**Deriving Positive Duties from Kant’s Formula of Universal Law**

**Guus Duindam**

**Abstract**

According to the objection from positive duties, Kant’s Formula of Universal Law is flawed because it cannot be used to derive any affirmative moral requirements. I argue that this objection fails and propose a novel way to derive positive duties from Kant’s formula. The Formula of Universal Law yields positive duties to adopt our own perfection and others’ happiness as ends because we could not rationally fail to will those ends as universal ends.

1. **Introduction**

The Universal Law Formula of Kant’s Categorical Imperative provides:

**Formula of Universal Law** (FUL): Act only according to that maxim through which you can at the same time will that it become a universal law (G 4:421).

Kant claims that FUL is the “supreme principle of morality” (e.g., G 4:392, MS 6:226), from which “all imperatives of duty can be derived” (G 4:421). Notwithstanding Kant’s ambitious claims, several interpreters have argued that FUL cannot be used to derive any affirmative moral requirements or “positive” duties (Allison 2011, 187–90; Wood 1999, 100–102; Kahn 2022, 2014). On this view, FUL could tell us only what we are prohibited from doing. Call this the *objection from positive duties*.

 The objection from positive duties can be motivated by the following concern. Kant’s applications of FUL suggest that we apply it by considering whether we could will that our maxim serve as a universal law. If we cannot—either because we cannot conceive of it as a universal law, or because we cannot rationally will it to be one—then we may not act on our maxim. If we can, then we *may* act on it. Because this process appears to leave no room for the obligatory, it seems that FUL is not equipped to derive positive duties (Wood 1999, 100–102; Kahn 2014, 96–97). This is a serious problem, for as Kant himself clearly recognizes (e.g. MS 6:392-394), morality imposes affirmative requirements. If FUL cannot be used to derive those requirements, then by Kant’s own lights, it could not be the supreme principle of morality, nor the principle from which all imperatives of duty are derived.

 This paper offers a new way to derive positive duties from FUL using the contradiction in the will test. First, I argue that FUL survives the objection from positive duties so long as it is capable of deriving positive duties of *virtue*, or wide positive duties. Second, I argue that the prevailing method for doing this does not work. Third, I show that FUL can nevertheless be used to derive positive duties and illustrate this by deriving Kant’s central positive duties: the duties to adopt one’s own perfection and the happiness of others as one’s ends (MS 6:392-394). We can derive such duties by showing that we could not rationally fail to will our own perfection and others’ happiness as universal ends. I conclude that the objection from positive duties fails.

1. **Positive Duties**

To understand what it would take to show that FUL survives the objection from positive duties, we must be clear about the role of positive duties in Kant’s ethics. It is uncontroversial that two positive duties play a central role: the duties to perfect oneself and to promote the happiness of others (MS 6:392-394, see also Herman 2007, 254–76; Timmons 2021, 98–111).[[1]](#endnote-1) These duties, which Kant also calls duties of virtue, are at the core of Kant’s *Doctrine of Virtue*. Kant repeatedly emphasizes that duties of virtue are duties to set ends: he explains that they cannot be coercively enforced (6:381), that they constrain only our “inner freedom” to set ends, not our “outer freedom” to engage in actions (6:396), and that they are “ends that are also duties” (6:385). Thus, the positive duties of virtue are distinct from what Herman has called duties of “performance,” i.e., duties to engage in particular (kinds of) actions (Herman 1993, 168–69). They are simply duties to set one’s own perfection and others’ happiness as one’s ends.

If FUL is to survive the objection from positive duties, then, it must certainly be capable of deriving the positive duties of virtue, or the two ends that are also duties. My claim is that this is *all* FUL would need to do in order to survive the objection. FUL does *not* need to be capable of deriving positive *performance* duties, i.e., positive duties to engage in actions. That is because it is perfectly possible to have a complete and plausible normative ethics without ever recognizing a single positive duty of performance.

To see why, consider the distinction between obligations and duties. A duty is a general rule or law of morality, such as “set your own perfection as an end” or “never make lying promises.” An obligation is an individual *instance* of being bound by a duty (on this distinction see Timmermann 2013, 42–43). Thus, a positive performance obligation would be an obligation to do a particular action at a particular time (keep the promise you made to Joey yesterday afternoon); a positive duty of performance would be a general moral requirement sometimes or always to do that action (always keep your promises).

It seems to me that while any plausible theory of morality would need to account for positive performance obligations, there is no need for positive duties of performance.[[2]](#endnote-2) Duties of performance are a blunt method for capturing our positive duties. Instead of saying that we must “sometimes donate to charity” for instance—a duty that would apply equally to the richest and the poorest among us—it seems far better to say with Kant that we must set the happiness of others as our end. Putting the duty in terms of required ends permits us to recognize that our affirmative obligations to act in pursuit of those ends can differ greatly depending on the circumstances. And duties to set ends still yield the performance obligations that morality *does* require. It seems clear, for instance, that when I can easily save a life without doing great harm to myself, my *duty* to have the end of others’ happiness yields an *obligation* to save that life (for more on how positive duties yield performance obligations see also Timmermann 2006, 302).

So long as it can derive positive duties to set ends, then, FUL could serve as a plausible supreme principle of morality and survive the objection. In the remainder of this paper, I will therefore focus on the duties to set ends. I will not consider whether FUL could be used to derive positive duties of performance (on this topic see O’Neill 2013, 162–71), nor whether Kant himself would have wanted to use FUL for that purpose.

1. **The CW-Test and Maxim Contradictories**

I’ve argued so far that FUL survives the objection from positive duties if it can yield duties of virtue, and in particular the duties to adopt one’s own perfection and others’ happiness as ends. Now as Kant explains in the *Groundwork*, FUL contains two distinct contradiction tests, corresponding to two distinct kinds of duties:

Some actions are such that their maxim cannot even be thought without contradiction as a universal law of nature; let alone that one could will that it should become such. In the case of others that inner impossibility is indeed not to be found, but it is still impossible to will that their maxim be elevated to the universality of a law of nature, because such a will would contradict itself. It is easy to see that the first conflicts with strict or narrower (unrelenting) duty, the second only with wider (meritorious) duty (G 4:424).

In other words, we can derive “strict” duties by considering whether a maxim could be conceived of as a universal law without contradiction (the “CC-test”) and we can derive “wide” duties by considering whether a maxim could be rationally willed as a universal law without contradiction (the “CW-test”). It is uncontroversial that the duties to set one’s own perfection and others’ happiness as ends are “wide” duties (e.g. G 4:423-424, MS 6:390-394; Timmons 2021, 92). Thus, if they are to be derived from FUL at all, they must be derived using the CW-test.

The prevailing method of deriving positive duties from the CW-test makes use of maxim contradictories. On this view, whenever the CW-test rejects a maxim, we have a duty to adopt its contradictory (Herman 1993, 63; Timmermann 2007, 85; Guyer 2007, 143). Thus, if, as Kant argues in the *Groundwork*, we cannot rationally will that a principle of willfully neglecting our own development become a universal law (4:423), then it follows that we must adopt the opposite attitude, i.e., of sometimes promoting our own development.

As Wood has argued, this strategy does not work (Wood 1999, 100–102). If adopting a maxim of principled slothfulness is forbidden, this means that we may not on principle refuse to perfect ourselves. It does not follow from this that self-perfection is obligatory, nor that we must adopt an affirmative principle of self-perfection. After all, refraining from a principled refusal to engage in self-perfection is not the same as adopting a principled commitment to self-perfection: it leaves open the possibility of adopting no particular attitude at all with respect to self-perfection. In other words, the maxim of self-perfection could be morally optional even if its contradictory—the maxim of slothfulness—were impermissible. Likewise, beneficence could be morally optional even if non-beneficence were shown to be impermissible. Thus, ruling out non-beneficence and slothfulness by means of FUL is not sufficient to show that we have affirmative duties of beneficence and self-perfection.

Kant himself recognized the moral difference between agents who *refuse on principle* to f and those who *merely fail* to f. In the *Metaphysics of Morals*, he distinguishes between a mere failure to live up to one’s positive duties and making it one’s “principle not to comply with” those duties (MS 6:390). The former evinces only a lack of moral worth, but in the latter case “intentional transgression has become a principle” and this is what is “properly called a vice” (6:390). One would therefore expect that FUL could be properly applied without conflating a refusal to do something with a mere failure to do it.

To show that we are required to f by rejecting a maxim-contradictory we would need to show not only that a maxim of principled refusal fails the CW-test, but also that a maxim of merely-failing-to-f fails that test. If we cannot do so, then the CW-test will not have ruled out a failure to adopt any attitude with respect to f, and we will not be warranted in concluding that there is an affirmative requirement to adopt such an attitude. It is impossible to test a maxim of merely-failing-to-f, however, for there could not be such a maxim. Attempting to adopt a maxim of merely-failing-to-f is like attempting to commit an unintentional crime: it cannot be done. The very distinction between merely failing to do something and refusing to do it is that only the latter is grounded in a maxim (MS 6:390). Accordingly, we cannot derive positive duties from the CW-test by showing that maxims describing the contradictory attitude fail the test.[[3]](#endnote-3)

1. **Deriving the Duties of Virtue**

I propose a different way of deriving positive duties. Although we can apply the CW-test by considering whether we could rationally will that a particular maxim become a universal law, this is not the test’s only possible application. We can also consider whether we could rationally *fail* to will that a particular maxim become a universal law. In other words: just as we can ask whether one could rationally will to live in a world where a particular maxim is a universal law, so too can we ask whether it could be consistent with rationality *not* to choose to live in that world. So understood, we can use the CW-test to derive positive duties by directly evaluating a positive maxim, that is, the maxim of adopting the attitude we believe to be required. Because this solution does not require first evaluating a contradictory attitude, it does not suffer from the problem discussed above.

As we’ve seen, Kant’s two central positive duties are duties to adopt ends. To show that we are morally required to set an end using the CW-test, we must show that we could not rationally fail to will that end as a universalend.[[4]](#endnote-4) I argue that we can determine whether we could rationally fail to will an end as a universal end by considering the normative standard of rationality Thomas Hill has dubbed *the* hypothetical imperative:

**The Hypothetical Imperative:** If a person wills an end, and certain means are necessary to achieve that end and are within his power, then he ought to will those means (Hill 1973, 431).

This principle plays a critical role in Kant’s understanding of rational agency (e.g. G 4:417, 4:419; Hill 1973) and therefore also in any application of the CW-test. After all, the CW-test asks us to determine whether we could rationally will some maxim as a universal law; any answer to that test necessarily relies on a standard of rational willing. I take it that standard is the hypothetical imperative.

In the context of the CW-test, we are effectively asked to imagine that we have the power to prescribe universal ends for everyone. Thus we can dispense with the ordinary requirement that the means are known and within our power (see e.g. Bratman 2009, 143). In a thought-experiment, all means are known and within our power. For purposes of the CW-test, we can state more simply:

**Simplified Hypothetical Imperative:** For any end *x*, if *y* is a necessary means to *x*, willing *x* rationally requires eitherwilling *y* or giving up on *x*.[[5]](#endnote-5)

It follows from this that if I maintain *x* as my end, a failure to will *y* would reveal an internal inconsistency, i.e., a contradiction in my will. Importantly, then, our maintaining *x* as an end implies more than a rational requirement to refrain from willfully refusing to do *y*: it requires affirmatively setting *y* as an end.Moreover, this requirement is transitive, such that if *z* were a necessary means to *y*,then rationality would require that we either will *z*, as well, or give up *x* as an end.To illustrate, assume that I am thirsty and set myself the goal of quenching my thirst. Drinking something is a necessary means to that end. It is not sufficient for me simply to refrain from *refusing* to drink. Rationality requires that if I am to adopt the end of quenching my thirst, I must *affirmatively adopt* the end of drinking something. And if we assume that some further steps are required to the end of drinking something, I must take those steps as well—or alternatively amend or give up my end.

 We can now consider the required ends of self-perfection and others’ happiness. Begin with self-perfection. Existing interpretations have rightly emphasized that our continued agency—our capacity to set and achieve ends—is a necessary means to any end whatever (Herman 1993, 52–62; Allison 2011, 189–90). We can only achieve any particular end if we are capable of achieving ends in general; therefore our continued agency is a *necessary end* (Herman 1993, 52–62; Allison 2011, 189–90; cf. Wood 1999, 91). Ordinarily the concept of a necessary end is used to derive the familiar negative version of the duty—prohibiting the vice of slothfulness (Allison 2011, 189–90; Herman 1993, 52–62). If we adopt the goal of principled slothfulness and neglect all of our talents, we will ultimately be incapable of achieving any of our ends, including our necessary ends. Therefore, we cannot will a discretionary end without also willing that we develop our own capacities at least to some minimal degree. But so long as we are agents, we will always will *some* discretionary ends. So, we cannot will to be an agent and simultaneously adopt the end of self-neglect. *A fortiori* we cannot will that everyone adopts that end.

 This approach is persuasive, and it may well represent the best reading of Kant’s *Groundwork* arguments. But it can be improved to directly show that we have an affirmative duty to adopt self-perfection as an end. Self-preservation and some degree of self-cultivation are necessary means to the necessary end of our continued agency. Moreover, as Herman and Allison show, our continued agency is itself a necessary means to any end whatever (Allison 2011, 189–90; Herman 1993, 52–62). Applying the hypothetical imperative, we see that every instance of end-setting separately commits us to the continuation of our agency, and in turn to our self-preservation and (some) self-cultivation. So long as we will anything, a failure to will self-preservation and some self-cultivation will result in a contradiction in our will. Accordingly, the CW-test affirmatively requires that we adopt the ends of self-preservation and (some minimal) self-cultivation.

 This argument is not yet satisfactory, for we are after a duty of self-perfection, not self-preservation. The duty of self-perfection requires that we set the promotion of our own rational agency as our end (MS 6:392-393). To establish that duty, we can again consider whether we could rationally fail to set (the promotion of) our own rational agency as an end. We cannot. As it seems to me Kant realized when he wrote the *Metaphysics of Morals*, understanding the necessity of this end involves considering our rational agency over time. Kant points out that we do not know which ends we will later adopt—we know neither what we might wish for in the future, nor what ends circumstances might force upon us. Accordingly, we must “cultivate [our] powers of mind and body so that they are fit to realize any ends you might encounter, however uncertain you are which of them could sometime become yours” (MS 6:392). In other words: we cannot rationally fail to will our own self-perfection, because that would necessarily entail our own failure to achieve at least some of our ends. So long as we are agents, we remain necessarily committed to the achievement of our ends, *including our future ends*. Hence we must will our own self-perfection.

 It should be clear that this is not simply a repetition of the argument for self-preservation. There, we saw that whenever we set any particular end, we commit ourselves to the continuation of our rational agency as a necessary means to that end. We can go one step further: insofar as we are agents, we are committed generally to the achievement of our ends, and that commitment extends not just to any end in particular but to *all* of our ends, including those we will adopt in the future. In light of that commitment, we cannot rationally fail to set our own rational nature as an end. Put more precisely: all rational agents share the second-order end that their first-order ends be accomplished. All agents are therefore rationally required to will their own self-perfection, for that is a necessary means to the universal second-order end. Accordingly, we could not rationally fail to will our own self-perfection.

It is important to notice that this conclusion follows from FUL’s CW-test as I have construed it. After all, it is not controversial that Kant’s ethics has the resources to derive duties to adopt ends. (For instance, Herman has argued along similar lines that the required end of one’s own perfection can be derived from a necessary end to “will correctly” (2007, 263).) At issue in the objection from positive duties is whether the *Formula of Universal Law* can be used to derive such duties. If, as I have argued, the CW-test requires us to will those ends that we could not rationally fail to will, then it can independently ground the affirmative duty of self-perfection.

We can arrive at the positive duty to set others’ happiness as our end in much the same way. As Kant points out in the *Groundwork*,it would be irrational to choose to live in a world of non-beneficence, because we could not accomplish our ends in that world (4:423). Interestingly, it is likely that we could not pursue even the end *of independence from others* wholly unaided, as Herman and O’Neill have argued (Herman 1993, 53–55; O’Neill 1989, 133–35). In any event, none would make it past infancy in a world of total non-beneficence. Now consider whether we could rationally *fail* to choose to live in a *beneficent* world, that is, whether we could rationally fail to will beneficence as a universal end. We cannot. It is just as irrational to choose to live in a world where everyone is non-beneficent as it is to fail to choose to live in a world where everyone *is* beneficent. Like our own perfection, the beneficence of others is a necessary means to at least some of our ends; therefore, it is a necessary means to the universal second-order end that our first-order ends be accomplished. In failing to choose to live in the beneficent world we are in effect failing to will that (at least some of) our ends be accomplished. As we have just seen, that is not something we can consistently will. Therefore, we cannot rationally fail to will that everyone adopt the end of others’ happiness. It follows that FUL’s CW-test affirmatively requires us to adopt that end.

Interestingly, both of these arguments are rooted in Kant’s assertion that all rational beings necessarily will their own happiness (G 4:415). Some commentators argue that because happiness appears to be the only such universal end Kant recognizes, only the required end of others’ happiness could possibly be grounded on FUL (e.g. Wood 1999, 102). But Kant defines “happiness” as “the state of a rational being in the world in the whole of whose existence *everything goes according to his wish and will*” (KpV 5:124, Kant’s emphasis). Thus, the universal end of happiness itself implies the second-order end that our first-order ends be accomplished. As we have seen, we can rely on that second-order end to derive both of Kant’s positive duties.

Having derived the two central positive duties in Kantian ethics from FUL, it will be useful briefly to return to the nature of those duties. If the ends of self-perfection and beneficence are required, and if rationality demands that we pursue the necessary means to these ends (since we may not abandon them), then it might seem that Kant’s system of positive duties devolves into an ethics of maximization—one according to which we must act so as to maximize our own perfection and the happiness of others. Such a view would come dangerously close to turning Kant into a utilitarian. But it does not follow from what I have argued here.

Beneficence and self-perfection are required pursuits, but they need not be our only ones, and nothing Kant says about positive duties suggests otherwise. True, we may not abandon necessary means to obligatory ends. Thus, for instance, our duty to self-perfection entails a duty of basic self-preservation, since the latter is a necessary means to the former (see MS 6:421). Determining precisely which means are minimally necessary to the ends of beneficence and self-perfection is the difficult question Kant takes up in his *Doctrine of Elements*. We need not answer it, however, to see that a rational requirement to take the *necessary* means to our obligatory ends differs from a requirement to take *each of the optimal* means to those ends. Only the latter requirement could amount to an ethics of maximization. Kantian virtue demands that our conduct be consistent with the required ends, and that we adopt a genuine commitment to furthering them. As Kant recognized, however, it is possible to adopt a genuine commitment to an end without devoting one’s life to maximizing it. How and when we pursue the ends of virtue will often be a matter of judgment—one which Kant says can leave significant room for latitude (e.g. MS 6:390; 6:393, 6:411).

Kant’s Formula of Universal Law therefore yields nuanced and attractive duties of virtue. Because they are duties to set ends—not to perform actions—they are context-sensitive and apply differently to different people. And because we are rationally required only to pursue the necessary means to our ends, Kant’s positive duties require consistent commitment to virtue while leaving “playroom” as to how that commitment is expressed (MS 6:393).

The CW-test, then, can be used to derive the central positive duties in Kant’s ethics, and we must conclude that the objection from positive duties fails. As I have tried to show, contradictions in the will can arise not only when we irrationally will an end, but also when we irrationally *fail to will* an end. Hence we can use the CW-test to determine not only whether an end is forbidden, but also whether an end is obligatory.[[6]](#endnote-6)

1. **Conclusion**

FUL is often characterized as the weaker, less useful sibling of the other formulations of Kant’s Categorical Imperative. The objection from positive duties is one important reason for FUL’s disfavored status. This paper has aimed to show that the objection from positive duties fails. Just as Kant claims, FUL can ground our central affirmative duties of virtue.

Guus Duindam, *University of Michigan*

435 S. State St., Ann Arbor, MI, 48109

gduindam@umich.edu

**References**

Allison, Henry E. 2011. *Kant’s Groundwork for the Metaphysics of Morals*. Oxford: Oxford University Press.

Bratman, Michael E. 2009. “Intention, Practical Rationality, and Self-Governance.” *Ethics* 119 (3).

Guyer, Paul. 2007. *Kant’s Groundwork for the Metaphysics of Morals: A Reader’s Guide*. London: Continuum.

Herman, Barbara. 1993. *The Practice of Moral Judgment*. Cambridge, Mass.: Harvard University Press.

———. 2007. *Moral Literacy*. Cambridge, Mass.: Harvard University Press.

Hill, Thomas E. 1973. “The Hypothetical Imperative.” *The Philosophical Review* 10 (1): 429–50.

Kahn, Samuel. 2014. “Can Positive Duties Be Derived from Kant’s Formula of Universal Law?” *Kantian Review* 19 (01): 93–108.

———. 2022. “Nary an Obligatory Maxim from Kant’s Universalizability Tests.” *Zeitschrift Für Ethik Und Moralphilosophie* 5 (1): 15–35. https://doi.org/10.1007/s42048-022-00115-0.

Marcus, Ruth Barcan. 1980. “Moral Dilemmas and Consistency.” *The Journal of Philosophy* 77 (3): 121–36.

O’Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. Cambridge: Cambridge University Press.

———. 2013. *Acting on Principle*. 2nd ed. Cambridge: Cambridge University Press.

Rippon, Simon. 2014. “Were Kant’s Hypothetical Imperatives Wide-Scope Oughts?” *Australasian Journal of Philosophy* 92 (4): 783–88.

Timmermann, Jens. 2006. “Kant on Conscience, ‘Indirect’ Duty, and Moral Error.” *International Philosophical Quarterly* 46 (3): 293–308.

———. 2007. *Kant’s Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge: Cambridge University Press.

———. 2013. “Kantian Dilemmas? Moral Conflict in Kant’s Ethical Theory.” *Archiv Fur Geschichte Der Philosophie* 95 (1): 36–64.

Timmons, Mark. 2021. *Kant’s Doctrine of Virtue*. Oxford: Oxford University Press.

Wood, Allen. 1999. *Kant’s Ethical Thought*. Cambridge: Press, Cambridge University.

Worsnip, Alex. 2015. “Narrow-Scoping for Wide-Scopers.” *Synthese* 192 (8): 2617–46.

**Notes**

1. In the *Doctrine of Elements*, Kant uses these duties to derive secondary positive and negative duties. The success of FUL does not hinge on those secondary arguments, so I set them aside here. [↑](#endnote-ref-1)
2. Although Kant is sometimes saddled with the view that we must always keep our promises (e.g. Marcus 1980, 123–24), he argues only for a negative duty against making lying promises (G 4:424). [↑](#endnote-ref-2)
3. Even if we *could* derive positive duties using maxim-contradictories, those duties would still be too weak: the requirement to set others’ happiness as an end involves far more than simply ‘sometimes helping others’ (Guyer 2007, 143). [↑](#endnote-ref-3)
4. The means to be used to achieve the ends are not relevant for purposes of these derivations. Duties of virtue concern only the moral status of ends, and they are derived by testing “maxim[s] *of ends*” (MS 6:395). [↑](#endnote-ref-4)
5. I take this to be a ‘wide-scope’ reading of the hypothetical imperative. For a defense of this reading, *see* Rippon 2014. For discussion of objections to this standard of rationality, *see* Worsnip 2015*.* [↑](#endnote-ref-5)
6. I am grateful to Robert Johnson, Sarah Buss, Allen Wood, Mariam Kazanjian, and the anonymous reviewer for this Journal for their comments on earlier versions of this paper. [↑](#endnote-ref-6)