

A Duty to Live: Kant on Suicide

Woojin Lim, Harvard University

Abstract: In this essay, I draw upon Kant's ethical theory to establish a duty to oneself, a duty to live. I work through three variations of the Categorical Imperative to argue at length against all forms of suicide. I distinguish between forms of 'self-killing' (*autochiria*) and 'suicide' (*suicidium*), and further between the act-type of 'suicide' and the action of 'suicide'. I conclude that on Kant's ethics, suicide is morally impermissible, but note that his moral framework may admit of one exception: self-killing (and perhaps 'suicide') could be morally permissible only when one must otherwise forfeit life as a rational and autonomous agent.

INTRODUCTION¹

Not to commit suicide, for Kant, is the "*first*, though not principal, duty of a human being to himself as an animal being" [MM 6:421].² Kant says that the voluntary act of destroying oneself can "never, under any circumstances, be regarded as allowable." [LE 27:603]. All the while, Kant admits of cases like Cato where suicide appears a "virtue" and may be defensible [LE 27:371].³ On the one hand, Kant lays down with an iron fist against all acts of suicide, and yet, on the other, Kant appears to permit select forms of suicide. Surely, suicide is impermissible when a person kills herself out of self-love, but what about for saving her country or sacrificing herself for the good of all humanity? On Kant's ethical framework, do cases like Cato license suicide? Even if not, are there any other cases where suicide, on Kant's account, seems permissible?

Drawing upon Kant's G and MM, I will begin by assessing Kant's two central arguments for why suicide is immoral. The first argument is based on the first variation of Kant's Categorical Imperative (CI), the Formula of Universal Law (FUL), while the second argument is based on Kant's second and third variations of CI, respectively, the Formula of Humanity as an End in Itself (FH) and the Formula of Autonomy (FA). After presenting

¹ I would like to thank the Roger Mudd Center for Ethics and all of the commentators at the Washington and Lee University Ethics Conference 2020. In particular, I would like to thank Professor Christine Korsgaard for her memorable class, Kant's Ethical Theory, where I began this paper. In addition, I would like to thank Justin Cavitt, Michael Cholbi, Thomas Hill, Rae Langton, and Allen Wood for providing comments on earlier drafts.

² I abbreviate Immanuel Kant's *Groundwork of the Metaphysics of Morals* (G), *Metaphysics of Morals* (MM), and *Lectures on Ethics* (LE). For an overview of the central questions posed here and a modified Kantian alternative that emphasizes the role of human dignity, see Thomas Hill, "Killing Ourselves: Suicide and the Appreciation of Life," in Steven Luper, ed., *Cambridge Companion to Life and Death*, Cambridge University Press, 2014, pp. 265-81.

³ Kant writes in full, "One must certainly admit of this example [Cato, who impales himself with a sword to bleed to death once he realized that it would not be possible for him to escape falling into Caesar's hands], that in such a case, where suicide is a virtue, there seems to be much to be said for it. It is also the one example that has given the world an opportunity of defending suicide." [LE 27:370]

both views, I will address the nuances—the ‘caustical questions’ from MM and the Cato example from LE. I will first distinguish between ‘suicide’ (which involves the deliberate intention to destroy oneself) and other forms of ‘self-killing’ (some of which might involve the intention to preserve oneself). Following Korsgaard, I will further distinguish between the act-type of ‘committing suicide’ and the action ‘committing suicide for ends-Y’.⁴ I will conclude that in the particular examples Kant provides, *prima facie*, suicide is, strictly speaking, morally impermissible, but I will come to show that his moral framework admits of one exception: self-killing (and perhaps ‘suicide’) could be permissible only when one must otherwise forfeit morality.

I.

A. Kant’s first argument against suicide

In G, Kant first argues that suicide violates FUL. Under FUL, rational beings ought to always “act as if the maxim of your action were to become by your will a universal law of nature” [G 4:421]. Kant asks whether the maxim “From self-love I make it my principle to shorten my life when its longer duration threatens more trouble than it promises agreeableness” (henceforth, ‘maxim S’) can be willed as a universal law [G 4:422]. Here, Kant says that a law which seeks to destroy life while at the same time impelling toward the furtherance of life would “contradict itself” and thus could not possibly be a law of nature [G 4:422].

The contradiction Kant mentions here cannot arise from logical or practical contradiction interpretations of FUL.⁵ There is nothing inconsistent or contradictory about imagining my willing of maxim S in a world where everyone is committed to maxim S for my same purposes. In such a world, it would still be possible for me to formulate and will maxim S—to end my life given the circumstances of self-love—and I do not thwart my own purpose in formulating maxim S when it is universalized. On the logical and practical contradiction interpretations of Kant’s FUL, no contradiction arises when maxim S is willed universally.

The teleological contradiction interpretation of FUL appears most fruitful. A contradiction arises when agents try to imagine a universalized maxim S in conjunction with a law of nature that holds “the furtherance of life” is the natural purpose of self-love. Suicide uses the impulse of self-love contrary to its natural purpose. Maxim S violates FUL, since it is impossible for agents—without contradiction—to be motivated by self-love to pursue “the furtherance of life” as part of a system of nature teleologically conceived while at the same time being motivated by self-love to commit suicide. Since self-love cannot motivate agents to use themselves both as a means to further their lives while also as a means to destroy their lives, maxim S contradicts itself.⁶

⁴ Korsgaard, “Acting for a Reason.” *The Constitution of Agency*. 2008, pp. 207-229. Oxford: Oxford University Press.

⁵ Korsgaard, “Kant’s Formula of Universal Law.” *Creating the Kingdom of Ends*. 1996, pp. 77-105. Cambridge: Cambridge University Press.

⁶ On an interpretive note, even Kant scholars digress on interpreting Kant’s three CI formulations. Allen Wood agrees with my conclusion—that Kant does not consider all act-types of ‘self-killing’ to be morally equivalent, and that particular actions of self-killing (and perhaps some ‘suicides’) could be morally permissible—but thinks that on the way to it, I had to deal with a number of factors that make the conclusion seem harder to draw than it is. Wood suggests that Kant’s FUL and law of nature [G 4:402-03 and G 4:421-424], and their intended use, are commonly

B. Kant's second argument against suicide

In MM, Kant argues that all rational beings have a duty to preserve themselves in their animal natures, since a person's body is the vessel through which humanity (and autonomy) manifests. For Kant, the body is tied to a rational being as a "condition of life" such that there is no way to dispose of the body without also disposing of the life to which the body is inextricably attached [LE 27:369]. For Kant, the willful killing of oneself stands contrary to FH and FA.⁷ To annihilate the very subject of morality in oneself is to root out the existence of morality—humanity and autonomy—from the world.

When rational beings commit suicide, they violate FH by using themselves merely as a means to destroy their humanity, and by failing to treat themselves as an end-in-itself. The human being is not merely a "thing"—"an instrument of animal gratification" [LE 27:1427]—and hence cannot be used merely as a means, but must always be regarded as a person, an end-in-itself [G 429]. Disposing of oneself as a mere means to some discretionary end (i.e. self-love) debases and dishonors one's humanity [MM 6:423].

To abandon at a stroke all the hardships of life is to completely subordinate humanity to "animal nature" [LE 27:1428]. The decision to commit suicide is under the sway of "animal impulse," not an autonomous will [LE 27:1428]. Kant writes, "To dispose over his person as a means is *not a choice open to him*" [LE 27:343], notably because freedom in its greatest condition must be consistent with "order in nature" and the "essential ends of mankind" [LE 27:346]. In committing suicide, according to Kant, humans employ their powers and freedom against themselves, to make themselves a carcass [LE 27:343].

Since free and rational beings are obligated to respect the humanity inherent in their persons, suicides are immoral because they stem from maxims that violate this duty to oneself, resulting in a failure to use one's freedom in a manner consistent with the duty to treat oneself an end-in-itself. In a sense, Kant's first argument blends in with the second: To adopt natural impulses (i.e. self-love) contrary to their natural ends (i.e. self-preservation) is to disrespect one's humanity and autonomy.⁸

misunderstood such that the only defensible outcome is that they fail to show anything. Kant's argument from FLN takes for granted that there is a general duty not to commit suicide and asks only whether the agent in that case would be justified in making an exception to it; this is not an argument purporting to show that suicide is always wrong. To read it that way is to misunderstand it. See Wood, *Formulas of The Moral Law (Elements in the Philosophy of Immanuel Kant)*. 2017, pp. 13-60. Cambridge: Cambridge University Press.

On the other hand, Michael Cholbi thinks that very few commentators have found the FULL argument convincing, under any interpretation. For reasons why it fails even under the teleological interpretation, see Cholbi, "Objections to the Formula of Universal Law." *Understanding Kant's Ethics*. 2016, pp. 145-173. Cambridge: Cambridge University Press.

⁷ FH states, "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" [G 4:429], and FA states, "Act only so that the will could regard itself as at the same time giving universal law through its maxim" [G 4:434].

⁸ I would like to thank Allen Wood for the following points. To add, Kant is explicit that arguments against suicide based on FH require good judgment and careful interpretation to prevent misunderstanding [G 4:429-430]. What FH justifies is that we must not dispose of our person. In cases where suicide would have the meaning simply of disposing of our person as a thing of no value, FH would condemn it. In other cases of suicide, the agent's rational capacities are diminished and so is her responsibility for her action. There certainly are cases of suicide where the act affirms

II.

Up to this point, I have reconstructed Kant's arguments that suicide is immoral. I have explained how suicide contradicts the natural purpose of rational beings for self-preservation and how suicide debases humanity and autonomy in one's person. Now, I will proceed to examine a number of caustical questions from MM and the story of Cato from LE to ask whether suicide is always impermissible.

In MM, under the article "On Killing Oneself," Kant has a subsection on "Caustical Questions". Here, Kant raises a number of questions. Is suicide permitted if it's done in order to save one's country? What about for the good of all humanity? What if a military general wants to avoid being ransomed on harmful conditions for his country? What if a man catches an incurable and highly infectious disease? In these cases, are individuals morally permitted to commit suicide? Though Kant never directly provides indication as to what purpose these questions serve, Kant does not appear to claim that such cases are valid exceptions to the immorality of suicide.

First off, it is important to note a distinction between 'self-killing' and 'suicide'. Ostensibly suicidal actions—for instance, amputating a limb or receiving smallpox vaccinations—are not 'suicide' if they are done for the purpose of self-preservation.⁹ The soldier who fights to his death on the battlefield instead of running away to save his life from the enemy: Kant explicitly says "that is no suicide" [LE 27:371]. Placing life in danger "to preserve life" belongs to "morals proper" [G 4:429]. Such actions might fall under the broader category of self-killing or putting oneself at harm's way. However, 'suicide' uniquely involves the "intention to do away with oneself," that is, a deliberate intention (*dolus*) to destroy oneself [LE 27:372]. Individuals who might be to blame for their deaths ('self-killing') do not commit 'suicide' if they lack this *dolus*. Such cases of 'self-killing', insofar as they are guided by duty (to oneself, to others, etc.), might be permissible, whereas suicide is not.

Speaking of suicide in particular, it is important to distinguish between 'act-types' and 'actions'—a distinction Korsgaard spearheads in "Acting for a Reason." An 'action' describes the entire formulation of a maxim (i.e. 'I will do act-A in order for end-E'), whereas an 'act-type' describes the first part of the maxim (i.e. 'I will do act-A'). Applied, the phrase 'committing suicide in order to X' refers to an 'action', whereas the phrase 'committing suicide' refers to an 'act-type'. The act-type of suicide may correspond to a

rather than denies the value of the agent's person: Cato is represented in Addison's play as such. It is a common misinterpretation of FH to argue from "Act A treats person P as a means" to "Act A violates FH." This inference is in general invalid.

Kant is perhaps offering a specious argument for the moral prejudices of his time and culture in suggesting that any act of suicide must treat your person only as a means and not also as an end. But the case of Cato, or of many people who choose to end their lives rather than descend into a dehumanized condition of degrading debility, are treating their persons as ends precisely by ending their lives in this way. Kant's discussion of the Cato example shows he was aware of this, even if the prejudices of his time and culture prevented him from saying so. Hume was more forthright in his essay on the topic, but he too, in a different manner, had to make concessions to the prejudices of his age when he withheld the essay from publication and let it be published only after his death.

⁹ In 'self-killing', one might also use oneself as a means to an end, but for self-preservation is in line with one's humanity and autonomy, and one does not use oneself merely as a means to an end.

number of different actions (with different ends-E). In constructing his arguments against suicide in G, Kant assumes that the one and only ends-E for suicide is for self-love. As Kant notes in ‘Caustitital Questions’, however, one may commit suicide out of rage, or out of a duty to avoid harming others, or for good of all humanity. The act-type ‘committing suicide’ bears a number of different maxims, and here, the ends-E (or purpose) determine the action’s moral worth.

Accordingly, one plausible explanation for the ‘Caustitital Questions’ section is that Kant encourages us to further deliberate on (1) whether or not an act of self-killing qualifies as suicide based on the maxim, and (2) what ends-E an action of suicide contains (and whether some ends-E may carry moral worth).

On one hand, it seems irrelevant for us to consider the ends-E of an act of suicide, given Kant’s fierce opposition to all flavors of suicide, no matter the circumstances. *Prima facie*, there is no way Kant can claim that suicide is morally permissible (or praiseworthy) without contradicting his stance on why all suicide is immoral, a position Kant holds consistently throughout G and MM. All the while, in LE, there appears to be some wiggle room. Consider the following passages.

In LE, Kant writes that the observance of morality is far greater than life. It is better to “sacrifice life than to forfeit morality” [LE 27:373]. Those who can no longer live honorably are no longer worthy to live at all, and since an agent should seek to preserve her life only insofar as she is worthy to live it [LE 27:371], she may no longer have a binding duty of self-preservation. It is plausible that suicide might be permissible, Kant admits, when the continuation of life rests upon circumstances in which the agent can “no longer live in accordance with virtue and prudence,” and must therefore put an end to his life from honorable motives [LE 27:370]. Though misery and other inclinations give no person the right to take their life, the fulfilment of self-regarding duties (i.e. to live honorably) may “[demand] the sacrifice of life” [LE 27:373]. After all, echoing my introduction, avoiding suicide may be the first, but “*not principal*” duty of a human being to himself as an animal being [MM 6:421, emphasis mine].

Here, Kant brings up the example of Cato [LE 27:370], a Roman aristocrat who kills himself after the defeat of the Pompeians at Thapsus, only once he realized it would not be possible for him to escape Caesar’s hands. Cato viewed his death as a necessity, since he would no longer be able to live as himself, in his proper humanity and autonomy, if captured by Caesar, and since he has a duty of beneficence to his fellow citizens.¹⁰ Kant encourages us to view Cato’s death as an enticing defense of suicide, though Kant says in retort that such suicide is a rarity—“one example of its kind” [LE 27:371].

Later in LE, Kant says that if Cato had still adhered to a steadfast mind, despite all the tortures that Caesar may have inflicted on him, remaining alive and enduring torture would have been the noble thing to do, instead of committing suicide [LE 27:374]. That said, in converse, it seems that if Cato knew with certainty that he could not possibly

¹⁰ Kant sets aside the question of one’s obligations to other people (i.e. the duty of friends to each other, of parents to their children, of subjects to superiors, of citizen to fellow citizens, or to God), since his central issue at stake is whether or not a duty against suicide remains, even after these relations are set aside [MM 6:422]. Hill (2014, p.271) phrases Kant’s project as follows, “Do we “owe it to ourselves,” in some sense, to live and die with dignity?” It is worth noting that a weighing of these conflicting obligations might be another intriguing project to explore in further papers, and whether Cato’s case is exceptional only because of this duty of beneficence.

adhere to a steadfast resolve, and that Caesar's torturing tools would not only degrade his humanity but also melt his rational and autonomous capacities, I raise the possibility on behalf of Kant that Cato's suicide (for another ends-E besides natural inclination) might be permitted as an act of duty to oneself in line with the natural law ("the furtherance of life").¹¹ Provided that Cato adopt ends-E consistent with the CI (i.e. not from inclination, but from duty), suicide might be permissible.

This situation appears paradoxical—a kind of non-identity problem. On the one hand, committing suicide prior to losing one's autonomous will and humanity violates FH and FA, whereas on the other hand, once an individual's rationality disappears, so does their ability to continue acting morally. Here, I leave open the option of Kant answering both ways. On the one hand, perhaps Kant's views on suicide remain locked and steadfast—it's better to hold onto the last shreds of rationality than to violate a perfect duty (i.e. of self-preservation) to oneself. On the flipside, perhaps Kant might concede that it would be better to kill oneself prior to losing one's ability to act rationally or morally for good.

There seems to be a deeper metaphysical question at work: can the destruction of one's 'humanity' or rationality come apart from her body and life, such that in select cases, it would be permissible for one to destroy her body in order to preserve her humanity?

Interestingly enough, in sustaining this discussion on the link between 'duty' and 'humanity', Kant never brings up the word 'suicide'. Perhaps, killing oneself *out of duty*—similarly to killing oneself *for self-preservation* (i.e. precarious vaccination, limb-amputation, etc.)—might just be another variant of the act-type 'self-killing', not 'suicide'. All the while, it seems that self-killing out of duty (ends-E) requires, nonetheless, a deliberative intention to do away with oneself ('suicide'), though for different purposes than for self-love or inclination. At the very least, it seems that the act-type 'suicide' should be separated from the various actions of 'suicide'—nearly all of which are forbidden, while some, in very exceptional cases like Cato's (where duty requires one to deliberately kill oneself), is not impermissible.

To recapitulate, when the maxim for self-killing (or more specifically, suicide) bears an end-E that contradicts CI-formulations, then those particular actions are immoral. When the maxim for self-killing (or suicide) bears an end-E that rational beings are generally obliged to follow *out of duty*, directed by the CI (i.e. when one

¹¹ There appears to be some ambiguity regarding what Kant means by "the furtherance of life" [G 4:422]. Typically presumed, Kant refers to the furtherance of the agent's own life. But it might still be possible that Kant could also refer to the furtherance of others' lives (i.e. of the country, of humanity, etc.). Though Kant, for the sake of space, pushes aside the thought of suicide as a "violation of one's duty to other people" [MM 6:422, see previous footnote], I raise the possibility that perhaps, the duty of beneficence to others might be so powerful as to outweigh the duty to oneself, for the sake of furthering others' lives.

Such a duty appears in line with CI, since it respects the natural law of 'furthering life,' and respects one's own and others' humanity and autonomy as ends-in-themselves. Provided that suicide is performed either to preserve one's self from a foreseeable downfall in one's humanity and autonomy, or to preserve others' lives, suicide might be morally justified only insofar as it arises out of duty.

must forfeit morality), such actions of self-killing (and maybe ‘suicide’) are not morally impermissible.¹²

CONCLUSION

In this essay, I have argued that Kant does not consider all act-types of ‘self-killing’ (nor ‘suicides’ for that matter) to be morally equivalent, and that particular actions of ‘self-killing’ (and perhaps some ‘suicides’)—for *self-preservation* or *out of duty*—are not morally impermissible. All else equal, I have argued that actions of self-killing (or suicide) that do not stem from duty are otherwise morally impermissible, for they stand in violation of CI. At the very least, I pose a challenge to Kant’s statement that all act-types of self-killings (and some suicides)—under his moral framework—are impermissible simpliciter, and I put forth the conceptual possibility of justifiable or commendable forms of self-killings (and perhaps some ‘suicides’) as a topic for further study.

References

- Kant, Immanuel. Collins Lecture Notes. In: Schneewind, JB, Heath P, eds. and trans. *Lectures on Ethics*. Cambridge: Cambridge University Press, 1997:37–222.
- . *Metaphysics of Morals* (1797). In: Gregor M, ed. and trans. *Practical Philosophy*. Cambridge: Cambridge University Press, 1996:353–604.
- . *Groundwork of the Metaphysics of Morals* (1785). In: Gregor M, ed. and trans. *Practical Philosophy*. Cambridge: Cambridge University Press, 1996:43–108.

¹² On suicidal obligations that extend beyond mere permissions and on the duty to die, see Cholbi, “Kant on euthanasia and the duty to die: clearing the air” *Journal of Medical Ethics*, vol. 41, no. 8, 2015; Dennis R. Cooley, “A Kantian Moral Duty for the Soon-to-be Demented to Commit Suicide” *The American Journal of Bioethics*, vol. 7, no. 6, pp. 37–44, 2007.