

The Philosophical and Practical Significance of Kant's Universality Formulations of the Categorical Imperative

Mark Timmons

I. Kant's Universal Law Formulation of the Categorical Imperative: Some Preliminary Remarks

In the second section of the *Groundwork*, Kant formulates a version of the categorical imperative (CI) – the ‘formula of universal law’ – as follows:

FUL Act only in accordance with that maxim through which you can at the same time will that it should become a universal law (G: 4:421).¹

He then reformulates FUL to include reference to the idea of *nature* – the ‘formula of universal law of nature’:

FLN Act as if the maxim of your action were to become by your will a universal law of nature (G 4:421).

The point of the shift from FUL to FLN is apparently so that the CI can more easily be applied to actions for purposes of deriving conclusions about the moral rightness and wrongness (the deontic status) of actions.² And in the *Groundwork*, just after introducing these universal law formulations of the CI, Kant proceeds to illustrate their application to the now famous cases of suicide, false promises, letting one's talents rust, and refusing to help others. The overall impression conveyed by these and other passages is that Kant proposes the universal law formulations of the CI as expressing a test of the deontic status of actions. In broad outline, the test (employed by an agent evaluating her own action) requires that she formulate her maxim of action and then ask herself whether she could consistently will her maxim as a universal law of nature (for short: whether the maxim is universa-

¹ All references to *Kant's* moral writings include an abbreviation for the relevant work followed respectively by the Akademie volume and page numbers. G = *Groundwork*; KpV = *Critique of Practical Reason*; MdS = *The Metaphysics of Morals*. I have used the translations that are to be found in *Practical Philosophy*, Mary J. Gregor (Ed. and Trans.), Cambridge: Cambridge University Press, 1996. KrV = *The Critique of Pure Reason* and I've used the translation by Paul Guyer and Allen W. Wood, Cambridge: Cambridge University Press, 1998. References to KrV follow the standard A/B citation practice.

² See G 4:421 and 4:437 and KpV 5:67 for Kant's explanation of the transition from FUL to FLN.

lizable). If not, then the action mentioned in her maxim is contrary to duty and hence morally wrong. If the maxim is universalizable, then the action mentioned in the maxim is not contrary to duty and hence not morally wrong. Furthermore, Kant claims that there are two ways in which one's maxim might fail to be universalizable. First, one's maxim might be such that one could not even *conceive* of it 'holding' (as Kant says) as a law of nature. But even in cases where one's maxim can be conceived as so holding, it may still be such that one cannot consistently *will* it to be such a law. This further distinction between two ways in which one's maxim may fail to be universalizable is the basis for distinguishing 'perfect' from 'imperfect' duty.

Were all of these claims about the power of the tests associated with the universal law formulations correct then Kant's universality formulations would represent a truly impressive philosophical achievement. For it would mean that there is a moral principle that has the following three characteristics:

Formality: the principle does not make reference (or otherwise appeal) to substantive claims about what is good or bad unlike those principles featured in standard natural law and consequentialist theories.

Self-sufficiency: the principle can function as a lone moral premise in moral arguments for substantive moral conclusions about the deontic status actions.

Fertility: the principle can be used in moral arguments to derive a broad range of moral conclusions about the deontic status of actions.

But Kant himself observes that 'The simplicity of this law in comparison with the great and various consequences that can be drawn from it must seem astonishing at first, as must also its authority to command without appearing to carry any incentive with it.' (MdS 6:225).

Now what Kant found astonishing, other philosophers beginning with Hegel³, have found unbelievable. For it would seem that the combination of any two members of the above triad is incompatible with the remaining member. For instance, if the universal law formulation of the CI is both formal and the only moral premise in a moral argument (self-sufficient), then it is hard to see how it can have fertile results.⁴ Again, if this formulation is taken to be both formal and yet can play a role in moral arguments with fertile implications, then it would seem that any such argument would need supplementary premises to get from the CI to moral conclusions. And finally, if the universal law formulation of the CI is taken to be both self-sufficient and of use in arguments yielding fertile results, then it is not clear how it can be purely formal. However, against these worries, Kant is not without his defenders. Some interpreters (including some who are our contemporaries) have tried to explain how, when properly interpreted (or reinterpreted), Kant's tests

³ G. F. W. Hegel, *Elements of the Philosophy of Right*, 1821, § 135. See also J. S. Mill, *Utilitarianism*, 1863, chapter 1.

⁴ This particular objection—the 'empty formalism' objection – is discussed further in section 4 below.

can be made to work.⁵ Debate over Kant's universal law formulations of the CI continues.

Notice that this dispute over the adequacy of the universal law formulation⁶ of the CI assumes that it is to be interpreted as a moral testing device, or what, in contemporary parlance, is often called a 'decision procedure'.⁷ And Kant certainly encourages this interpretation in a number of places in his writings. But suppose the skeptics are right and that universal law formulation of the CI can't be made to work as an adequate decision procedure for arriving at correct deontic verdicts about actions? This pessimistic view of the universal law formulation does not mean that Kant's moral theory is to be rejected. Far from it as Allen Wood⁸ has recently argued. After all, Kant also expresses the CI by other formulae and some of them seem to be largely independent of the universality formulations and are perhaps comparatively more defensible. But still, suppose that the pessimism about the adequacy of Kant's universality formulation of the CI is correct and suppose also that other elements in Kant's ethics (including his other formulations of the CI) might still be defensible. May we then conclude that nothing much would be lost from his moral philosophy were we to re-do Kant's ethics without the universality formulation of the CI? If its only role in the overall economy of Kant's ethics is to function as a decision procedure, and if it can't do that job, then isn't its significance in relation to this theory lost?

One might think so. But I think that making such an inference would be too hasty. When I read those passages in section 1 of the Groundwork in which Kant is moving from certain theses about the good will to the universal law formula of the CI, it seems evident that the idea of universal law featured in this formula is really part of an interconnected web of ideas, including the ideas of the supremacy (of moral considerations) and of respect (as a complex attitude). What I wish to argue in this paper is that the true philosophical significance of Kant's universality formulation of the CI lies not in its alleged role as a decision procedure but rather in the fact that the concepts implicated in it serve collectively as formal constraints on what can count as a substantive morally relevant reason for action – the kind of reason grounding moral obligations. Thus, I shall propose what I call the 'formal constraints' interpretation of universal law formulation of the CI. But I also wish to argue that the universality *tests* that are expressed in the universal law formulation have an important practical significance in moral thinking other than function-

⁵ For more detail on this matter, see my, 'The Categorical Imperative and Universalizability', in Christof Horn and Dieter Schönecker (Eds.), *Kant's Groundwork of the Metaphysics of Morals: New Interpretations*, Berlin & New York: De Gruyter, forthcoming.

⁶ Henceforth, I will simply refer to *the* universal law formulation since the difference between FUL and FLN will not matter for what follows.

⁷ Talk of a 'decision procedure' is to be construed broadly enough to allow compatibility with the need for non-rule based moral judgment to play a role in any such procedure.

⁸ Allen Wood, *Kant's Ethical Thought*, New York: Cambridge University Press, 1999, pp. 97–110.

ing as a moral decision procedure. In setting forth and defending the claims, I will first explain more clearly the differing roles that a moral principle might play in the overall economy of a moral theory. Then, in the third and fourth sections, I will proceed to explain the interpretation I am proposing. Because of limitations of space I will not be able to fully develop and defend the interpretation in question, but I do hope to provide a convincing enough case for what I think is a fruitful way of understanding the philosophical and practical significance of Kant's universality formulations of the CI that differs significantly from how they are typically understood.

II. Decision Procedures, Moral Criteria, and Moral Principles⁹

In philosophical ethics, a fundamental moral principle is often cast in a dual role – as expressing both a moral criterion and a decision procedure. But, as many recent defenders of consequentialist and, in particular, utilitarian moral theories have insisted, a moral principle might play one of these roles without having to play the other. Here, let us restrict our attention to principles of right conduct (as opposed to principles of value). A 'criterion' of right conduct (as this term is here being used) picks out some underlying feature of an action *in virtue of which* the action has a particular deontic status. Let us assume just for the sake of simplicity of discussion that there is some *one* underlying feature, F, possession of which by an action makes it right and the lack of which makes an action wrong. On this assumption, the property or feature F of actions in virtue of which an action is right represents a *fundamental right-making feature*. This feature (supposing there is only one) is a 'criterion' of right action.¹⁰

Now moral theories have, as one of their primary aims, the task of inquiring into the nature of right and wrong action in an effort to uncover what it is about them in virtue of which they have the deontic properties they do. On our assumption that there is one such underlying feature, we can formulate a generalization that expresses the criterion in question by the formula: *An action A is right (wrong) iff (and because) A has feature F*. Classical act utilitarianism features a fundamental moral principle having this form:

U An act token A (performed by an agent S in some circumstance C) is (was, would be) right iff (and because) S's performing A in C would produce as great a balance of happiness over unhappiness (considering all of those who will be affected by S's Aing) as would any available alternative act token open to S in C.

⁹ This section of the paper is a re-packaging of section 1 of my 'Decision Procedures, Moral Criteria, and the Problem of Relevant Descriptions in Kant's Ethics', *Jahrbuch für Recht und Ethik* 5, 1997.

¹⁰ Below, in section 3, I will refine this characterization of a moral criterion.

But, as utilitarians point out, U is intended to express a moral 'criterion of right action and it need not be taken as also expressing a procedure to follow in moral deliberation. Understood as a decision procedure, U would have an agent who is engaged in moral deliberation and choice try to figure out the overall net effects on everyone's happiness (and unhappiness) of all of the alternative actions open to her in some circumstance. Not only might this be a very poor way of going about making reliable moral choices (given utilitarianism), but following the procedure in question might be prohibited in some circumstances, since deliberating is an action which may produce worse consequences than doing something else, including using some other method for arriving at moral decisions. Indeed for a utilitarian, if it were to turn out that were everyone to attempt to maximize her *own* happiness the overall result for everyone's happiness would be highest, then the utilitarian would find herself recommending a principle of egoism as a proper moral decision procedure while maintaining that utilitarianism is a correct moral criterion. Whatever sort of 'schizophrenia' this combination of views might represent in moral theory, it illustrates the point that a moral principle might play the role of a moral criterion and not the role of a decision procedure (and vice versa).

Now all this is familiar stuff from contemporary moral philosophy. However, one might wonder how this distinction between a moral principle as a criterion and a moral principle as a decision procedure might play out in Kant's ethics. Elsewhere I have defended what I call the 'differential roles' interpretation of the CI, according to which (roughly) different formulations of the CI play different roles in Kant's ethics.¹¹ On this interpretation, the universal law formulation of the CI is cast in the role of a decision procedure – which, as explained earlier, surely fits with much of what Kant says about this formulation and how he uses it in his moral writings. However, we should not assume that it also expresses a fundamental criterion of right action. That is, we should not assume that what *makes* an action right (or wrong) is the fact that its associated maxim is universalizable (or not).

Rather, I claim that it is the humanity formulation of the CI which expresses a fundamental moral criterion in Kant's moral theory. Here is how Kant expresses this formulation in Section II of the Groundwork:

FH So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means (G 4:429).

On my understanding, what FH says (in effect) is that the rightness or wrongness of an action is *explained by* facts about how one's action 'bears' on what Kant calls 'humanity'. Kant's term 'humanity' (*Menschlichkeit*) refers to our rational capacities as agents who not only are able to deliberate and act on the basis of what they take to be good normative reasons, but are capable of autonomy which, as Kant says, 'is the property the will has of being a law to itself (independently of

¹¹ See my 'Decision Procedures, Moral Criteria' paper cited above in note 8.

every property belonging to objects of volition' (G 4:440). It is our rational capacity, then, that is the most fundamental morally relevant feature bearing on the deontic status of actions.¹²

Here is not the place to revisit these claims about the varying roles of different formulations of the CI. What is important for what follows are the following claims which collectively summarize my previous remarks. First, the universality formulation of the CI seems to be cast in the role as a decision procedure involving a consistency test on maxims. But second, it need not be understood as also expressing a moral criterion. Rather (third) it is the formula of humanity that serves as a moral criterion which thus purports to indicate a fundamental right-making feature of actions.

III. The Formal Constraint Interpretation

Let us suppose then, that the skeptics are right and the universal law formulation of the CI cannot be made to work as an adequate decision procedure. One reaction to this thought might be (and had been) to understand the universalizability constraint featured in the CI as a mere 'logical' constraint on moral thought and discourse. Henry Sidgwick expressed the constraint as follows:

LU If a kind of conduct that is right (or wrong) for me is not right (or wrong) for someone else, it must be on ground of some difference between the two cases, other than the fact that I and he are different persons.¹³

LU is formal, but it cannot be used to derive moral verdicts about actions and, indeed, it is compatible with most varieties of competing moral theory. If the true significance of Kant's universality formula is reduced to LU, then it is of some, but not great significance, in Kant's moral theory, and certainly not one of its distinguishing features. The main idea of this paper is that there is something in between having to understand Kant's universality formulation of the CI as a formal, self-sufficient, and fertile decision procedure and, alternatively, understanding it as giving us nothing more than LU. My claim (to repeat) is that even if the universality formulation of CI fails to represent an adequate decision procedure in ethics, it can be understood as involving an interconnected set of formal constraints on what can count as a fundamental morally relevant right-making feature of actions. In other words, think of Kant's universality formulations of the CI as specifying some formal feature that any substantive right-making feature must have *given our ordinary*

¹² This is not to say that because FH represents a moral criterion that it cannot also be used in moral deliberation or as a basis for reasoning one's way to moral conclusions. In fact, it is the humanity formulation that Kant almost always employs in the Tugendlehre in arguing for various duties to oneself and to others.

¹³ Quoted from *H. Sidgwick, The Methods of Ethics*, 7th edition, 1907) p. 397, Indianapolis: Hackett Publishing Co., 1981. See also, pp. 209–210 where Sidgwick seems in effect to suggest that Kant's universal law formulations of the CI only convey what is expressed by LU.

concept of a moral obligation. This, at any rate, is the idea that I will be developing in what follows.

As a start, let us distinguish between the philosophical significance and the practical significance of the universality principles. The philosophical significance of the idea of universal law in the overall economy of Kant's ethics is that a proper understanding of it reveals to us an interconnected set of considerations that collectively serve pick out *moral* reasons and distinguish them from non-moral, including prudential, reasons for action. In this way, Kant's universality principles (or rather, the interconnected set of ideas implicated in such principles) are stronger than LU: they involve formal notions that collectively constrain the available candidates for being fundamental moral reasons. The practical significance of Kantian universalizability is that it can serve as a form of ad hominem argumentative device that, when properly applied, reveals a kind of duplicity in the thinking of those agents who act contrary to morality. In the remainder of this section, I will be developing the philosophical significance of the universal law formulations and then, in the following section, I will turn to their practical import.

The philosophical significance of universality

The philosophical significance of universality in Kant's ethics has to do (so I claim) with three interconnected topics of philosophical inquiry: the *content*, *ground*, and *normativity* of moral obligations. Let me first explain what philosophical tasks these topics involve and explain how I understand the relation between Kant's conception of universality and the tasks in question.

In the general introduction to the *Metaphysics of Morals*, Kant tells us that '*Duty* is that action to which someone is bound. It is therefore the matter of obligation, and there can be one and the same duty (as to the action) although we can be bound to it in different ways' (MdS 6:223, see also G 4:439). This sentence contains the distinction between what I am calling the content of moral obligation (duty) and the ground of obligation – the considerations in virtue of which some action is a moral duty. As Kant remarks, one can be morally required to perform the same action on a variety of distinct grounds. Let me now connect considerations of content and grounds of moral obligation with moral principles.

As I understand moral principles, they are of two sorts: high-level and mid-level. The most abstract, high level moral principles, such as the principle of utility and the humanity formulation of the CI specify the fundamental grounds of moral obligation. That is, they specify for any action (abstractly specified simply as an action) what it is about the action that, most fundamentally, makes it right or wrong. By contrast, mid-level moral principles (often called 'moral rules') such as the Ten Commandments and Ross's list of prima facie duties, specify action types (e.g., lying, harming, self-improvement, etc.) that pick out those types in terms of morally relevant descriptions that 'build into' the characterization of the action a

ground of obligation. To characterize an action as a lie is to characterize an action type in terms of a consideration (intentionally communicating to another a proposition that one does not believe with the intention of deceiving the recipient of the communication) that represents a ground of obligation (or *prima facie* obligation). Such mid-level moral principles or rules, then, specify both content and grounds of duty in their specification of action types that are required or prohibited. I believe that principles of both sorts play important roles in Kant's ethics, a topic, consideration of which, would take us too far a field for present purposes. For now the point to which I am calling attention is that one philosophical task for a moral theory is to specify the content and ground of obligation. Doing so is what is involved in providing a moral criterion.¹⁴

The question about normativity concerns how it is that moral considerations have the normative force or authority they have and, in particular, how they have the kind of normative supremacy as reasons that they are commonly taken to have. One would expect that this issue is deeply related to the issue about the content and grounds of duty: in order for some consideration to be one that 'calls for' and hence grounds some duty, it must have a kind of normative authority – indeed a normative supremacy – for those agents who are subject to moral requirements.¹⁵

How, then, is Kant's universality principle connected with the three philosophical tasks just mentioned? As I've said, my idea is that the universality principle involves an interconnected network of formal elements that a feature must have if it is to count as fundamental right-making feature and how we are to respond to it, at least given our common sense conception of duty. And (to cut to the chase) the feature that such formal elements serve to pick out is humanity as a ground and a certain kind of response to humanity as consisting of duty. So the picture I have in mind is one in which the humanity formulation of the CI represents a moral criterion, while the universalizability formulation is perhaps best understood as (in effect) encapsulating a set of formal requirements on what can count as a fundamental moral criterion. At least this is the idea I now want to develop.

Let me admit that I am taking some liberties with Kant's texts. He doesn't present the universality versions of the CI as I am proposing to do. However, I do think that my proposal meshes nicely both with some of the dominant themes in Kant's moral philosophy and with what I take to be the practical significance of universalizability as a testing device as this idea is featured in Kant's writings, particularly the Groundwork.

¹⁴ A moral criterion may specify some single necessary and sufficient right-making feature (or duty-making feature), but a criterion might only specify a *prima facie* sufficient duty-making feature as we find with Ross's principles of *prima facie* duty. Hence, talk of a moral criterion allows for both monist and pluralist accounts of right-making features.

¹⁵ At least this tight connection between the content of moral obligation and normativity holds if one is a so-called internalist about obligation and normative reasons. For a defense of Kant as this kind of internalist, see my 'The Possibility of Moral Motivation in Kant's Ethics', *Southern Journal of Philosophy* 23, 1985, pp. 377–398.

With these preliminary remarks out of the way, we may now proceed to examine some of the detail of the formal constraints interpretation.

The formal constraints

There are three formal features that we find introduced in section I of the Groundwork that I believe are deeply interconnected in what Kant takes to be common rational moral thinking. They are the features of *law-likeness*, *supremacy*, and *respect*. Each of them can be understood to represent a constraint that a moral criterion must satisfy. Let us proceed to briefly consider the various constraints in order.

Constraint 1: Law-like character of moral thought

In the preface to the Groundwork, Kant claims that moral laws must hold necessarily for all rational agents as such:

Everyone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command "thou shalt not lie" does not hold only for human beings, as if other rational beings did not have to heed it, and so with all other moral laws properly so called; that, therefore, the ground of obligation here must not be sought in the nature of the human being or in the circumstances of the world in which he is placed, but a priori simply in concepts of pure reason . . . (G 4:389).

In this passage (and others like it)¹⁶, Kant is appealing to what he takes to be the common sense notion of duty and he is making a point about the 'ground' of obligation. Notice that in the first sentence of the passage, Kant speaks of a *law* as being a ground of obligation and mentions a mid-level moral principle (rule) as an example. As just explained, such rules specify the content and ground of a basic obligation by referring to an action by way of a term that encapsulates a morally relevant description. Now as I understand this passage, Kant is making two points: one about the grounds of obligation, and one about appropriate philosophical methodology for investigating these grounds. Let us take these one at a time.

The point about grounds involves two claims: (1) If moral laws (at least the most basic ones) are to be 'valid' for all rational agents (including, but not limited to human agents), then there must be 'grounds' of obligation – considerations in virtue of which an action is obligatory – that serve as reasons for all rational agents. Such reasons are, as Kant would say, 'strictly universal' in scope. (This point is an implication of the common sense assumption that such laws are 'absolutely necessary'.)¹⁷ (2) Such grounding considerations or facts cannot have speci-

¹⁶ In the Groundwork, see 4:408, 4:412, and 4:425.

¹⁷ See KrV B3-B4 where Kant explicitly connects the notions of necessity and strict universality. He writes: 'Necessity and strict universality are therefore secure indications of an *a priori* cognition, and also belong together inseparably'.

fically to do with human beings and their distinctively human nature, otherwise such considerations may not provide reasons for all rational agents, and thus would not be strictly universal in scope.

The methodological point (which won't concern us here) is that (3) given the modal status of fundamental moral principles and the kinds of reasons they must feature, the search for the grounds of obligation must proceed a priori 'simply in concepts of pure reason.'

So in this passage, Kant is appealing to the kind of 'absolute' necessity characteristic of moral principles – a characteristic which is partly constitutive of a *law* – and what we learn is that such necessity requires there to be facts that ground moral laws, or to put it as I would prefer: this necessity requires there to be facts that necessarily provide reasons for action to all rational agents. Let me spell this out a bit more in three steps. First, we have the idea that moral principles are practical laws which in turn implies that they hold (and must hold) for all rational agents. (Moral principles on Kant's view are a species of rationality norm and so these norms are partly constitutive of what practical rationality is all about.) Second, because such laws are laws of practical rationality, they must be grounded in reasons – considerations that count in favor of (at a minimum) the action mentioned in the law. If we now put the first two points together, we arrive at the third: there must be considerations that count as reasons for all rational agents which ground moral laws. Call such reasons 'categorical' in their authoritative force.

Now let us return to our project of using formal constraints to fix a moral criterion – at least a criterion that common sense presupposes. So, the question is: how much constraining power does the idea of categorical grounding reasons have? What candidates for a grounding reason does the 'absolute necessity' of moral principles rule out? Here, I confess that matters become complicated and answering this question would require an investigation into the various candidates for moral criteria featured in competing normative moral theories which I cannot undertake here. However, I think we can say this much on Kant's behalf: any putative ground of action whose normative authority depends only on inclination is ruled out. The problem is determining which normative grounds or reasons for action get ruled out here. For instance, why can't we say that the perfection of rational agents or their well-being, or both, are grounding reasons that all rational agents have reason to promote or at least maintain and thus that either one of these is a viable candidate for a grounding reason? After all, arguably the value and normative authority of such considerations are not (or need not be) accounted for as *being* valuable or authoritative by being objects of inclination.¹⁸

¹⁸ At least not without argument. In the second section of the Groundwork at 4:428, Kant distinguishes between items that have relative worth and items having absolute worth according to which the former get their worth, so to speak, by being objects of inclination. What Kant needs is an argument for the claim that only humanity has absolute worth. Again, space does not permit delving into Kant's Groundwork II arguments for this claim.

So far as I can tell, then, the absolute necessity of moral laws (and the categorical reasons that this kind of necessity implies) is not strong enough to rule out as grounding considerations all competitors to the property of rational agency (or what Kant refers to as 'humanity'). Furthermore, in one way, Kant's moral theory does allow considerations of perfection and well-being to ground duties to oneself and duties to others respectively. But here we must be careful to distinguish 'fundamental' grounds from 'derived' grounds. In Kant's scheme, considerations of perfection and well-being do ground duties to self and other respectively, but they themselves are grounded in something moral fundamental. And our quarry here is fundamental grounds for moral laws. Thus, the necessity and strict universality of moral considerations implicated in the very idea of a moral law do not serve to rule out certain available candidates for being a fundamental moral criterion. So let us now proceed to the second formal constraint.

Constraint 2: Supremacy of moral reasons

In order for a consideration to have normative authority, it must count as a reason for (or against) some action. Some considerations having normative authority are solely prudential, others are moral. Arguably, common sense moral thinking presupposes that moral considerations have a kind of normative superiority such that in cases where moral reasons and nonmoral reasons conflict, the former have greater normative authority. How to work out this idea of conflicts among moral and nonmoral considerations is something I won't get into here. For our purposes, what is important is the common sense presumption about the normative supremacy of moral considerations. Since we are interested in fundamental grounding considerations behind moral principles, common sense moral thought presupposes that whatever considerations are fundamental here, they must also possess the kind of 'normative supremacy' just indicated. What does this formal feature tell us about the content of fundamental moral reasons?

Before proceeding further, notice that the constraints of law-likeness and supremacy are distinct considerations: satisfying one does not entail satisfying the other. The fact that some consideration will always trump competing considerations does not (so far as I can tell) entail that such considerations represent categorical reasons. Suppose modest rational egoism as a theory of reasons were true, according to which (1) one may have non-self-interested reasons for action but (2) whenever such non-self-interested reasons seriously conflict with reasons of self-interest, the latter always prevail in their normative force. If we assume that self-interested reasons are not categorical – my well-being provides me with a reason for action but need not provide anyone else with a reason – then we have here a theory of practical reasons where normative superiority does not entail that such reasons are categorical. (And we have already noted that from the fact that some reason is categorical, it does not follow that it is normatively supreme.)

To return to the supremacy criterion: let us suppose (on Kant's behalf) that reasons for action are themselves traceable to what has value. Thus, something's

being good provides (or can provide) someone with a reason. On this assumption, it would seem to follow that if some type of normative reason were normatively supreme, then there must be something of value that is itself of supreme value and which explains the supremacy attached to the grounding reason. So, according to the criterion of supremacy, grounding considerations must be something of supreme value – a value that would explain why the thing in question provides supremely authoritative reasons to all agents.

In relation to our formal constraints project, the question is what does this constraint rule out as possible grounds of moral laws? Now if we consider various passages in Kant's texts, it is clear that he thinks that in order for something to play the role of providing a supremely authoritative reason, it must have a kind of worth – a 'dignity' – that only humanity can have. Kant's argument for this claim is the subject of much contemporary controversy and we cannot pause here to consider the matter in any detail. But it is of interest to consider whether, were Kant's argument for the supremacy of humanity cogent, we could conclude that the humanity formulation is the moral criterion presupposed by common moral cognition.

So, if we put the constraints of law-likeness and supremacy together, how far have we progressed in our task of using formal constraints to zero in on a fundamental moral reason (assuming there is just one)? You might think that these two features get us to our destination: humanity is presumably a feature that provides reasons for all rational agents, and given its unique value, we are home free: humanity has a kind of dignity that makes rational nature a supremely authoritative categorical reason for action – the sort of reason needed to ground moral laws.

But this verdict would be too hasty. Granted, the two formal constraints in question do help fix a fundamental grounding feature of moral obligation, but we are after a moral criterion that involves a specification of ground and *content* of moral obligation, and specifying the former is not sufficient for specifying the latter. After all, taking humanity to be a supremely authoritative ground of morality is compatible with a moral theory that is consequentialist in structure and which takes rational nature as the most basic morally relevant consideration bearing on action. Consequentialism is the view that (1) the deontic status of an action is entirely dependent upon the overall intrinsic value of the consequences of the action and (2) intrinsic value is to be maximally promoted.¹⁹ So in addition to fixing on the proper ground of moral laws there is in addition determining what response is fitting or proper on behalf of rational agents in response to the worth of humanity. Presumably, Kant thinks that ordinary moral thought is non-consequentialist in nature – promoting humanity in the sense of bringing about more of it is not the (sole) proper response to this value. So the formal criteria of law-likeness and supremacy

¹⁹ Here, I am ignoring the potentially important distinction between act and rule versions of consequentialism; nothing that I say turns on this difference.

do not collectively rule out important competitors to the moral criterion featured in Kant's moral theory.

Constraint 3: Respect as a proper response

In the Groundwork, Kant tells us that respect is 'a representation of a worth that [not only] infringes upon my self-love' (G 4:401n), but that this item of worth 'far outweighs any worth of what is recommended by inclination' (G 4:403). He also identifies this item of worth when he tells us that the proper object of respect is the moral law. There are two things to note here. First, when Kant says that the law is the proper object of respect, I read him as in effect claiming that the sorts of considerations that necessarily provide reasons for action to all rational beings is a proper object of the attitude of respect. (More on this below.) Second, it is the supremacy of the law (better: moral considerations) that 'call forth' or demand respect²⁰ and, in particular, the fact that moral considerations check our self-love and 'strikes down' what Kant calls 'self-conceit' – in effect a kind of smug comfort one might take in herself without having made moral considerations fundamental in one's estimation of self worth. If we now ask what sort of contentful consideration merits this kind of response, one obvious candidate is our rational nature. And perhaps, if we dig deeper into the nuances of the attitude of respect, we might find that *only* rational agency can inspire in rational agents this kind of complex response.

So, the attitude of respect at least fits nicely with the supremely authoritative and categorical nature of rational agency as a reason bearing on action. But perhaps more importantly, the attitude of respect as a fitting attitude toward rational nature helps to eliminate a consequentialist account of the moral criterion for which Kant is searching. Arguably, respecting rational nature is not equivalent to merely *promoting it*, though exactly what it demands of us is a subject of investigation that requires just the sort of project Kant undertakes in the Tugendlehre: namely, establishing ends that we have a duty to have (in light of our distinctively human nature) and working out a system of requirements in light of those ends.²¹ These requirements are not simply a matter of promoting the value of humanity. To take one example, Kant claims that the duty of gratitude requires that we *honor* a benefactor, where honoring of this sort does not involve promoting that person's welfare.²² Again, these remarks about the attitude of respect are far too brief to make a convincing case for the idea that it can serve as a formal constraint on any

²⁰ See, for instance, KpV 5:80.

²¹ According to what *T. M. Scanlon* calls 'teleological' views of value, intrinsically good states of affairs are to be promoted. See *his*, *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, chapter 2. Scanlon argues against teleological views and, in particular, with regard to valuing human life.

²² Otherwise, this duty would really be an instance of the duty of beneficence – 'to promote according to one's means the happiness of others in need, without hoping for something in return' (MdS 6:453).

moral criterion that adequately captures common sense moral thinking. This, I leave for another occasion.

With the elements of the formal constraints interpretation on the table, let us now sum up. Moral reasons qua moral must satisfy an interconnected set of formal constraints: they must be such that (1) all rational agents have such reasons for action, (2) they must be supremely authoritative, and (3) they must be a proper object of respect. I am suggesting that these tightly interconnected ideas are (in effect) encapsulated in Kant's universality formulations of the CI. I have not of course been able to supply all the argumentative gaps that would need filling in order to show just how much collective constraining power these constraints yield. As I've said, these details will have to be left for another occasion. But let us proceed for the time being upon the perhaps Panglossian assumption that the gaps can be filled and take stock of where we are.

*The formal constraint argument*²³

What I've been suggesting is that the various formal constraints might plausibly be viewed as the central part of an argument for the claim that the humanity or, more specifically, treating humanity always as an end and never as a mere means represents a moral criterion. The argument would go as follows:

1. According to 'common rational cognition', in order for some consideration bearing on the rightness of action to count as a moral reason, it must satisfy formal constraints 1–3 listed above.
2. Humanity as an end in itself satisfies all the relevant formal constraints.
3. Humanity as an end is the only consideration that satisfies the formal constraints.

Thus,

4. According to common rational cognition, humanity as an end in itself is a fundamental moral reason.

Since a moral principle cast in the role of a moral criterion is in the business of specifying some one (or more) fundamental morally relevant features that explain why an action has the deontic status it does and serve as normative reasons for action, the import of this argument is that it takes us from formal constraints on

²³ Here is an appropriate place to mention the attempt by *Samuel J. Kerstein* in his recent book, *Kant's Search for the Supreme Principle of Morality*, Cambridge: Cambridge University Press, 2002, to defend what he calls a 'criterial' interpretation of Kant's attempt to justify the CI. Space does not permit a discussion of Kerstein's very interesting interpretation, however, what I am calling the 'formal constraints' interpretation is similar in spirit (but not in letter) to Kerstein's criterial interpretation. Kerstein proposes a list of 8 criteria that a supreme principle must satisfy – a list that does not include the criteria of supremacy (of moral reasons) and respect.

moral reasons to a fundamental substantive moral principle – the humanity formulation. Of course, whether this argument can be defended, especially its third premise, cannot be examined here. However, let me conclude this section with three observations about my formal constraint interpretation.

First, in the Groundwork at 4:436, Kant tells us that '*Autonomy* is . . . the ground of the dignity of human nature and of every rational creature'. So, although having dignity is the fundamental value that grounds moral requirements, it is the autonomy of rational beings – roughly, their capacity for 'giving' law that represents a feature of such beings that is most fundamental in Kant's theory of right conduct and value. At the beginning of this section, I mentioned that in addition to philosophical questions about the *content* and *ground* of moral reasons, there are questions about the *normative authority* of such reasons. I also mentioned that these issues are deeply related in Kant's thought: in order for a consideration to count as a moral reason bearing on action, it must have a kind of normative supremacy vis-à-vis competing reasons. Kant's account of normative authority has to do with his notion of our natures as 'law giving' creatures, and so a full and complete story about moral reasons in Kant would require exploration of his notion of autonomy. This I leave for another occasion.

Second, if my formal constraints interpretation is at all plausible (though not necessarily as a reading of how Kant does in fact argue, but one that is compatible with the main doctrines in Kant's ethics), then one might expect that the very conception of duty has built into it the features that would help us determine the content of a moral criterion. In section 1 of the Groundwork, Kant defines duty as 'the necessity of an action from respect for law' which either explicitly or implicitly (as Kant's texts makes clear) involve the three formal notions described above. And this seems to be keeping with the argumentative spirit of the first two sections of the Groundwork.

Third, as explained earlier, one of the interpretative desiderata for any plausible interpretation of Kant's notion of universalizability is that the kind of constraint expressed by this notion be stronger than the logical thesis of universalizability. This desideratum would be nicely satisfied were Kant's universalizability tests adequate tests of the deontic status of actions. If they aren't, there is still an important philosophical role they might play. My formal constraint interpretation takes Kant's conception of universality as a metaethical constraint on reasons being moral reasons, and they are supposed to have enough power collectively to pick out certain considerations, in terms of their content, as being moral reasons. So, on the one hand, it is stronger than the logical principle of universalizability. But on the other hand, understood as I propose, it constrains without serving as a formal and self-sufficient decision procedure from which we can derive an adequate system of duties. The content that the humanity formulation adds to the categorical imperative is significant for purposes of reasoning one's way to a system of duties. And so it is not surprising that when Kant turns his attention to deriving a rich set of duties in the *Tugendlehre*, he does so, on the basis of the humanity formulation.²⁴

IV. The Practical Significance of Universality

There is one obvious loose end that my proposal must deal with and this concerns the *practical* significance of Kant's universalizability constraint. There is evidence in Kant's texts (that I alluded to earlier) particularly in the Groundwork, that he thought of the universal law formulation of the CI as a self-sufficient decision procedure; and this I do not doubt. For instance, in section II, just after introducing FUL, Kant remarks that 'if all imperatives of duty *can be derived* from this single imperative . . . we shall at least be able to show what we think by it and what the concept wants to say' (G 4:421, my emphasis). And in the sample applications we find bits of reasoning that certainly suggest that FLN is being used as a self-sufficient decision procedure; that consideration of one's maxim as a possible universal law of nature can, together perhaps with auxiliary non-moral assumptions, generate moral verdicts.

But let me suggest that even if the universal law formulation of the CI can't play this role as a self-sufficient moral principle in a decision procedure, the tests associated with it are of use in moral deliberation; specifically, they serve to help reveal a kind of duplicity in the thinking of those who might attempt to 'justify' their own immoral behavior.²⁵ To explain this more fully, I will make a number of observations first about the point and purpose of Kant's use of the universalizability tests in his four famous Groundwork examples, and then I will call attention to some features of his examples that fit nicely with what I take to be their practical role.

I believe a useful place to begin in thinking about the overall purpose of Kant's use of the universalizability tests in the Groundwork is by considering what Kant says about the tests in the second paragraph (G, 4:424) following his presentation of the four examples. Because of the length of the paragraph in question, I will quote it in three installments, making comments as I go. The paragraph begins:

[1] If we now attend to ourselves in any transgression of a duty, we find that we do not really will that our maxim should become universal law, since that is impossible for us. But that the opposite of our maxim should instead remain a universal law, only we take the liberty of making an *exception* to it for ourselves (or just for this once) to the advantage of our inclination.

²⁴ I might be accused at this point of false advertising. At the outset, I promised to provide an interpretation of universalizability, but what I have really done is offer an interpretation of the notion of universal law and what I claim are conceptually associated notions (normative supremacy and respect as a fitting attitude). This bait and switch was accomplished by vague talk of 'universality', so the complaint might go. Reply: if one wants to use 'universalizability' associated with Kant's moral theory to refer exclusively to a consistency test on maxims involving a hypothetical scenario, I won't complain. In the next section, I explain what I take to be the true significance of this kind of thinking. What I have been doing, then, can be understood as presenting what I take to be the philosophical significance of the central concepts implicated in Kant's universality tests.

²⁵ This point is also made by *Allen Wood*, *Kant's Ethical Thought*, cited above in note 7, pp. 107–110.

Notice that the focus here is on an agent who 'transgresses' duty and that person's associated psychology of attempting to justify his behavior by making exceptions. The passage continues with more about this agent's psychology:

[2] Consequently, if we weighed all cases from one and the same point of view, namely that of reason, we would find a contradiction in our own will, namely that a certain principle be objectively necessary as a universal law and yet subjectively not hold universally but allow exceptions. Since, however, we at one time regard our action from the point of view of a will wholly conformed with reason but then regard the very same action from the point of view of a will affected by inclination, there is really no contradiction here but instead a resistance of inclination to the precept of reason (*antagonismus*), through which the universality of the principle (*universalitas*) is changed into mere generality (*generalitas*) and the practical rational principle is to meet the maxim half way.

Notice that the contradiction 'of reason' that Kant mentions in the first sentence of [2] is one of both holding some principle to be universally binding on all agents and yet (in effect) denying this universality in making an exception of oneself. In cases of moral transgression, what happens is that the point of view of inclination dominates and the universality characteristic of a moral principle is thereby compromised by (in effect) treating the principle as have merely general (and thus defeasible) authoritative force.

The final part of the paragraph indicates what I think is crucial to Kant's examples: the fact that the agents in those examples (and agents generally) recognize battery of considerations that have a non-prudential normative authority:

[3] Now even though this [treating moral principles as merely defeasible] cannot be justified in our own impartially rendered judgment, it still shows that *we really acknowledge the validity of the categorical imperative* and permit ourselves (with all respect for it) only a few exceptions that, as it seems to us, are inconsiderable and wrong from us. [emphasis added]

Let me now bring together the observations we've made in relation to the quoted paragraph. The practical significance of Kant's universality tests seem to lie in what they reveal (or can reveal) to an ordinary agent who is contemplating some course of action that she or he is reflecting upon. There are three related features constitutive of such thinking. (1) These tests help reveal a distinct source of practical reasons – reasons that can be 'accessed' so to speak by taking an impartial stance toward one's proposed course of action. (2) Reasons we have that are revealed by taking this impartial stance often conflict with reasons grounded in considerations of prudence. And (3) the agent's recognition that the immoral course of action cannot be justified by impartial reason suggests that the agent recognizes the normative supremacy of moral requirements. So the fact that the agent takes moral requirements to have this kind of supremacy is the basis for a kind of duplicity in the agent who (typically) rationalizes her behavior making an exception of herself. What I wish to call attention to here is that for the universality tests to play this role, they need not represent self-sufficient decision procedures for deriving conclusions about one's duties.

What I have been saying about the practical import of Kant's universality tests fits nicely with other features of Kant's four examples – features that fuel a variety of objections to Kant's tests when construed as self-sufficient decision procedures. Let me state the objections and then proceed to explain why granting the points featured in them do not undermine what I take to be the practical significance of Kant's tests.

(1) *Empty formalism objection.* The universality test either smuggles substantive moral assumptions into the examples, and thus the CI does not represent a self-sufficient moral principle, or, if the test is interpreted so that it is not guilty of such smuggling, then its formality will mean that it is empty – lacking in substantive moral implications. For instance, in the case of suicide, the appeal to a teleological law seems to involve an appeal to moral considerations about the moral appropriateness of a motive. In the case of lying promises, the moral acceptability of the institution of promising is taken for granted. And in the talents and beneficence examples, appeals to 'rational' maxims – maxims that any non-finite agent qua rational would necessarily adopt and thus maxims that any finite rational agent ought to adopt – involve background moral assumptions. Thus, the idea that the universal law formulation of the CI is *self-sufficient* is mistaken.²⁶ Furthermore, without the moral assumptions in question, the principle (and the associated test) is empty in the sense that it does not yield a determinate set of moral verdicts.

(2) *Disguised egoism objection.* In the rusting talents and beneficence cases, Kant seems to be appealing to egoistic considerations, which is at odds with his attempt to show that there are non-egoistic reasons for action. For instance, in the example of refusing to help others, Kant observes that although one can consistently conceive of one's maxim as universal law, one cannot will it as law because '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 will, he would rob himself of all hope of the assistance he wishes for himself' (G 4:423). Commenting on this example, W. D. Ross claims that 'here again the appeal is to results, and in this case to definitely hedonistic results; benevolence is justified by an appeal to far-sighted selfishness.'²⁷

(3) *Limited scope objection.* One of the aims of Kant's example applications of the categorical imperative is to show that there are categorical reasons for not engaging in such activities as suicide, false promising, allowing one's talents to rust, and refraining from helping others. But in many of these examples (if not all of them), Kant's arguments fail to demonstrate that *all* agents have categorical reasons. For instance, in the beneficence example, Kant's application of the CI assumes that all agents do will (or ought rationally to will) that they be helped when in need. But, it is doubtful that all actual agents (e.g., the die-hard individualist) in fact will that they be helped, and it is questionable whether such agents are thereby guilty of a failure of rationality. Similar remarks apply to the rusting talents case. Thus, so the objection goes, at least some of Kant's sample applications of the CI fail to show that all agents have categorically binding reasons for action.

In response to these objections, I think that we ought to concede the various observations about Kant's tests featured in them – Kant's tests do involve moral

²⁶ As noted earlier, strictly speaking it is the combination of formality and self-sufficiency that fuels the 'empty formality' (lack of fertility) objection.

²⁷ W. D. Ross, *Kant's Ethical Theory*, Oxford: Oxford University Press, 1954, p. 47.

assumptions, they do involve an appeal to how the agent would be affected in the hypothetical scenarios, and strictly speaking they do have limited scope. However, conceding all of this (if properly understood) does not impugn what I see as the true practical significance of Kant's universality tests. Let explain further.

As a start, we need to keep in mind the audience and main purpose of the examples in question. The audience to whom the examples are addressed are individuals who, like the anonymous characters in all of the examples, have a well-functioning moral conscience in the sense that in cases requiring moral thought, they have 'enough conscience' to ask themselves whether what they propose to do is right (and not just in their self-interest). And as lately noted, the main purpose of the examples is to make clear a mode of practical thinking that brings out an agent's recognition of the supremacy of moral reasons vis-à-vis prudential reasons. Thus, the point of the tests on the reading I am proposing (contrary to the empty formalism charge) is not to illustrate a self-sufficient decision procedure. Viewed in this way, it does not matter that the various examples in which Kant invokes the universality test already involve background moral assumptions – assumptions about the proper function of the motive of self-love, the morality of the practice of promising, and so forth.

As for the disguised egoism objection, it is true that in Kant's false promising, rusting talents, and refusing help examples, the effects of willing one's maxim as universal law as it bears on the agent's own projects is crucial in those examples. But, as many commentators have noted, Kant's appeal in these examples is not based on an appeal to egoism: Kant's arguments are not based on the *disvalue* for the agent of willing her maxim to be universal law. Rather, the appeal is to volitional consistency. For instance, in the false promising example, to will that one's maxim become universal law (i.e., to will that everyone adopt and act on the same sort of maxim) is to will a system in which one's attempt to get money on a false promise is frustrated. It is the appeal to willing a sort of inconsistency featuring self-frustration – willing to get money by a false promise in a system in which the purpose of promising has been undermined and so, in effect, willing that one's aim of getting money not be obtainable by a false promise – that is the operative idea in Kant's example.²⁸ Still, even if Kant's examples are not based on egoism, they are what we might call 'self-focused' in the sense that it is the volitional consistency of the agent who is contemplating acting on this or that maxim that is crucial. And, given what I've explained is the practical import of these examples, their self-focused element is entirely fitting. They function to make manifest a kind of duplicity *in an agent* who would transgress duty.

²⁸ Here, I am working with an 'inconsistency in intention' interpretation of Kant's test, sometimes called the 'practical inconsistency' interpretation that we find in Korsgaard. See C. Korsgaard, 'Kant's Formula of Universal Law', in: *Pacific Philosophical Quarterly* 66, 1985, pp. 24–47. Reprinted in C. Korsgaard, *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, 1996, and in P. Guyer (ed.), *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, Lanham, MD: Rowman & Littlefield, 1998.

Finally, if, as I have said, the universality tests are addressed to 'normal' human agents – agents who can be presumed to have a normal set of interests, desires, and needs – then the fact that the talent and beneficence arguments would be lost in their practical significance on anyone who doesn't care about her future or on anyone who is a die-hard individualist is not damaging to the practical 'force' of those examples.

I believe there is much more that can be said about what I'm calling the practical significance of Kant's universality tests, which will have to wait for another occasion. But I hope that I have said enough to defend the main idea that even if Kant's universality tests do not represent self-sufficient tests of moral rightness, they are useful in helping to confirm the distinct authority of moral reasons and, in the context of moral reasoning, help reveal a kind of duplicity in an agent's attempt to justify immoral action.

V. Conclusion

I began with the question: 'suppose Kant's universal law formulation of the CI cannot be made to work as a moral decision procedure; would anything significant be lost if we were to jettison these formulations from Kant's overall moral theory?' My answer has been that there would be loss. First, the universal law formulation encapsulates a set of interconnected concepts that serve collectively to help constrain the available candidates for title of fundamental criterion of right action. Of course, we could re-write Kant's theory and preserve these concepts and not bother with the universal law formulations. And granted, for purposes of narrowing in on a criterion presupposed in common sense rational moral cognition, we could easily do so. However, as I have lately explained, these formulations do serve a practical function in Kant's ethics, even if they can't be made to function as a moral decision procedure.

Zusammenfassung

Kants Verallgemeinerungsformeln des Kategorischen Imperativs werden üblicherweise interpretiert als Ausdruck einer Art von an sich selbst hinreichendem Entscheidungsverfahren oder Test, den ein Handelnder verwenden kann, um auf vernünftige Weise zu Entscheidungen über seine moralischen Verpflichtungen zu gelangen. Gleichwohl ist dieses Testverfahren vielfacher Kritik ausgesetzt worden, wobei die wohl bekannteste die ist, es sei „leer“, da sich mit seiner Hilfe – im Unterschied dazu, was Kant sich von ihm erhoffte – kein reiches Arsenal an moralischen Verpflichtungen gewinnen lasse. Angenommen, diese Kritik sei zutreffend, so stellt sich doch die Frage, ob dieses pessimistische Verdikt über die Verallgemeinerungsformeln impliziert, dass diesen Formulierungen jede philosophische und praktische Bedeutung für Kants (oder zumindest eine Kantianische) Moral-

philosophie fehlt. Im vorliegenden Beitrag wird das Gegenteil behauptet. Im Hinblick auf die philosophische Signifikanz dieser Formulierungen wird vorgeschlagen, was hier als Auffassung des „formalen Rahmens“ bezeichnet wird und bedeuten soll, dass (1) die Ideen der Gesetzesähnlichkeit, der normativen Autorität und des Respekts von diesen Formulierungen impliziert werden, und dass (2) diese miteinander verbundene Gruppe von Ideen gemeinsam als formaler Rahmen dafür dienen, was man als substantielle und fundamental rechtssetzende Gründe für das Handeln bezeichnen kann, sowie dass (3) sie idealerweise zumindest dazu dienen, den Begriff der „Menschheit“ als fundamentales rechtssetzendes Merkmal für Handlungen herauszuarbeiten. Im Hinblick auf die praktische Signifikanz der Verallgemeinerungsformel wird die These vertreten, dass sie dazu dient, die Anerkennung der normativen Autorität moralischer Überlegungen im Unterschied zu Klugheitsüberlegungen durch den Handelnden zu entwickeln, und damit zugleich dabei hilft, eine Art von Doppelung im Denken von Handelnden zu entwickeln, die moralische Anforderungen verletzen.