Kant and the Categorical Imperative

William O’Meara
Professor of Philosophy
James Madison University
USA
An Introduction to
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research
This paper should be cited as follows:

Kant and the Categorical Imperative

William O’Meara
Professor of Philosophy
James Madison University
USA

Abstract

Kant’s application of his first two formulations of the Categorical Imperative, the Formula of Universal Law (FUL) and the Formula of Humanity as an End-in-Itself (FH), to his famous set of four moral problems reveals the weaknesses of his defense of these formulae.

Korsgaard correctly analyzes Kant’s application of the FUL to the first two problems. However, Korsgaard fails to consider, for example, that an agent could generalize to a unique class which has only one possible member in the following way without any contradiction in concept, “All who are uniquely Joe Jones with my unique genome and phenome may commit suicide or make a false promise. Of course, that maxim is immoral, but not because it violates any law of logic. What it does violate is universal respect for humanity as an end, and so the FUL must rely upon the FH. Kant’s defense of the FH will be evaluated as inadequate also.

Korsgaard correctly analyzes Kant’s application of the FUL to the last two problems. However, Kant’s attempt fails because there is again no logical contradiction in one’s willing if one generalizes to a unique class with only one member who resolves never to accept help from any other person. So again, Kant’s argument here must be that the FUL needs the FH. But Kant himself violates his rule that humanity must be conceived as an independently existing end whose value ought not be harmed. Kant’s FH, for its best understanding, is dependent upon belief in freedom of the will, which cannot be rationally proven, as Kant himself admits and Korsgaard emphasizes, and hence FH must be conceived as fundamentally dependent upon one’s own deep existential choice of human freedom and of valuing one’s humanity as the supreme value, both not to be harmed and also to be advanced.

Keywords: Kant, Categorical, Imperative

Contact Information of Corresponding author:
Kant’s famous four moral problems offer us an opportunity to evaluate his defense and his application of his first two formulations of the categorical imperative, the Formula of Universal Law (FUL) and the Formula of Humanity as an End in Itself (FH).

**Moral Problems 1 and 2**

The first moral problem is a proposed act of suicide in order to escape an unhappy life, and the second moral problem is a proposed false promise in order to obtain money. Korsgaard correctly sees Kant as attempting to derive perfect duties of avoiding both suicide and the making of a false promise. For these two proposed actions cannot even be conceived as universal laws of nature without a contradiction.

In evaluation of this argument, we must grant that Kant is right if one attempts to generalize one’s proposed action to all rational agents, then one is involved in a contradiction. Korsgaard focuses well on the contradiction. A person is attempting to universalize the following maxim, for example, ‘If I desire the end of obtaining money, then I may make a false promise,’ into ‘If anyone desires the obtaining of money, then anyone may make a false promise.’ But such a universalization would render every person’s false promise useless as a means to obtain one’s desired end, but one who desires an end necessarily wills the means. Korsgaard concludes that . . . ‘the lying promiser who attempts to will the universalization of his maxim wills the denial of the analytical principle on which he himself proposes to act, and the denial of an analytical principle is a contradiction’ (Korsgaard 15).

However, in further evaluation, we may argue that no conceptual contradiction is involved if one generalizes one’s maxim to a restricted group as to how that group treats others outside the group. For example, if I belong to group ‘X’ and say that all who are in group ‘X’ may make false promises to other groups, then I am not contradicting myself, especially if all those, including me, in group ‘X’ have promised never to make a false promise to any other member of the group. People who belong to group ‘X’ will inevitably come into conflict with members of other group and they will be prosecuted and punished, but such conflict, prosecution, and punishment are an empirical difficulty that results from false promising, not a conceptual contradiction, as Kant has argued.

In further evaluation, we may note that it is logically possible to generalize one’s proposed action of making a false promise into a class or group which can have only one unique member, saying ‘All who are uniquely me with my genetic endowment and all my personal experiences may make false promises whenever they wish to do so.’ There is no conceptual contradiction in such a generalization.

---

1 See also Samuel J. Kerstein, “Deriving the Supreme Principle of Morality from Common Moral Ideals,” The Blackwell Guide to Kant’s Ethics, edited by Thomas E. Hill, Jr. (Oxford: Blackwell Publishing, 2009) p. 132: ‘Consider the bizarre principle, BP: “Act only on that maxim such that you cannot, at the same time, will that it become a universal law.” Assuming
action must be willing logically to be logically consistent and to generalize one’s action in meeting similar situations. However, it is not a requirement of logic that one must generalize to the class of all rational agents; one may generalize to one’s own group ‘X’ which treats all members within the group morally but not those outside the group. Indeed, one could even generalize to a unique group with only one possible member. Of course, such choice of a special group ‘X’ or, even worse, of oneself as above and apart from all other people in allowing only oneself to make false promises is an immoral choice of giving special dignity to one’s special group ‘X’ or to oneself. What morality claims is true dignity for all as persons. The FH is precisely the reason why the FUL must be upheld.¹

Kant goes on to argue that the value of humanity as an end in itself requires the rejection of the proposed suicide and the proposed false promise. As to the proposed suicide, Kant argues that the proposed suicide treats the self as a mere means to the end of happiness in being willing to kill oneself when happiness is no longer possible (Kant 1909 47). As to the proposed false promise, Kant argues that this proposed false promise treats the others as mere means who are used and abused for one’s own purposes (Kant 1909 48).

Although Kant himself has explicitly rejected all proposed suicides and proposed false promises, we might apply his principle of humanity and draw different conclusions in these two cases if an agent were facing great natural or moral evil. For example, if an agent in a terminal illness were facing unbearable pain, the agent might think to herself that continuing to live in such unbearable pain is an assault on human dignity and hence suicide is justifiable as a means of protecting human dignity.² In fact, such an agent could rationally generalize her proposed action without any conceptual contradiction to all such rational agents who face such unbearable pain in a terminal illness. Again, if an agent of the French Underground is captured by the Gestapo during WWII and knows that she will not be able to withstand torture, she might conclude that protecting the dignity and lives of her fellow members

¹Korsgaard, p. 178, confirms this argument by noting that Kant himself points out that each ‘action has an end, and choice is always determined by an end (G 427; MPV 381, 384-85; R’). So a maxim of action, or of the means to an end, is adopted freely only when you have adopted the maxim of holding that end. . . . How can it be necessary to have certain maxims? The answer is that if there are ends that are duties, there will be maxims that it is a duty to have: maxims of actions that promote those ends. . . .[specifically in the “Supreme Principle of the Doctrine of Virtue” which runs “act according to a maxim whose ends are such that there can be a universal law that everyone have these ends” (MPV 395).]

²See Stephen Darwell, “Why Kant Needs the Second Person Standpoint,” The Blackwell Guide to Kant’s Ethics, edited by Thomas E. Hill, Jr. (Oxford: Blackwell Publishing, 2009) p. 150, who writes in evaluation of Kant’s argument for the self as an end in itself: ‘I could quite intelligibly deliberate under the assumption that, although my rational thought has great value now, it might not some time from now. It is not obvious, of course, how I could justify such an assumption. But there is nothing incoherent in it. . . . I could certainly think that the value of my thinking and acting rationally in the future is overridden by other values, if, for example, I were given a Hobson’s choice in which I could continue as a well-functioning rational agent only at the cost, say, of my children’s lives.’
morally permits her to commit suicide. Furthermore, she would be willing to universalize this proposed action to any other agent in a similar situation. Consequently, even if in facing great physical or moral evil we do draw different applications from the FUL and the FH, the strength of these two formulations of the categorical imperative remains for our wise applications to our many different situations.¹

**Moral Problems 3 and 4**

The third moral problem is a proposed refusal to develop one’s own talents even though they might be useful to humanity’s advancement, and the fourth moral problem is a proposed refusal to assist others in great need even though one’s own circumstances are comfortable. Korsgaard correctly sees Kant as attempting to derive only imperfect duties of avoiding refusing to develop one’s own talents and refusing to assist others in great need. She notes that Kant acknowledges that there is no inherent conceptual contradiction in universalizing the refusal to develop one’s own talents and in refusing to assist others (Korsgaard, 14).

However, Kant goes on to argue and Korsgaard emphasizes that every rational agent would never will to universalize these proposed maxims because of the effects of such a universalization. Any agent who universalized not developing her talents and not helping others in need cannot guarantee that in all her future life that she would never need to helped by the well-developed talents of others, for example, in a surgery, or never need to helped by others during a great natural disaster. An agent who universalized not developing talents and not helping others would find herself contradicting herself at that later time when she would likely need the talents and the help of others and would will a different maxim as the guidelines for all others.

In evaluation, we may certainly choose to agree with Kant’s argumentation here as highly probable in light of the contingency of future events and our likely need for help from others, and I do so, but it is not logically necessary that the agent who refuses to develop talents now or to help others now must will in the future to be helped by the talents of others in the future. The agent could simply stick to her resolution that no one needs to develop talents or to be helped by others. She could say, ‘Live and let live. I choose to live completely by own present abilities and never to make a claim on any others in the future, and they have no claim on me.’ She might end up in great difficulty, but she might not. Neither future events nor logic necessarily requires her to universalize the development of talents or to help others in great distress. Indeed, she could generalize her ability to develop her talents only for a special group ‘X’ or generalize her ability to help others in her special group ‘X,’ but there is no future event nor rule of logic that

¹See Thomas E Hill, Jr., “Humanity as an End in itself,” *Ethics*, Vol. 91, No. 1 (October 1980) p. 95: ‘although suicide contravenes a “perfect duty,” still Kant leaves open “casuistic questions,” for example, whether it is wrong to kill oneself in anticipation of an unjust death sentence, or in order to save one’s country, or to escape an impending madness resulting from the bite of a rapid dog (MEV 84-85 [423-24]).’
necessarily requires her to universalize the development of human talents by all people so that they be useful to all of humanity or to universalize the principle that all agents who are comfortable must morally help other agents in great stress.\(^1\)

Especially, then, logic does not require the universalization of developing one’s talents or of helping others. Kant is wrong. Morality is not founded upon logic. However, Kant is right that the formula of Humanity as an End in Itself is at the heart of morality. As Kant argues in problems 3 and 4, the agent who refuses to develop her own talents or who refuses to help others in distress does not directly harm the humanity of others, but she is not advancing the development of humanity. As Kant notes in problem 4,

\[\ldots\text{[H]}\] humanity might indeed subsist even though no one should contribute anything to the happiness of others. However, such a state of affairs would only harmonize negatively, not positively, \textit{with humanity as an end in itself} if every one does not also endeavour, as far as in him lies, to forward the ends of others. For the ends of any subject which is an end in himself, ought as far as possible to be my ends also if that conception is to have its \textit{full} effect with me (Kant 1909: 48-49).

This moral principle that one assist others in distress does not require that the helping agent do everything for those in need but that one help the others in restoring their basic abilities and basic wellbeing to be productive human beings through their own rationality and freedom. All the ends of these people in need for whom I have an obligation to help are being respected and have their full effect upon me since I am assisting them in retaining and building upon their own rational and free initiative.

\section*{Evaluation of the Formula of Humanity}

In evaluation of Kant’s application of the Formula of Humanity to problems 3 and 4, I completely agree with the resulting moral resolution of these cases. However, I find great difficulty in Kant’s understanding of the FH. Kant wishes to defend the value of humanity as not based in a contingent human desire and/or in a contingent decision to value oneself as a rational and free agent. The value of humanity as an end in itself is not borrowed from experience, Kant claims, because of two reasons: ‘\textit{firstly}, because it is

\[^1\text{See Richard Dean, “The Formula of Humanity as an End in Itself,” The Blackwell Guide to Kant’s Ethics, edited by Thomas E. Hill, Jr. (Oxford: Blackwell Publishing, 2009) p. 93, where he evaluates Kant’s FH as follows: ‘There is a large gap to be filled in the move from saying that each agent must treat her own rational nature as an end in itself to saying that each agent must treat every rational nature as an end in itself.’ See also Samuel J. Kerstein, “Deriving the Supreme Principle of Morality from Common Moral Ideals,” The Blackwell Guide to Kant’s Ethics, edited by Thomas E. Hill, Jr. (Oxford: Blackwell Publishing, 2009) p. 132: ‘Another . . . example of a principle Kant would be unable to dismiss on the basis of his criteria is the following principle of weak universalization, WU: “Act only on that maxim which, when generalized, could be a universal law.” WU is not equivalent to the Formula of Universal Law. And Kant himself suggests that a maxim of non-beneficence could, when generalized, constitute a universal law . . .’} \]
universal, applying as it does to all rational beings whatever, and experience is
capable of determining anything about them; secondly, because it does not
present humanity as an end to men (subjectively), that is as an object which
men do of themselves actually adopt as an end; but as an objective end, which
must as a law constitute the supreme limiting condition of all our subjective
ends . . . ’ (Kant 1909 49).

Recognizing that a true universal value of humanity cannot be rooted in
the experiential desire, Kant affirms that the value of humanity is inherent in
the very nature of rational and free agency. It is the pre-existing condition of
any agency whatsoever, and it is primarily a limiting condition upon our
human action. We must never violate, degrade, harm or destroy the dignity of
humanity. However, Kant’s full answer to the third problem acknowledges
that a person might refuse to develop his talents and that this refusal does not
violate humanity in one’s own person as an end in itself. Kant says that action
should not only not violate humanity as an end in itself but that action should
also advance this end. Hence, it is immoral not to develop one’s talents when
they may be useful to others.

It is clear that this solution by Kant violates his own guideline that moral
decisions should not be based on the consequences of action. Kant had said that
an act should not be judged in the light of the consequences or ends to be
effected (Kant 1909 56). Kant had said that the principle of humanity should
not be used as a positive goal but as a negative principle which prohibits
actions against the value of the human being. Rader evaluates Kant’s argument,
writing, ‘But is it possible to carry out the formula, “so act as to treat
humanity, whether in your own person or in that of another, always as an end
and never as a means only,” without a view to the effects of actions? Must we
not have some positive idea of the end of man and how to achieve them?’
(Rader 564). For example, how do we know what the conditions of political
and economic freedom should be, unless we see the effects upon the positive
development of the community and individuality and creativity?

When I attempt to respond for Kant to my objection, this is how I would
argue. Rationality and freedom as human faculties are never finished products,
nor are they discovered in my analysis of my agency as finished products. I
discover them as processes that I value inherently whenever I value anything
else that I desire, for all other things that I desire are desired as means to the
value of humanity in myself and others. And humanity, especially in our
abilities of being rational and free, is essentially in process. I would not truly
value my humanity as an end in itself unless I both reject as immoral the doing
of any injury to humanity and I reject as immoral the refusal to develop
rationality and freedom as processes in need of further development. For no
person’s reasoning or choices are perfect at age 20 or 30 or 40, or, indeed, at
any age. So Kant is right that the value of humanity is the pre-existing
condition and ultimate value of any human choosing, and this pre-existing
condition is inherently in process and in need of further actualization.
Consequently, the Formula of Humanity as an End in Itself requires not only
the prohibition of any deliberate harming of the human person but also the
advancement and actualization of humanity in positive and creative ways.
Although my response on behalf of Kant’s FH is appealing as a defense of the inherent value of humanity both as a value not to be harmed and as a value to be advanced, Kant himself reveals that his ethical system relies upon the assumption of teleology in nature and humanity, when he affirms that plants are created for plant-eating animals, and plant-eating animals for flesh-eating animals, and all of nature for the sake of humanity, ‘the ultimate purpose of creation here on earth, because he is the only being who can form a concept of purposes and who can, by his reason, make an aggregate of purposively formed things a system of purposes’ (Kant 1951:426-27).

Furthermore, Kant argues that there must be some ultimate end worth being chosen for its own sake because otherwise all actions would simply be means chosen to ends which themselves would be means and so on into infinity and restlessness. He argues as follows:

That the existence of something must be an end in itself, and not all things can be merely means, is just as necessary in the system of ends as Ens a se is in the existence of efficient causes. A thing that is an end in itself is a Bonum a se. What can be considered merely a means has its value as a means only when it is used as such. There must be therefore a being that is an end in itself. A thing in nature is a means for another; that goes on forever, and it is necessary at last to think of a thing that is itself an end, otherwise the series would come to no conclusion (quoted in Guyer p. 152 from Naturrecht Feyyerabend, 27:1321).

Kant’s presupposition of humanity as the Bonum a se, as ‘the ultimate purpose of creation here on earth,’ itself depends upon the assumption of human freedom as the condition which makes moral obligation possible. But Kant himself insists that freedom of the will is a noumenal presupposition which he himself cannot demonstrate. Korsgaard emphasizes that Kant reveals a central point of his ethics in the practical postulates of immortality, freedom of the will, and the existence of God (Korsgaard, 172-173). She writes:

The standpoint from which you adopt the belief in freedom is that of the deliberating agent. You are licensed to believe in the practical postulates because they are the necessary conditions of obeying the moral law. Thus it is primarily your own freedom that you are licensed to believe in. . . . (Korsgaard, 174)

Kant has assumed three points, (1) a teleological description of the plant, animal, and human kingdoms, which culminates (2) in human beings having

---

1See also Thomas Auxter, Kant’s Moral Teleology (Macon, Ga.: Mercer University Press, 1982) p. 165: ‘we are “endowed with freedom (of causality)” and consequently “we find in ourselves . . . a moral teleology.”

‘This teleology stems from our moral identity and defines the nature of human choice. It “has to do with the reference of our own causality to purposes.” Because we have a moral identity, we are essentially connected with the rest of the world in two basic ways. First, other beings who share this identity must be treated as ends, and we therefore must “pass judgment” on what will uphold this value in interactions with them. . . . This final purpose or goal must be the comprehensive moral order, the ultimate moral teleology, that issues out of a systematic organization of the human life-world “with reference to our own causality” to moral ends. (Kritik der Urteilskraft, KgS, 5:447-48 (298)).’
the absolute moral purpose of realizing the value of humanity as an end in itself, which itself assumes (3) the noumenal freedom of our human will. However, all three points can be challenged.

First, a teleological description of nature in physics, chemistry, and biology is not a presupposition of these sciences. There are billions of galaxies with billions of stars, and the chance occurrence of a star with a planet with temperate climate and water which can by chance evolve living forms may be simply a fortuitous occurrence. Furthermore, from the viewpoint of evolutionary biologists there is no necessity that the human species, homo sapiens sapiens, had to evolve, but it may be our lucky accident that our specific species evolved from pre-existing species of homo.

Second, the assumption of human beings as having the absolute moral purpose of realizing the value of humanity as an end in itself is itself, as Kant tells us, ‘a moral principle [which] is nothing but a dimly conceived metaphysics, which is inherent in every man’s rational constitution’ (Kant 1983 376). Such a dimly conceived metaphysics can be challenged by those deeply affected by the absurdity of life in the 20th and 21st centuries. As Camus argues, ‘At a certain point, on his path the absurd man is tempted. History is not lacking in either religions or prophets, even without gods. He is asked to leap. All he can reply is that he doesn’t fully understand, that it is not obvious’(Camus 52-53).

Furthermore, even Kant’s own adoption of a teleological perspective on nature, plants, animals, and humans, is, he points out, not a constitutive metaphysics of things in themselves, ‘but only . . . a regulative principle of the cognitive faculty’(Kant 1951 5:197). However, as Guyer emphasizes for Kant, ‘By seeing our freedom as the ultimate end of nature, we can give ourselves a dignity that we lack as mere organisms of nature, or elevate ourselves above nature, but nature itself cannot force us to dignify ourselves in this way, nor can any theoretical proposition about nature force us to see ourselves in this way’ (Guyer, 169). This affirmation of human dignity is precisely a choice, a choice which the absurd person may accept, but it is not compelled by evidence.

Indeed, the person who has accepted absurdity can easily reject Kant’s argument that there must be a Bonum a se since otherwise there would be only means to further means. There is nothing absurd in a series of efficient causes there goes on in regression endlessly, as Kant himself has argued about the phenomenal world; we do not have to conclude to an Ens a se as the Uncaused Cause of a series of efficient causes. So also, we do not have to conclude to a Bonum a se because all things chosen as ends could also be chosen as means to

---

1Cf. John E. Hare, The Moral Gap: Kantian Ethics, Human Limits, and God’s Assistance (Oxford: Clarendon Press, 1997) p. 94 connects Kant’s ethics of reason with moral faith, arguing that moral faith in the validity of moral values is necessary for practical reason if it can be shown that theoretical reason has not disproven universal moral values. However, he weakens his argument by admitting that ‘moral faith is consistent with some doubt about whether your continued well-being is consistent with your trying to live a morally good life. “Lord I believe; help thou mine unbelief” is a possible frame of mind. It is possible especially when . . . faced with a particularly glaring example of the suffering of the innocent and the triumph of the guilty.’
further ends and further means even though the human heart would ever be restless in seeking new goals which would never finally satisfy human desire.¹

We might very well wish that all action is ultimately for a *Bonum a se*, specifically, in human freedom, but there is no proof of human freedom.²

Third, Kant’s assumption of the noumenal freedom of our human will can be challenged. Whereas Kant has attempted to root absolute morality in human rationality and freedom alone, other philosophers such as David Hume have attempted to root morality precisely in human feeling without the assumption of noumenal freedom of the will. Such empirical attempts of Hume and others to emphasize human feelings can make better sense out of such cases as this:

Rachel Bachner-Melman, a clinical psychologist at Hadassah University Medical Center in Jerusalem who specializes in eating disorders, has seen the impact of extreme selflessness on the anorexic young women who populate her ward.

“They are terribly sensitive to the needs of those around them,” she said in an interview. “They know who needs to be pushed in a wheelchair, who needs a word of encouragement, who needs to be fed.”

Yet the spectral empaths will express no desires of their own. “They try to hide their needs or deny their needs or pretend their needs don’t exist,” Dr. Bachner-Melman went on. “They barely feel they have the right to exist themselves.” They apologize for themselves, for the hated, hollow self, by giving, ceaselessly giving (Angier).

Reasoning alone seems insufficient to reestablish a deep sense of value in their own value. Such spectral empaths need rather a deep feeling for and/or a profound existential choice of the value of their own selves. If correct reasoning were enough to establish this deep value, then mere conversation with them would be sufficient to convince them rationally of the value of themselves. For just as they value other human so highly in their deep empathy for others, so also they logically should value a deep empathy for themselves. However, mere logical conversation cannot heal this disease of feeling. They must learn experientially to feel for themselves as they feel for others. Such a genuine feeling

---

¹Richard Dean, “The Formula of Humanity as an End in Itself,” *The Blackwell Guide to Kant’s Ethics*, edited by Thomas E. Hill, Jr. (Oxford: Blackwell Publishing, 2009) p. 86, points out, “Morality is the only condition under which a rational being can be an end in itself. . . .” . . .In the *Critique of Judgment*, Kant says that “it is only as a moral being that man can be the final purpose [end, or Zweck] of creation.” So, I argue, it is not humanity as an end in itself that is the foundation of universal morality, but it is the emotional preference for and existential choice of universal morality that truly establishes humanity as an end in itself of creation.

²See Korsgaard, *Creating the Kingdom of Ends*, p. 176. She argues that for Kant it does not matter ‘if we are in fact (theoretically) free. . . . Reason becomes an efficient cause by telling us how a free person would act [that is, with respect for the Kingdom of Ends]. . . . for if the moral law does indeed provide the positive conception of freedom, then we know how a person with a completely free will would act. . . . But if we are able to act exactly as we would if we were free, under the influence of the idea of freedom, then we are free. . . . By acting morally, we can make ourselves free.’
of the value of both oneself and others is essential for a general morality that would embrace all humans, the Humean would argue.

If reasoning alone cannot transform such people, then either such transformation can occur only through feeling or through existential choice. Either such transformation can be found in the deepest feelings for the value both of self and all other rational agents which deeply felt emotional therapy may help a person recover, or such transformation can occur both with the feeling for and profound existential choice of the value both of self and all other rational agents Kant’s defense of the Formula of Humanity as an End in Itself needs to be supplemented either with the empiricist’s deep usage of human feelings as a key source of human morality and/or with the existentialist’s profound choice of the value of self and others as ends in themselves.

Bibliography
