

Lyle Crawford

What If Everybody Did That? Kant's Test Of The Universalized Maxim

In Section I of the *Groundwork of the Metaphysics of Morals*, Kant argues that actions have moral worth if and only if they precede from "respect for the moral law" (4:400)¹, or the rational apprehension of duty. Determining the *form* of the moral law, whether or not it is ultimately binding for rational beings as such, is his main project in Section II. Our actions are acts of will, but the basis of any action is the "representation of laws, that is, in accordance with principles" (4:412) as causally determining their outcome, and this is the sphere of "practical reason." In order to derive the supreme moral law, Kant works out a theory of "imperatives" governing practical reason. An imperative is a command we must obey insofar as we are rational, and may be either "hypothetical" or "categorical." A hypothetical imperative commands us to do something "analytically." What we ought to do can be determined by analysis of some goal or end which we do in fact will. If analysis of this end in context reveals that it can be achieved only by some particular means, then it violates practical reason not to will the means. Though a hypothetical imperative is rationally binding, whether the imperative obtains at all is contingent upon the willing of some end, and we may not will the end even if we still want it, in the sense of knowing that we would like it, or merely hoping that it comes about. Many actions are governed by hypothetical imperatives, but Kant thinks that the supreme principle of morality is categorical. As a categorical imperative, this principle is not contingent upon or derived from the willing of an end; it is an unconditional "ought" binding for every rational being as such.

The sole moral principle which does not derive its "ought" from outside itself, according to Kant, runs as follows: "*act only in accordance with that maxim through which you can at the same time will that it become a universal law,*" or "*act as if the maxim of your action were to become by your will a universal law of nature*" (4:421). In Kant's subsequent explanation of this principle, and his examples of its application, it becomes clear that the maxims ruled out by this test are those which entail some kind of *contradiction*. The contradiction may be one of two kinds, usually called "contradiction in conception" and "contradiction in the will," which govern, respectively, "perfect" (enforceable) and "imperfect" (virtuous but unenforceable) duties.

Questions about how the contradiction tests are to be understood have received a great deal of scholarly attention. One recent commentator, Christine Korsgaard, has devoted special attention to an overview of the major interpretations, and has argued persuasively for one of them. Korsgaard points out that the issue does not appear to be decidable merely by an examination of Kant's wording: "language supporting all [three interpretations] can be found in Kant's texts, and it seems possible that he was not aware of the differences among them" (80)². If there is a single coherent account of the tests, it must be thought out carefully. The contradiction entailed by the universalization of an immoral maxim is generally interpreted in one of three ways, Korsgaard says: logical, teleological or practical. As we will see, Korsgaard finds deep problems with the first two interpretations and argues very strongly for the Practical Contradiction reading of the test. However, this reading of the supreme moral principle invites a significant objection, namely, that the categorical imperative is not really categorical at all, but is rather a hypothetical imperative of the last kind identified by Kant himself -- one of "prudence"³ (4:416).

Proponents of the Logical Contradiction interpretation of Kant's test claim that an immoral maxim entails "something very like a logical or physical impossibility" (81) when universalized. Indeed, with the so-called contradiction in conception test, Kant strongly suggests that he is thinking of a logical contradiction, although the contradiction in the will test, as formulated, explicitly denies that any "inner impossibility" (4:424) is to be found in some immoral actions. Of the four examples following the primary statements of the categorical imperative, the case of the lying promise is the best contender for this interpretation. If the maxim of the lying promise were universalized, says Kant, "it would make the promise and the end one might have in it itself impossible" (4:422). The contradiction will be a logical one if this means, as it is usually taken to mean when Kant's term "impossible" is stressed, that in a world in which this maxim were universal, there could be *no such practice as promising*, since no one would even consider whether the words "I promise. . ." actually guaranteed anything.

Korsgaard develops an objection to this interpretation based on a distinction between "natural" and "conventional" actions. Conventional actions are based on "practices," and the Logical Contradiction view, as illustrated in the lying promise example, seems to account well for immoral conventional actions: "a practice has a standard purpose, and if its rules are universally violated it ceases to be efficacious for this purpose, and so ceases to exist" (85). The problem arises when we try to account for immoral natural actions, that is, actions whose possibility does not depend on a practice, but only on laws of nature. Acts of violence are actions of this kind. The universalization of the maxims of many immoral natural actions is quite conceivable, since nothing upon which the actions depend is thereby "ushered off the scene" (84). Korsgaard cites an attempt to account for the wrongfulness of this kind of action on the Logical Contradiction interpretation by showing that it fails the contradiction in the will test. The problem is that Kant is clear that duties derived from this test are "wide (meritorious)," not having the same enforceability as "narrow" duties (4:424). On this account, the person who commits a wrongful act of violence, in failing to regard others as ends in themselves, will lack some proper virtues or attitudes. But as Korsgaard points out, we usually want to say more than this -- the violent criminal does something unjust, and the duty not to act in this way is enforced, not just encouraged. If the contradiction in conception test is to account for immoral natural actions, it cannot be with a logical contradiction.

The Teleological Contradiction interpretation, like the Logical Contradiction view, finds numerous supporting or corroborating passages in Kant's texts. Korsgaard introduces the view as the attempt to derive a contradiction from immoral actions using Kant's second formula, the "Law of Nature" formula cited above, and she summarizes it in the following way: "we find some way to assign natural purposes to various instincts and types of actions and then find the contradiction when universalized maxims involve uses of those instincts and actions that defeat the natural purpose" (88).

By focussing on "purposes" instead of "practices," this view looks as though it will be able to account for both conventional and natural immoral acts. In the case of the conventional action of the lying promise, we "assign" to promising the natural purpose of "establish[ing] trust and confidence and the cooperation which they make possible" (89). The universalized maxim of the lying promise entails a teleological contradiction because promising becomes a poor means of achieving its own natural purpose. (Of course, this is something of an understatement if universalized lying would obliterate the whole institution of promising.) Where the Teleological Contradiction view apparently has an advantage over the Logical Contradiction view is in the case of a natural action like suicide, Kant's first example. "Self-love" is supposed to have as its natural purpose or "destination"⁴... the furtherance of life" (4:422). Suppose a suicide is motivated to end his own life from self-love. If this maxim were universalised, it would result in a system of nature in which the instinct undermined its natural purpose. "It is then seen at once," says Kant, "that [this system of nature]... would contradict

itself and would therefore not subsist as nature" (4:422).

Korsgaard identifies a critical slip in this construal of Kant's contradiction test. If we return to Kant's original formulae, we see that the agent's will is invoked to establish a contradiction. The case against the natural action of suicide will unravel on this point. The "system of nature" may contain a teleological contradiction, but the telos need not be willed (by this agent) at all. One cannot will the universalization of the lying promise because one relies on, and therefore must will, the existence of the practice of promising, but the suicide is perfectly able to reject the natural purpose of self-love while ending his own life out of self-love- "neither his own purpose nor anything else commits him to the purpose [of the preservation of life]" (90). For Kant's argument to work, he needs an agent's own purpose to commit him to willing some natural purpose which is undermined by the universalization of his maxim. Indeed, Kant's own definition of "maxim" as "the subjective principle of volition" (4:401) seems to make an agent's own purpose central. But if an agent can will his own purpose without willing his action's natural purpose, then there is no contradiction of practical reason, and the maxim doesn't fail the contradiction test.

Korsgaard defends the Practical Contradiction interpretation of Kant's test. This view, she thinks, avoids the problems of the previous two interpretations by understanding the contradiction to be between an agent's *own purpose as willed in the maxim of action* and the conditions which would obtain in a world in which this maxim gave the "standard procedure" (92) for realizing this purpose. The test reveals actions which depend for their efficacy on their status as exceptions, in other words, actions which are "not fair." On this account, the contradiction is not something which cannot be conceived, but something which would not work.

The greatest advantage of this interpretation, according to Korsgaard, is that it is continuous with Kant's own distinctly practical conception of the contradiction "contained" in the violation of hypothetical imperatives. In the most important passage, Kant writes that "in the volition of an object as my effect, my causality as acting cause, that is, the use of means, is already thought" (4:417). To will is to regard oneself as a cause, or as the initiation of a causal chain whose other links are the means one can use to achieve some object. To will the object (the effect) without willing the means (the cause) is a practical contradiction, and is *no weaker* than a theoretical contradiction. Thus, an immoral maxim can still result in a full-blooded contradiction on the practical view. So, in the case of the lying promise, the agent wills the falsification of the hypothetical imperative on which he relies in the same act of will that he wills his own particular maxim, thus thwarting his own maxim's purpose. Herein lies the contradiction, since "in the world of the universalized maxim, the *hypothetical* imperative from which the false promiser constructs his maxim is no longer true" (94). With the Practical Contradiction view worked out, we can see that no special account of the contradiction in the will test is required for it. We need only identify some purpose essential to the will, such as "general effectiveness in the pursuit of its ends" (96), and maxims which fail this test, violating "wide duties," can be understood as those which would undermine this basic purpose when universalized.

Korsgaard is clear that the Practical Contradiction view, by itself, has trouble with some kinds of natural actions, namely ones for which there is no particular purpose other than the expression of an emotional state. However, this view is nowhere near as limited as the Logical Contradiction view because, as we have seen, it takes into account an agent's purpose, and most natural actions do have an exterior purpose which would clearly be thwarted if the maxim were universalized. To see how this is so will usually require only a basic analysis of the purpose. So, for example, universalized stealing for the purpose of possession would contradict itself because the purpose "to possess a thing" amounts to something like "to be secure in the possession of this thing," which would be impossible in a world where one must expect to be a victim of theft oneself. Similarly, committing an act of violence for the purpose of gaining some advantage would contradict itself

because the advantage gained would in turn make oneself a target of violence, which would destroy basic conditions, such as being healthy or alive, of enjoying a given advantage.

I am persuaded that the Practical Contradiction reading of Kant's test is the most consistent, and makes the most sense out of Kant's own examples. If it is right, though, I think that it forces us to look at Kant's supreme moral principle quite differently. While my suspicion has been held by others⁵ with no reference to the Practical Contradiction interpretation as identified and defended by Korsgaard, this reading seems to make the objection even more plausible.

Some of Kant's language in the examples used to illustrate the categorical imperative provides particularly good support for the Practical Contradiction view. He argues for the impossibility of willing a maxim of failing to provide assistance to those in need. It is said to be impossible "since many cases could occur in which one would need the love and sympathy of others and in which, by such a law of nature arisen from his own will, he would rob himself of all hope of the assistance he wishes for himself" (4:423). But if an agent does not want the assistance of others -- if he is confident of his natural ability to defend himself and get what he wants, for instance -- then the maxim need not thwart its own purpose and embroil the agent in contradiction. Odder still than this apparent loop-hole is the explicit appeal to the agent's (contingent) self-interest to motivate moral action, an appeal made in hypothetical, not categorical, imperatives. If Kant's independent argument for moral motivation already seems oddly incomplete, the wording of this example only raises further doubts about how the categorical imperative is really to motivate a rational agent.

Korsgaard is very clear that for the Practical Contradiction view to work, we must carefully identify a maxim's purpose *for the agent* and include this in the maxim to be universalized. Her arguments show quite convincingly that this is indeed what we need to do, but in the process they reveal -- and this is highlighted in the above example -- that we need to insert a hypothetical imperative of prudence into the equation. *If* we are concerned for our own well-being, *then* we must act in such a way that, on an assumption of general reciprocity, will ensure our protection. While this makes sense of Kant's text, it also seems to undermine his account of the moral worth of actions. If the categorical imperative relies on a hypothetical imperative to indicate an adequate test for our maxims, is it really categorical? It does not seem that an agent can be said to undertake or abstain from some action out of the purely rational apprehension of duty as determined by the categorical imperative if this determination must invoke our self-interest. Korsgaard presents as a virtue of her position the continuity of the Practical Contradiction interpretation with the practical nature of hypothetical imperatives. However, by emphasizing the crucial role an agent's own interests play in the formulation of Kant's test, she has inadvertently underscored a weakness in Kant's description of how the categorical imperative is to function. Officially, Kant cannot intend the practical consequences of universalized immoral maxims to motivate moral action, since his account of moral motivation must coincide with his view of the moral worth of actions. But the language he is forced to use in arguing his example undermines the selfless rationality he needs for genuinely moral action, and this is compounded by Korsgaard's convincing defence of the Practical Contradiction interpretation.

Notes

1 *Groundwork of the Metaphysics of Morals*. (Ed. Mary Gregor). Cambridge University Press, 1997.

2 *Creating the Kingdom of Ends*. Cambridge University Press, 1996.

3 Klugheit. The German also suggests "cleverness" or "shrewdness."

4 Bestimmung. Throughout the *Groundwork*, the noun usually means "determination," and the verb,

bestimmen, "to determine."

5 A version of this criticism was first offered by Schopenhauer in his essay *On the Basis of Morality*. His most famous objection in the long section on Kant is that his hero's ethics is essentially theological: "Duty is the necessity of an action out of respect for the law,' would... read in natural and undisguised language, "Duty signifies an action which ought to be done out of obedience to the law" (SSS6). He goes on, however, to expose "the concealed hypothetical nature [of Kant's supreme principle], by virtue of which it is even based on pure egoism, and this is the secret interpreter of the instruction given in the principle" (SSS7).