

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications - Department of
Philosophy

Philosophy, Department of

1993

What Is Wrong with Kant's Four Examples

Nelson T. Potter Jr.

University of Nebraska - Lincoln, kjohnson6@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/philosfacpub>



Part of the [Continental Philosophy Commons](#), and the [Ethics and Political Philosophy Commons](#)

Potter, Nelson T. Jr., "What Is Wrong with Kant's Four Examples" (1993). *Faculty Publications - Department of Philosophy*. 20.

<https://digitalcommons.unl.edu/philosfacpub/20>

This Article is brought to you for free and open access by the Philosophy, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications - Department of Philosophy by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

What Is Wrong with Kant's Four Examples

Nelson Potter

University of Nebraska-Lincoln

Abstract

Kant gives four examples to illustrate the application of the categorical imperative immediately after introducing its "universal law" formulation in Chapter Two his *Groundwork of the Metaphysics of Morals*. These examples have been much discussed to gain an understanding of how the categorical imperative applies to derive specific duties. It is argued that the discussions found in these examples do not accord well with Kant's fuller account of that application in his later work *The Metaphysics of Morals*. That [later] work has quite different, sometimes better, arguments for the same moral conclusions, and never mentions the argument against making a lying promise (the second example). Giving exclusive or excessive attention to these four examples has distorted our understanding of Kant's moral philosophy.

I

For a good many years there has been a great concentration on the four examples of the application of the categorical imperative that Kant gives us in the *Grundlegung zur Metaphysik der Sitten*, immediately after he introduces the "universal law" formulation of the categorical imperative in Chapter Two. Perhaps it all started with C. D. Broad's discussion of Kant in *Five Types of Ethical Theory*.¹ When Robert Paul Wolff's well known anthology of writings on Kant first appeared some 25 years ago, five of the eight essays there on Kant's ethics were at least in part concerned with giving an explication or critique of these four examples.² Many textbook discussions of Kant center around Kant's arguments for the four examples (e.g., Hospers, Feldman); as a result whole generations of students have been taught Kant based on what he shows or fails to show about the meaning and applicability of the categorical imperative in the famous four examples.³ In this literature friends of Kant have usually tried to defend his reasoning in the four examples, and those hostile to his general views have criticized it.

In this paper I wish to take a different approach, and, although I will mention a few points of criticism, mainly I wish to urge that these famous few paragraphs do not very well state or reflect Kant's more considered

and detailed discussions of these same issues in other ethical works of his critical period, especially *The Metaphysics of Morals*. Thus, they are not a very reliable guide to Kant's views on the application of the categorical imperative. After I discuss each of the four examples as presented after the "universal law" formulation, I will also discuss (1) the reprise of these examples after the second formulation, concerning respect of persons (G, IV, 427-430).⁴ This coverage of the same four examples has been much less discussed than the presentation of these examples after the first formulation (G, IV, 421-3). Finally, (2) I will discuss what Kant has to say about similar duties in the *Doctrine of Virtue* (*Tugendlehre*) part of the *Metaphysics of Morals*. I conclude that we should give the famous four examples in their first occurrence, just after the first formulation of the categorical imperative, a much needed rest, and learn to look elsewhere in addition to, and sometimes in place of, these examples to discern Kant's views on the application of the categorical imperative.

There is another line of criticism of the emphasis on the four examples that I believe is correct, but which I will not try to defend here. It is as follows: If we take the four examples discussion as quite important or central to Kant's moral philosophy, we are typically at least making the assumption that the examples are the endpoint and goal of Kant's introduction and presentation of the categorical imperative in the *Grundlegung*, i.e., that Kant's moral philosophy is in the first instance a deontological theory of morally right or wrong action, perhaps after the model of W. D. Ross or H. A. Prichard. This is the interpretation of Kant as a pure deontological formalist. It is this line of interpretation that makes many practitioners of virtue ethics view Kant as an enemy. Again, I do believe that this common line of interpretation is incorrect, but I will not be directly criticizing it here. Rather, insofar as this interpretation is based on an assumption of the centrality of the four examples, I will be indirectly undermining it.

My procedure will be next, in II, to make some general remarks about the argumentative structure and intentions of the *Grundlegung* as a whole. Then I will discuss the four examples in the order 1, 3, 2, 4, in order to discuss the two examples of duties to oneself together (1 and 3 in III, followed by 2 and 4 in IV). For each discussion, after presenting and commenting on Kant's initial presentation after the universal law formulation, I will contrast it with both the reprise of the example following the second formulation, and the fuller discussion of the example in the *Tugendlehre*, when there is one. Finally, in Part V I will draw some general conclusions from these discussions.

II

In this section I wish to make two general points about the *Grundlegung* as a whole. The first relates to Kant's general intention in this small book,

the first work on ethics from Kant's critical period. In the introduction Kant, after telling us that he intends to devote a later work to the application of the categorical imperative, a work to be entitled *Metaphysics of Morals* (a work actually published in 1798), writes as follows about the limited purpose of the present work:

The sole aim of the present *Groundwork* is to seek out and establish *the supreme principle of morality*. This by itself is a business which by its very purpose constitutes a whole and has to be separated off from every other enquiry. The application of the principle to the whole system would no doubt throw much light on my answers to this central question....All the same, I had to forego this advantage, which in any case would be more flattering to myself than to others, since the convenience of a principle in use and its seeming adequacy afford no completely safe proof of its correctness. (G, IV, 392)

After this statement it might seem surprising that Kant even included his four examples, for this statement suggests that the *Grundlegung* will contain no information about application. And yet it makes sense to include at least a few examples of applications. Otherwise it may seem that the categorical imperative, especially in its first "universal law" formulation, might have no applications at all, and it might seem unspeakably mysterious and abstract without at least a few suggestions about its application. So I suggest that it was as a result of a sort of compromise with the common understanding that the examples were there at all. They are meant to be just what Kant says they are: examples, intended to give merely a brief and preliminary indication of how such applications will go, and, as Duncan argued, they provide a kind of interlude to the main argument of the *Grundlegung*. Hence, Kant has himself given us an indication that the four examples are not to be taken too seriously, and he has told us that the place to look, if one wishes to understand the application of the categorical imperative is to the planned *Metaphysics of Morals*.

The second point I wish to make about the *Grundlegung* as a whole is that in a number of other respects it appears that when Kant wrote it he did not have his system of moral philosophy entirely worked out. There are some departures from the doctrine and terminology of the later works on moral philosophy, to a much greater degree than can be found among those other later works in relation to each other (I have in mind primarily the second *Critique*, the *Religion*, and the *Metaphysics of Morals*). Here are some instances to illustrate this: (1) Kant apparently almost immediately undertook to revise the doctrine and proof of freedom that was offered in Chapter Three of the *Grundlegung*, having quickly become dissatisfied with it. The new version appeared quickly in the second *Critique*.⁵ (2)

There are a number of doctrines and points of terminology that appear in the *Groundwork* that are never used later. The terminology of "objective ends" and "ends in themselves" is seldom if ever used in the later works, though this change of language may not reflect any doctrinal change. (3) The very doctrine of three formulations of the categorical imperative, formulations which Kant tells us are equivalent to each other, receives little to no confirmation in later works. (4) The terminology for "motives" or "incentives" changes, and the terminology of the *Grundlegung* is not adhered to later on. There are possible explanations available in some of these cases which are adequate to explain away any such shift. But when we put all of these differences together, I conclude that the *Grundlegung* is more different from Kant's later works in moral philosophy than those later works are from each other. The differences among the later works can often be accounted for simply in terms of the fact that each of them is taking up different topics. If Kant later revised some of his views, this is another reason for not taking the *Grundlegung* as the sole or even the most reliable or most important text in understanding Kant's views.

III

Now let us consider each of the examples in turn from G, IV, 421-3, bringing out the limitations and inadequacies of each of the examples, regarded as representative arguments for moral applications of what Kant takes to be the basic moral principle. I'm not primarily interested here in philosophical critiques of the arguments themselves, something which many others have attempted before me, but some criticisms will be made in the process of interpretation, and as an adjunct to interpretation. The "what is wrong" from my title that I wish to show, is not primarily philosophical inadequacy of the application arguments, but unhelpfulness for developing a general interpretation of Kant's views on the application of the categorical imperative.

We turn now to Examples One and Three, the two examples of duties to oneself.

(1) For the first example, where the agent is considering suicide, the proposed maxim is mentioned, "From self-love I make it my principle to shorten my life if its continuance threatens more evil than it promises pleasure." (G, V, 422) Kant tells us that this maxim is not a possible law of nature: "It is seen at once that a system of nature by whose law the very same feeling whose function (*Bestimmung*) is to stimulate the furtherance of life should actually destroy life would contradict itself and consequently could not subsist as a system of nature." (G, IV, 422)

This seems not a very satisfactory argument. For one thing, there doesn't seem to be any contradiction in a natural impulse that now pushes us to do one thing, and at another moment something quite different. The im-

pulse to avoid uncomfortable temperatures one time has me moving out of the hot sun, and later (next winter) covering myself with clothes to keep warm. Or the impulse of thirst, which at one moment moves me to seek out drink, once satiated moves me to refuse offers of drink.

Perhaps Kant's claim is more plausible if the law in question is regarded as a purely *teleological* law, one which mentions *the* purpose of a given organ, function, or ability. In a *teleological* system perhaps it would be correct that a given faculty or impulse could not have contrary purposive outcomes within that singular system, and still be such a law. The teleological law here would be something like: "The natural end of the feeling of self-love is the promotion of life in the agent."

Here in an interlude mostly confined to the present paragraph, I indulge in some philosophical critique of this view, a brief departure from the above expressed resolve to abstain. The claim that there are such teleological natural laws seems hard to defend. Such laws seem to mix together the functions of explanation and value justification. And even if we do posit such laws, there still seem to be problems with plurality of ends, or of candidate ends. For example, is the natural end of the sexual faculty (a) reproduction, (b) experience of sexual pleasure, (c) expression of sexual love, or is it some combination? In an ethical theory based on natural ends, different answers to this question will produce different recommendations for behavior. If even some of Kant's applications of the categorical imperative (Paton said all)⁶ require the use of teleological laws, that is an outcome more difficult to defend than the categorical imperative itself. Also, if such laws play an essential role in application arguments, they arguably preempt the universal law principle itself as the substantive moral principle that is the source of the specific obligation; the universal law requirement begins to seem like a fifth wheel. Finally, there seems little prospect of using arguments with teleological premises to arrive at uniquely correct moral conclusions. So if this were the correct way to understand Kant's views on the application of the categorical imperative, then this part of his theory can probably be described as of historical interest only.

However, when we turn to the *Metaphysics of Morals*, we see that the teleological version of the argument against suicide is entirely abandoned, and a different, and more defensible argument is presented:

Man cannot renounce his personality as long as he is a subject of duty, hence as long as he lives; and it is a contradiction that he should be authorized to withdraw from all obligation, that is, freely to act as if no authorization were needed for this action. To annihilate the subject of morality in one's own person is to root out of existence morality itself from the world, as far as one can, even though morality is an end in itself. Consequently, disposing of oneself as a mere means to some

discretionary end is debasing humanity in one's person (*homo noumenon*), to which man (*homo phenomenon*) was nevertheless entrusted for preservation. (MdS, VI, 422-3)

This argument can be elucidated with a partially analogous example. Suppose that Jones, who has been working hard to achieve a certain important political reform, and who has gathered many others to work with her on behalf of this cause, suddenly withdraws from the effort because she decides that she was tired of it. We sometimes criticize persons who do such things by calling them "quitters." The general point is that one doesn't have the moral right arbitrarily to terminate these ongoing life commitments for the sake of (morally) arbitrary fulfillment of personal pleasure or avoidance of pain. I think such an argument does not manage to show suicide wrong in every case, but it seems likely to be an objection at least to some such cases.

There are two points of disanalogy between the example just presented and Kant's argument that should be made explicit: (1) My example might seem to be a violation of a duty to others. I think Kant would reply that when we opt out of ongoing duties to others in such a way we also violate a duty to ourselves, just because such an opting out seems to reflect a lack of moral seriousness, and hence also has an unavoidable self-reference to one's own moral personality. (2) My example concerns the undertaking of an *optional* end. Kant's argument against suicide in contrast concerns the arbitrary termination of goals that were not optional, but obligatory, like any Kantian end that is also a duty. Presenting an argument for there being non-optional moral ends is too big an undertaking for this paper. Let me instead just make the point that a life that involved *no* ongoing moral projects, whether optional or obligatory, would be a moral life so minimal as to be incompatible with moral excellence.

This argument makes no reference to teleological laws, and also is entirely different from the structure of the *Grundlegung* argument. In fact there are no such teleological law arguments relating to duties to oneself in the *Tugendlehre* at all. Thus the *Grundlegung* discussion of suicide with which we began is not even very closely related to the later, more detailed discussion of the same topic, and the central argument against suicide is quite different. The big similarity, which might tend to obscure the differences between the arguments, is that in both texts Kant does conclude that suicide is morally wrong, but the main arguments are quite different.

In the *Tugendlehre* teleological language is not entirely absent, but the teleology is not natural teleology, but moral teleology. We are to abstain from suicide because continued life is an indispensable means to all the moral ends of humans, just as, as we will see below, talents are to be developed because after development they provide indispensable means to

certain of our moral ends. There is a reference in Kant's introduction to the topic of duties to oneself (see MdS, VI, 420) to "the impulses of nature having do with man's *animality*." The first of these is the one relevant to suicide, the impulse through whichnature aims at [the agent's] self-preservation..." (MdS, VI, 420) So the teleological view of human animal nature is retained, and is arguably used to set the stage for the final description of duties to oneself. But still the use of such teleological assumptions in specific application arguments does not occur in this later, fuller discussion in the *Metaphysics of Morals*.⁷

The *Tugendlehre* argument is more closely approximated by Kant's *Grundlegung* discussion of suicide in connection with the *second* formulation of the categorical imperative. There Kant writes,

If he does away with himself in order to escape from a painful situation, he is making use of a person merely as a *means* to maintain a tolerable state of affairs till the end of his life. But man is not a thing — not something to be used *merely* as a means: he must always in all his actions be regarded as an end in himself. Hence I cannot dispose of man in my person by maiming, spoiling, or killing. (G, V, 429)

However, notice that this argument is rather *incomplete*, since Kant does not further explain here how or why the action of suicide is contrary to humanity in our own person, as he does in the *Tugendlehre*.

There is also, it might be thought, a defect in Kant's reformulation of the moral imperative here: "Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end." (G, V, 429) If I ask someone at table please to pass the salt, I may seem to be violating this formulation, at least unless I simultaneously inquire how his life is, and whether there is anything I can do to help *him*, for failing such additional actions I may seem to be using him merely as a means to my having the salt. And yet asking someone to pass the salt is normally morally unobjectionable. It seems more likely that the second formulation is intended to rule out actions that violate the *rights* of another, or fail to harmonize with the wills of others, or, in the case under discussion, if I can speak this way, violate the rights of humanity in one's *own* person. There are serious questions about how we determine what actions would be ruled out or required of us; sometimes, at least arguably, they could be determined by an application of the *first* formulation of the categorical imperative, though that works better for duties to others than for duties to oneself. If we accept this interpretation of the second formulation, it does, so far as it goes, support Kant's claim of the equivalence of the first and second formulations. It would be when we act so as to violate someone's rights (including our

own) or fail to harmonize our action with the will of others and only then that by such an action we regard this human being as nothing more than a means to our ends, a utensil, as it were. Hence only such violative actions would be morally objectionable, under either formulation. Although Kant surely believed when he wrote the *Grundlegung* that the idea of humanity grounds rights, the specific nature of that connection remained unclear. Until the *Tugendlehre* structure of ends that are also duties is in place, Kant is not well positioned to explain how or why an act of suicide would violate any such right rational agents may have against themselves.

(3) Let us now turn to the third example, the obligation to develop talents. Here the first version of the *Grundlegung* argument is teleological in the same way as the discussion of suicide was. The maxim is one of neglecting talents for an agent who does not need to develop them to get by:

He then sees that a system of nature could indeed always subsist under such a natural law, although (like the South Sea Islanders) every man should let his talents rust and should be bent on devoting his life solely to idleness, indulgence, procreation, and, in a word, to enjoyment. Only he cannot possibly *will* that this should become a universal law of nature, or should be implanted in us as such a law by a natural instinct. For as a rational being, he necessarily wills that all his powers should be developed, since they serve him, and are given him, for all sorts of possible ends. (G, IV, 423)

The teleological law here is based on the idea that talents are “given” to us for our use, and that we are out of harmony with this purposive system if we fail to act so as to develop them. And this application reflects the distinction between contradiction in conception for perfect duties (including suicide) and contradiction in the will for imperfect duties. The above critical comments about the teleological interpretation making Kant’s theory of purely historical interest, also apply here.

When Kant discusses this example in connection with the second formulation he says that our actions must not just not conflict with the end of humanity in us; they must harmonize with this end. “Now there are in humanity capacities for great perfection which form part of nature’s purpose for humanity in our person. To neglect these can admittedly be compatible with the *maintenance* of humanity as an end in itself, but not with the *promotion* of this end.” (G, IV, 430) This is very close to the parallel discussion of talents that follows the first formulation, and is even more explicit in its reference to considerations of natural teleology. For this third example, unlike the other three, there is no improvement to be found in the discussion of the issue that comes after the second formulation.

In the *Tugendlehre*, as has been noted above, teleological language is not entirely absent, but the teleology found in moral application arguments is not *natural* teleology, where in the present instance the reference would be to the end of the talents themselves, or of the projected giver of the talents (e.g., God) in giving them to us, but to *moral* teleology. After saying talents are not just to be developed for the sake of personal advantage, Kant adds, "Instead, it is a command of morally practical reason and a duty of man to himself to cultivate his capacities (some among them more than others, insofar as men have different ends) and to be in a pragmatic respect a man equal to the end of his existence." (MdS, VI, 445) Talents are to be developed because developed talents are important means to the *moral* ends of human beings. For example, one may train to be a nurse with the moral aims of becoming self-supporting and helping others. There is a thorough analogy with the above discussion of suicide, except, of course, for the fact that the suicide duty is a negative duty to abstain, and the talents duty is a positive duty to develop. The comments made above concerning suicide can be repeated here: the arguments are quite different in *Tugendlehre*, and the *Grundlegung* argument is less defensible. Also, since they are quite different from each other the *Grundlegung* argument does not at all accord with Kant's *Tugendlehre* views.⁸

IV

(2) We turn to the lying promise argument, Kant's second example (G, IV, 422). This is an interesting argument, for a number of reasons. First, it seems like a good, defensible moral argument, that is very clearly based on the universal law principle, unlike those for the duties to oneself. It also has the advantage of pretty clearly illustrating what Kant means by the contradiction in conception test. In fact Kant at the time of writing the *Grundlegung* seems to have been enormously impressed with this argument, which for that reason we might think to have been a recent discovery. He not only gives us the example at G, IV, 422; he has already gone through the same argument for the same example, also in considerable detail, in Chapter One! (G, IV, 402) Toward the end of that chapter, he also arrives at the statement of the supreme principle of morality (as he does in Chapter Two at G, IV, 420-1)—it will be called the "categorical imperative" only after the requisite terminology has been introduced in Chapter Two—and then gives an almost equally elaborate development to the same lying promise example. In addition, he alludes to the lying promise argument at two other places where he is looking for a good application example: G, IV, 419 and 441.

This argument so impresses him that it appears that the "contradiction in conception" test is based on it primarily. (I leave aside perfect duties to oneself, also said to be cases of contradiction in conception because the

contradiction is more difficult to show there.) For among the other moral examples he gives in his works, only one other could be taken to illustrate such a test, perhaps: the case of the unreturned deposit (from *Critique of Practical Reason*, V, 27), and even that may not quite be such a case since no resistance to making deposits would develop, given that the promisee is deceased. In contrast resistance to accepting promises to repay would develop (in the second example argument) after the practice of making lying promises became a universal law, as people came to realize that promises were no longer reliable indicators of performance.

As wonderful, interesting and impressive as this argument is, it should be noted that it has quite a limited scope. Barbara Herman has argued quite persuasively in "Murder and Mayhem"⁹ that the moral obligations to refrain from murder and mayhem will have to be *imperfect* duties, according to Kant's criterion. Such obligations would usually be considered to be perfect duties, and so such an outcome seems surprising. If we accept some other common account of perfect duties such as those duties for which there are corresponding rights (J. S. Mill), or those duties which are so basic that if they are commonly violated, organized society ceases to be possible (H. L. A. Hart), and there is a return to nature, or those duties which are legally enforceable by criminal laws, or even those duties that are enforceable and hence fall within Kant's *Rechtslehre*, then the "contradiction in conception" test applies only to a limited subset of such duties, those which, as Herman says, involve some social convention, or as I would say, those duties which when violated involve acts that are parasitic upon others who are acting differently. What would be some other duties that have this character? Some likely instances are the duty to pay taxes, the duty not to steal, and perhaps the duty not to cheat in fair competitions.

What happens to such arguments for duties in the *Metaphysics of Morals*? They are no where to be found! Some such obligations, e.g., some promise-keeping obligations, no doubt get "absorbed" into the *Rechtslehre*, viz., those duties that are legally enforceable, such as the duties based on enforceable contracts (see MdS, VI, 220). But this is an example of a perfect duty to others, and such duties are not part of the *Tugendlehre*. And of course the account of such duties in the *Rechtslehre* is quite different. Enforcement there involves the social provision of non-moral incentives such as imprisonment or fines. There are no individual arguments to specific moral conclusions, except perhaps for the general argument for property, or the argument that capital punishment is the appropriate punishment for murder. Kant indicates that agents have a *moral* obligation to obey the law, but the parasitic character of violations is no part of Kant's argument for any duty in the *Rechtslehre*. In the *Tugendlehre*, lying, which is perhaps related to making a lying promise, is briefly discussed as a duty to *oneself*,

where again the argument is quite different. The passages are not entirely clear, but the parasitic character of violations does seem to be a part of Kant's discussion of lies in the late little essay, "On a Supposed Right to Tell Lies from Benevolent Motives." (See *Akademie*, VIII, 425-30)

Thus, although this "contradiction in conception" form of argument is interesting in its own right, it at best picks out only a limited subset of external, *Recht*-type obligations to others, leaving out murder and mayhem, for example, and probably also enforceable written contracts, and leaving only a narrow slice of nonenforceable duties to others: oral, legally unenforceable promises, small lies, and wrongful impositions that do not rise to the level of legal wrongs. This narrow class is not at all developed in Kant's later, fuller discussion of applications, the *Metaphysics of Morals*. So though it might have seemed to be a promising prelude to the development of Kant's later, most fully worked out theory of moral applications, this interesting argument disappears and is not even to be found in the *Metaphysics of Morals*.

When Kant discusses the lying promise case in the restatement of the four examples after the second formulation (G, V, 429-430) he argues that making a lying promise entails that the person lied to could not possibly share the promisor's purpose, and that hence a lying promise treats him merely as a means to the promisor's end. He then continues with some comments attempting to broaden the scope of this sort of argument:

This incompatibility with the principle of duty to others leaps to the eye more obviously when we bring in examples of attempts on the freedom and property of others. For then it is manifest that a violator of the rights of man intends to use the person of others merely as a means without taking into consideration that, as rational beings, they ought always at the same time to be rated as ends — that is, only as beings who must themselves be able to share in the end of the very same action. (G, IV, 430)

Kant here appears to be asserting that the kinds of action that Barbara Herman discusses as examples of murder and mayhem are violations of perfect duty to others, and he also makes explicit here the point that such actions involve the violation of the *rights* of others. It may seem, then, that the scope of perfect duties to others under the contradiction in conception test, which derives from the first formulation, is narrower than the second formulation test for such perfect duties. We have here again, as with duties to oneself discussed above, some evidence of non-equivalence between the first and second formulations. But since pursuit of this issue is aside from the main aim of this paper, we cannot develop that interesting question further now.

Let me make one other point about the lying promise argument. As I argued in detail in another essay¹⁰ the contradiction in conception has to do with the maxim, "To the end of getting money I need, I make a lying promise." The individual liar is successful in getting the desired money because of the trust of the loaner, a trust built on promises kept by people on other occasions in the loaner's experience. If everybody is making lying promises, the present lying promise will no longer be effective in getting money. So we have two contradictory law-like statements: In the individual case, the lying promise *causes* the loaner to hand over the money. In the case where everyone has been lying, the lying promise *does not cause* the loaner to hand over the money. This I argued is the contradiction Kant is talking about. If this is a correct interpretation, it means that the agent's maxim has internal to it a statement of a means-end relation, and that commitment to a maxim involves commitment to an end as well as a means. The point, that maxims include ends, is explicit in the *Tugendlehre* (see MdS, VI, 385), but not at all explicit in the *Grundlegung*. So even for the *Grundlegung*'s lying promise argument the later work must be referred to if we are fully to understand the argument for the most characteristic moral example of the *Grundlegung*.

(4) What about the obligation of mutual aid? (G, IV, 423) This is a good example of the contradiction in the will criterion. As stated, it seems not as broad as an obligation of universal beneficence, and so again the duty under discussion is somewhat limited. The fourth example gives us an important argument. It seems based on the factual claim that we are finite beings who have needs, and when we actually have such a need, we could not will that we not be helped. In contrast, the argument in the *Tugendlehre* for analogous imperfect duties to others is broader, fuller, and not limited to mutual aid, and there is still a different structure of argument, based on the other end that is the same time a duty, viz., the happiness of others. Further, the *Tugendlehre* discussion contains such new arguments as the idea that I am permitted to be benevolent to myself on the condition of my being benevolent to every other as well. (MdS, VI, 451)

The fourth example is the one that is closest to the counterpart *Tugendlehre* discussion. Barbara Herman's essay¹¹ is an authoritative, full interpretation of this powerful *Grundlegung* paragraph; it is based almost exclusively on the *Grundlegung*, and does not mention the characteristic *Tugendlehre* idea of the end that is also a duty. Without going further into a discussion of this duty in the *Tugendlehre*, it could be mentioned that this sort of duty is one of the best and clearest applications of the idea of ends that are also duties in that work.

The discussion of the fourth example after the second formulation (G, IV, 430) also helps connect the *Tugendlehre* and the *Grundlegung* discussions. Kant notes that "the natural end which all men seek is their own

happiness," and we have an obligation to agree "positively with humanity as an end in itself For the ends of a subject who is an end in himself must, if this conception is to have its *full* effect in me, be also, as far as possible, my ends." (G, IV, 430) Again this seems to indicate a broader duty than the same example discussed after the first formulation, and could even be read as seeming to hint at the *Tugendlehre* idea of the happiness of others as an end that is also a duty.

Usually the discussions of the four examples placed after the *second* formulation are much less attended to in the literature, but I conclude that in three out of four instances (Examples 1,2, and 4), those later discussions point us better towards and tell us more about Kant's later and fuller views on the application of the categorical imperative, the *Tugendlehre* doctrine.

There is also the following difference between the first discussion of the fourth example and the discussion of the *Tugendlehre*. The former seems to depend on a more specific point about human beings and their practical natures, viz., that we are all beings who can find ourselves in need of assistance from others. This rather specific need determines a specific and narrow duty of mutual aid. In the *Tugendlehre*, the basis of discussion of duties to others is the end which is at the same time a duty of the happiness of others. The fact that we humans have the end of happiness is a consequence of our being finite moral beings, beings not completely self-contained and self-adequate, beings with needs. This *Tugendlehre* claim is a less specific kind of fact about humans – that we have some needs – and is in fact a claim compatible with the proposition that we can each take care of our own needs, without assistance. Perhaps this is an advantage of the *Grundlegung* argument: the need claimed there better grounds a more specific, narrower, and arguably more important obligation. On the other hand the obligation thus justified is narrow and isolated, with little or no indication given of how this sample-argument could provide a model for other imperfect duties to others.

Let us here discuss a bit more the contrast between the "contradiction in conception" moral test of maxims, and the "contradiction in the will." (See G, IV, 423-4) This is a contrast that specifically arises from examples two and four as discussed after the Second Chapter introduction of the first formulation of the categorical imperative. We have already seen that the contradiction in conception test applies only to a small part of the entire range of perfect duties to others (unless we merely define such duties in terms of the idea of contradiction in conception). For this reason it does not provide a good characterization of the range of perfect duties to others according to many of the common preexisting conceptions of such duties, so that e.g., they include the duties to refrain from assault, rape, or murder. The idea of contradiction in conception does not help us understand any argument against suicide, nor is the idea of contradiction in the will

helpful in understanding the proposed obligation to develop one's talents, nor any broader obligation to pursue the happiness of others. This is because the fourth example argument relies on the obliged agent having a potential for experiencing the same need as the person needing assistance, whereas such a specific correspondence is no part of the account of our obligation to pursue the dutiful end of the happiness of others.

V

Since both of the arguments for duties to oneself contain premises making reference to teleological laws, which may not exist, or if they do exist do not seem likely to provide a secure basis for uniquely correct claims of moral obligation, since laws mentioning competing ends in such proposed laws seem equally plausible, Kant has simply not succeeded in the *Grundlegung* in making a good case for his long-held view that there are very significant duties to oneself. Some might say that this is all to the good, since this is anyhow an indefensible part of Kant's theory. Here I'll only comment that to determine adequately such indefensibility, one needs to consider the different account of such duties offered in the *Tugendlehre*, which at least makes a better, fuller case than the *Grundlegung*.¹²

Both of the duties to others arguments are good, defensible arguments. The one relating to perfect duties to others is based on the categorical imperative, first formulation, understood as a principle of justice or fairness, and the intuitive moral objection is that maxims that entail a contradiction in conception are parasitic on other people acting differently. This model argument has a narrower scope of application than one might at first expect. It does not extend to all perfect duties to others, where that phrase is understood as all duties correlated with rights, or all socially enforceable duties. The argument of the fourth example is also an argument for a quite specific duty, one of mutual aid. It may not be broadly generalizable as a model argument, as a sort of reciprocity argument, to a range of other imperfect duties to others. Kant in the *Tugendlehre* in any case does not build on this approach, but instead gives an alternative and much more broad ranging account of imperfect duties to others, in terms of the end that is also a duty of the happiness of others.

When we turn to the four examples reprised after the presentation of the second formulation in the *Grundlegung*, they do better for us. The argument for the duty to oneself against suicide comes somewhat closer to being cogent, and both of the arguments for duties to others are presented in a way that suggests possibilities for their broader generalization. But still it is indispensable to consider the more fully worked out versions of the application from the *Tugendlehre* for a full, accurate understanding of Kant's views. And compared to the first traversal of the four examples, the second traversals are neglected texts.

I have here little to say about *why* the *second* traversals in the *Grundlegung* of the four examples should be better and closer to the later, fully worked out views of the *Tugendlehre*. Some might, accepting this point, wish to use it to argue for the superiority of the *second* formulation of the categorical imperative over the first. Also, it is perhaps understandable that the *first* traversal has gotten more attention. As first, it is more prominently placed in the text than what might have seemed likely to be a mere re-traversal of the same ground. Further, there may have seemed to be more interesting intellectual challenges to understanding how a “universal law” formulation is to be applied; in contrast the second formulation has seemed to many to be fuzzy and unclear. But whatever the explanation, the main point remains that the text that has received the greatest attention is the least satisfactory indicator of Kant's final views.

I conclude that the four examples connected to the “universal law” formulation in the *Grundlegung* do not deserve the detailed (and isolated) attention that has been lavished on them. They do much less to tell us Kant's views about the application of the categorical imperative than many other texts, but especially the *Tugendlehre* part of the *Metaphysics of Morals*, provide. In fact, if one goes to the *Metaphysics of Morals* expecting a simple expansion and development of the ideas presented in Kant's first traversal of the four *Grundlegung* examples, one is in for quite a number of surprises — at least four.¹³

ENDNOTES

¹1930, London, Routledge and Kegan Paul, Ltd.

²Robert Paul Wolff, ed., *Kant: A Collection of Critical Essays*, 1967, Garden City, New York: Anchor Books. The essays by Ebbinghaus, Harrison, Kemp, Harrison, and Matson.

³See A. R. C. Duncan's neglected book *Practical Reason and Morality* (1957, London: Thomas Nelson and Sons) for an interpretation of the argument of the *Grundlegung* that claims that the part of that work in which the four examples are given is an “ethical interlude” that stands aside from the main function and purpose of the work. Duncan's main thesis, that the *Grundlegung* is not a work in ethics, I have no wish to defend, but his account of the argumentative and analytical structure of that work is in its details often insightful.

⁴References to Kant's works will be given in the text; references are to the volume and page of the passages referred to in the Prussian Academy edition of Kant's works (*Kants gesammelte Schriften. herausgegeben Von der Deutschen (formerly Koeniglichen Preussischen) Akademie der Wissenschaften*, 29 volumes (Berlin: Walter de Gruyter (and predecessors), 1902). Abbreviations for two of the main works referred to: G=*Grundlegung zur Metaphysik der Swell.* MdS=*Metaphysik*

der Sitten. I have used the translation of G by H. J. Paton, *Groundwork of the Metaphysics of Morals* (3rd edition, 1956, republished in paperback in 1964 by New York: Harper & Row.) I refer to this work using the first word of the German title throughout this essay, because it has been variously translated as "Groundwork," "Foundations" and "Grounding." I have used the new, complete translation of MdS by Mary Gregor, *The Metaphysics of Morals*, 1991, Cambridge: Cambridge University Press; this work consists of two parts, the *Rechtslehre*, the doctrine (or fundamental principles) of law or right, and the *Tugendlehre*, the doctrine of virtue. Both of these translations include the *Akademie* pagination in their margins.

⁵ For a review of this shift of views, see the opening chapters of Lewis White Beck, *A Commentary on Kant's "Critique of Practical Reason,"* 1960, Chicago: University of Chicago Press.

⁶ Paton presents his views on the role of teleological law in the application of the categorical imperative in his: *The Categorical Imperative, A Study in Kant's Moral Philosophy*, 3rd edition, 1958, London: Hutchinson & Co. For a fuller critique of his interpretation of Kant, see my "Paton on the Application of the Categorical Imperative," *Kant-Studien*, 64. Jahrgang, Heft 3, 1973, pp. 411-422.

⁷ The whole introductory discussion, MdS, VI, 417-420, can be usefully studied in considering the role of what I am distinguishing as natural vs. moral teleology in Kant's fullest presentation of his doctrine of duties to oneself. For directing my attention to these passages and suggesting I discuss more fully the teleological setting of duties to oneself I thank one of the referees for *Journal of Philosophical Research*. Also in the *Grundlegung*, G, IV, 394-396, in a discussion of "the purpose of nature in attaching reason to our will as its governor" (G, IV, 394-5), Kant argues against the view that the most plausible natural end of reason is to teach us the direction to seek our happiness, and that instinct would be a better guide to happiness than reason. He concludes that reason was given us to become a good in itself, i.e., a good will. This striking passage finds no analogues in Kant's later works on moral philosophy.

⁸ There is a lot of resistance to the very idea of duties to oneself, and if we limit our attention to the G, Kant does not make a very good case for them at all. One of the best and best-known of such cases is Thomas Hill's essay on servility, which does draw on MdS, at least for its topic, if not for the textual detail of its argument. For this reason work by a scholar as major as Jerome Schneewind in the new companion to Kant (p. 322, *The Cambridge Companion to Kant*, edited by Paul Guyer (Cambridge, Cambridge University Press, 1992)) seems to indicate, in a sketch of the application of the categorical imperative that it applies only to the duties to others, making no reference to the fact that Kant at least *thought* that there was also a realm of duties to *oneself* that was quite important. He presents Kant's own arguments for specific duties in ways that limit *them* to duties to others. It is difficult or impossible to apply the universal law formulation to assist in deriving any duty to oneself. It may seem as though the universal law formula is really limited to kingdom of ends situations, where there are thus a plurality

of agents in some sort of commerce. This is perhaps equally true of the discussion of duties to oneself in the *Tugendlehre*, as it is true of their discussion in the *Grundlegung*. The second formulation works better with duties to oneself.

Let me add that there is considerable literature on the equivalence or lack thereof of the different formulations of the categorical imperative. In this paper I am taking no position with respect to that issue. I am simply noting that at a fairly intuitive level one formulation works better, or seems to, for deriving certain kinds of duties. The explanation for what I am noting, from the perspective of someone who wishes to defend basic equivalence might be as simple as the following point from Onora O'Neill's contribution to this topic: "The incompleteness of FUL [formula of universal law] and FEI [formula of end in itself] is not a matter of their needing to be supplemented if they are to be applied; the point is rather that neither is fully explicit, and that they are inexplicit in complementary ways." (p. 143, *Constructions of Reason: Explorations of Kant's Practical Philosophy*, 1989, Cambridge, Cambridge University Press). Her entire paragraph, from which this quote is taken, is apropos to this point.

⁹ "Murder and Mayhem: Violence and Kantian Casuistry," in *The Monist*, 72:P3 (1989), pp. 411-432.

¹⁰ "How to Apply the Categorical Imperative," *Philosophia*, 5:4 (October 1975), pp. 395-416.

¹¹ "Mutual Aid and Respect for Persons," *Ethics* vol. 94 (1984), pp. 577-602. The reader can also be most usefully referred to especially Chapters 10 and 12 of O'Neill (op. cit.), and also to her earlier work *Faces of Hunger: All Essay on Poverty Development, and Justice* (1986, London: George Allen and Unwin). In both books she presents a plea for the importance of imperfect duties to others and for a morality based on duties rather than rights.

¹² I now have an essay in preparation that gives an interpretation of Kant on duties to oneself in the *Tugendlehre*.

¹³ This paper benefitted greatly from comments on an earlier version by Mark van Roojen, and the anonymous referees for the *Journal of Philosophical Research*.