular blood donors receive preferential treatment when they come to need a transfusion. They are not even issued with little gold stars to wear. Maybe I am naïve, but I find myself tempted to see it as **a genuine case of pure, disinterested altruism.**" (230; my own emphasis)

There is no doubt that Kant too would not think that the process of perfection can be taken as a process of overcoming our biological conditions. It is rather a case of surrendering to them. If we really want to overcome our biological conditions, we have to overcome the aim and meaning of life themselves which are determined by them. The way of it lies in reason "the vocation" of which "must be to produce a will that is good, not perhaps as a means to other purposes, but *good in itself*, for which reason was absolutely necessary" (*Groundwork*, 52; cf. 50-51). And we can be sure that this is true. Otherwise, how would it be possible for us to find the process of perfection problematic? If our reason is really determined by Mother Nature only to serve the aim and meaning of life which are determined by our biological conditions, then it is not possible to have any principled suspicion about engaging in the process of perfection. But we do have such suspicions. Correspondingly, in the *Groundwork* Kant explains repeatedly how and what we would think about ourselves when we see that our reason exclusively serve that aim and meaning of life. For example he says:

"they [basic moral principles reason dictates] expect nothing from the inclination of human beings but everything from the supremacy

of the law and the respect owed it or, **failing this, condemn the human being to contempt for himself and inner abhorrence**” (77; cf. 51; my own emphasis)

I hope that all my considerations so far show what we are in fact doing in pursuing the process of perfection and what problem in principle it shows if we find it problematic in principle. I will close this paper by giving an example of how such a process happens from Viktor Frankl’s *Man’s Search for Meaning*.

## V. Closing part

In a concentration camp like Auschwitz in the 2nd World War what counted for most people there is only survival, and so their minds must have been full of thoughts like the following:

F1) “I remember a personal experience. Almost in tears from pain (I had terrible sores on my feet from wearing torn shoes), I limped a few kilometers with our long column of men from the camp to our work site. Very cold, bitter winds struck us. I kept thinking of the endless little problems of our miserable life. What would there be to eat tonight? If a piece of sausage came as extra ration, should I exchange it for a piece of bread? Should I trade my last cigarette, which was left from a bonus I received a fortnight ago, for a bowl of soup? How could I get a piece of wire to replace the fragment which served as one of my shoelaces? Would I get to our work site in time to join my usual working party or would I have to join another, which might have a brutal foreman? What could I do to get on good terms with the

Capo, who could help me to obtain work in camp instead of undertaking this horribly long daily march? (73)

All thoughts of this kind are not at all trivial in the camp but rather necessary and concern things that are most important and so reasonable. How to act is decided according to certain strategies for survival and welfare. But, nevertheless, I am sure that all of us know that there is every reason to think that there are always exceptions of people who do not think nor behave in that way and so show their dignity even in such an extremely miserable situation. What might be the thoughts like which they would come up with in order to come to decide to do so? It would be like what Frankl describes:

F2) “The way in which a man accepts his fate and all the suffering it entails, the way he takes up his cross, give him ample opportunity-even under the most difficult circumstances-to add a deeper meaning to his life. It may remain brave, dignified and unselfish. Or in bitter fight for self-preservation he may forget his human dignity and become no more than an animal.” (67)

This is a conflict between following the law of the pure will in Kant’s sense and following strategies based on the law of ruthless selfishness, whether at a level of the individual or the group in the camp. Now the last impetus to coming to decide to act according to the law of the pure will be something like what Frankl adds immediately after (F1):

“I became disgusted with the state of affairs which compelled me, daily and hourly, to think of only such trivial things.”(73)

This shows that the thought of those important thoughts and things described in (F1) as necessary and important might suddenly change into the thought of them as “trivial thoughts and things”. How could this happen? The reason Frankl gives is that of “becoming disgusted.” Sure, there might be various different reasons for becoming disgusted, and, accordingly, different corresponding results might come out. But it is reasonable to think that if the result of the change of the thought is to be like what is “to remain brave, dignified and unselfish” as described in (F2), then the reason for becoming disgusted must be to have forgotten “his human dignity” and to have “become no more than an animal” in the ruthless struggle for existence described in (F2).

## References

- Bostrom, N. Version 2.1 2003, *The Transhumanist FAQ*,  
www.nickbostrom.com
- Dawkins, R. 2006, *The Selfish Gene*, 30<sup>th</sup> Anniversary Edition, Oxford University Press.
- Frankl, V.E. 1992, *Man's Search For Meaning*, Bacon Press, Boston, transl. by Ilse Lasch.
- Hobbes, T. 1996, *Leviathan*, ed. by Gaskin, Oxford World's Classics.
- Kant, I. 1952, *Grundlegung zur Metaphysik der Sitten*, ed. by Vorlaender, K. Felix Meiner in Hamburg.
- \_\_\_\_\_ 1996, *Groundwork of The metaphysics of morals*, transl. by Gregor, M.J. in *Practical Philosophy*, ed. by Gregor, M.J., The Cambridge Edition of The Works of Immanuel Kant.
- Sandel, M.J. 2007, *The Case against Perfection*, The Belknap Press of Harvard University Press.
- Singer, P. 1995, *How are we to live?*, Prometheus Books, Amherst, New York
- Tolstoy, L. 2012, *Anna Karenina*, transl. by Louisa and Asylmer Maude, Vintage Books, A Division of Random House, New York.
- 안성두 2010, 「진화론과 불교사상의 접점」, 『한국불교학』 56집.
- Lee, Jong Kwan. 2017, “Future of Work and its temporality in the Age of 4<sup>th</sup> Industrial Revolution”, presented on November 7th 2017 in the second International Conference entitled “Philosophy between Past and Future” at Sungkyunkwan University.
- Kass, L.R. 2003, “Ageless Bodies, Happy Souls: Biotechnology and the Pursuit of Perfection”, www.TheNewAtlantis.com

Bostrom, N. 2005, “In defense of Posthuman Dignity”, *Bioethics*, Vol.19, No. 3, pp.202-214, [www.nickbostrom.com](http://www.nickbostrom.com)

Dónald P.O'mathúna 2013,“Human Dignity and the Ethics of Human Enhancement”-<https://bioethicsireland.files.wordpress.com/2009/09/omathuna-2013-enhancement-dignity.pdf>

원석영 2009, 「홉스의 도덕철학과 유물론」, 철학적 분석 20호

## 도덕 형이상학에 이르는 길

원석영 (성균관대학교)

본 논문의 제목은 독자들에게 칸트의 『도덕 형이상학 정초』를 연상시킬 수 있다. 이는 의도된 것이다. 이에 상응하게, 나는 그곳에서 칸트가 도덕 형이상학에 대한 아이디어와 그것의 필요성에 이르게 된 과정에 대한 반성적인 버전을, 어느 정도 자유롭게 사변적이기는 하지만 칸트의 정신에 충실한 버전을 제시하고자 한다. 이를 위해 나는 흄스와 도킨스 그리고 프랭클의 도움을 빌 것이다. 나는 그러한 반성이 그 자체로서 의미를 가질 수 있다고 믿는다. 더 나아가, 나는 그것이 인간의 지적 신체적 능력을 인공적으로 완전하게 만드는, 따라서 결국 모든 인간이 완벽하게 동형 동질화 되어갈 수 있는 과정에 대한 논쟁, 이른 바 트랜스 또는 포스트 휴머니즘 논쟁 일반과 관련해서 중요한 의미를 함축하고 있다고 생각한다. 내가 파악한 바로는, 트랜스 휴머니즘 운동의 대표적인 인물 가운데 하나인 닉 보스트롬이 완전화 과정을 끝까지 추구해야 하는 그래서 마침내 전능한 슈퍼 인텔리전스를 만들어야 하는 궁극적인 이유로 제시하는 것은 예측할 수 있는 위험과 이에 따른 두려움에서 기인된 불가피성이다. 즉 그러지 않을 경우, 사악한 마음을 지닌 자들이 우리보다 먼저 슈퍼 인텔리전스를 만들어 우리를 노예와 하게 될 것이라는 위험과 이에 따른 두려움 말이다, 나는 현재 우리가 처해있는 상황이 실제로 이렇다면, 우리는 참으로 참담한 상황에 놓여 있다고 생각한다. 그렇다면 우리가 우리 자신을 되돌아보는 것은 의미 있는 일이라고 생각한다. 왜 우리가 완전화 과정을 추구하게 된 것인지, 그것을 추구하면서 실제로 우리가 무엇을 하고 있는 것인지, 그리고 그것이 어떤 근본적인 문제들을 가지고 있는지에 대해 보다 나은 이해를 얻도록 노력해야 할 적기라고 생각한다. 칸트가 도덕 형이상학에 대한 아이디어와 그 필요성에

철학탐구 제51집

이르게 된 과정에 대한 반성은 우리에게 도움을 줄 수 있다. 이 때문에, 나의 반성적 버전은 그러한 목적에 맞추어져 있다.

주제분류 : 형이상학, 도덕, 인간의 완전화 과정

주요어 : 칸트, 도덕 형이상학, 흄스, 도킨스, 프랭클, 보스트롬

원석영 E-mail: [sucky@skku.edu](mailto:sucky@skku.edu)

투 고 일	2018년 07월 30일
심 사 일	2018년 07월 30일
게재확정	2018년 08월 19일