teorema
Vol. XXIII/1-3, 2004, pp. 17-37

Moral Health, Moral Prosperity, and Universalization in Kant's Ethics*

Donald Wilson

RESUMEN

Usando un análisis de la distinción entre obligaciones perfectas e imperfectas sugerida en *La Metafísica de las Costumbres*, argumento que el imperativo categórico (IC) de Kant requiere que las máximas sean universalizables en el sentido de que puedan ser vistas como leyes universales consistentes con la integridad y el ejercicio efectivo de la acción racional. Este punto de vista, sostengo, tiene un cierto número de ventajas sobre la interpretación Korsgaardiana del IC basada en la contradicción práctica, tanto en términos del criterio de evaluación que Korsgaard usa como en términos de consideraciones textuales más amplias.

ABSTRACT

Drawing on an analysis of the distinction between perfect and imperfect duties suggested by *The Metaphysics of Morals*, I argue that Kant's Categorical Imperative (CI) requires that maxims be universalizable in the sense that they can be regarded as universal laws consistent with the integrity and effective exercise of rational agency. This account, I claim, has a number of advantages over Korsgaard's practical contradiction interpretation of the CI both in terms of the criteria of assessment that Korsgaard uses and in those of broader textual considerations.

I

In a previous paper[1] I argued that the ambiguity and continuing controversy surrounding various key notions in the *Groundwork* suggest there is room for a different approach to the interpretation of this central text in Kant's moral philosophy. Given ongoing disagreements about the meaning and scope of basic notions like universalization and the apparent impossibility of deciding between different accounts on the basis of the *Groundwork* discussion, I suggested that we hold these questions of interpretation in abeyance and look first to Kant's other works in moral philosophy for further insight.

The basic idea here is a simple one. Kant explicitly reminds us in the introduction to *The Metaphysics of Morals* [MM] that the CI remains the supreme principle of morality and tells us that this work is concerned with the application of this fundamental principle to the specific nature of human agency "to show in it what can be inferred from universal moral principles"

[MM, 6:217]. Insofar as they are meant to illuminate the practical implications of the formal account of morality set forth in the *Groundwork* [GW], more practical works like *The Metaphysics of Morals* ought then to be an important source of insight into this account. In particular, I suggested that we could work "backwards" from an account of the nature of perfect and imperfect duties arrived at by consideration of the *Metaphysics of Morals* to an account of the CI's requirement that maxims be universalizable.

This is the aim of the present paper. I begin with a brief review of the central elements of the account of perfect and imperfect duties that I defended in the earlier paper, and then show how this account can be used to inform the nature and scope of the CI's universalization requirement. Next, I discuss various implications of this account and consider it further in relation to four criteria suggested by Christine Korsgaard's defense of the practical contradiction account of the CI. Comparing it with this account, I claim that it has a number of advantages with respect to the kind of cases it can accommodate and the overall interpretation of Kant's view it offers.

II

When Kant turns to a discussion of particular duties of virtue owed to ourselves in the latter half of the Doctrine of Virtue, he distinguishes between (negative and formal) perfect duties and (positive and material) imperfect duties, suggesting that:

> The first belong to the moral health (*ad esse*) of a human being as object of both his outer sense and his inner sense, to the *preservation* of his nature in its perfection (as *receptivity*). The second belong to his moral *prosperity* (*ad melius esse, opulentia moralis*), which consists in possessing a *capacity* sufficient for all his ends, insofar as this can be acquired; they belong to his *cultivation* (active perfecting) of himself [MM, 6:419].

I argued previously that perfect duties are understood here as duties concerned with ensuring our moral health in the negative sense that they prohibit actions that compromise our capacity to govern ourselves on the basis of reason rather than self-love (our capacity for "inner freedom"). So, for example, the perfect duty prohibiting the drunkenness requires us to forgo this vice because it tends to undermine our capacity to use our powers purposively in thinking and deliberation in general and hence, indirectly, our ability to make specifically moral use of these powers (in reflecting on what morality requires and choosing to constrain our pursuit of interests of self-love accordingly).

The same kind of concern can be seen in the initially somewhat obscure duty we owe to ourselves prohibiting self-deception. The vice involved in

this case, I claimed, is best understood as a kind of insincerity or lack of conscientiousness in reflecting on one's actions and character. Kant thinks of moral deliberation as a kind of conversation with oneself in which a conception of our actions or character is submitted to our conscience for scrutiny and judgment. Thus understood, this process is vulnerable to a kind of special pleading on behalf of our desires to which Kant thinks we are particularly susceptible. He thinks that we tend to be powerfully drawn by interests of self-love and are given to complacency in thinking about the influence these interests have on our choices and the way we understand our lives [*Religion Within the Boundaries of Mere Reason* [RR] 6:37]. The danger then is that, preoccupied with our desires, we fail to impartially consider what we propose to do, and, instead, allow our desires to dictate the understanding of our choices and character that we submit to our conscience for judgment. This kind of insincerity or lack of conscientiousness in reflection and deliberation causes us to bring to our conscience a distorted account of our actions and our lives, threatening our capacity for inner freedom (and thus our moral health) at the deepest level by compromising the deliberative processes on which it depends.

Compared with these negative duties, imperfect duties associated with concerns with ensuring our *moral prosperity* or *flourishing* are regarded as duties focused, more positively, on promoting the *effective exercise* of this kind of inner freedom in choice and action. So, for example, Kant identifies imperfect duties requiring us to cultivate talents and abilities that will make us more effective as agents in general (and hence more effective in securing ends we seek on the basis of duty), and to perfect our moral character by actively cultivating our receptiveness to the voice of duty and our ability to follow through properly on commitments we set for ourselves on this basis.

According to the analysis I have suggested, then, perfect duties prohibit vices that tend to compromise our capacity for inner freedom[2] and hence our ability to function as rational agents, and imperfect duties require the cultivation of virtues facilitating the effective exercise of rational agency. The division of duties here thus serves to identify the *distinctive moral problems* associated with maxims contrary to these duties and, in doing so, can be used to illuminate the nature of the CI's universalization constraint. Whether we think that the CI is intended to serve as a guide to individual moral deliberation or, as Barbara Herman suggests, as a procedure yielding a series of standing deliberative presumptions, the fact that a maxim fails to meet the universalization requirement the CI imposes is taken to indicate that it is morally problematic. Given an independent account of the form of moral problem associated with perfect duties like the one I have defended, it ought to be *this feature* of a maxim that universalization calls to our attention. That is, the sense in which a maxim is said to be contradictory ought to be such that it indicates it has *this problematic feature* and not some other, *since it is this characteristic in virtue of which it is wrong*.

Having arrived at this account of the distinction between perfect and imperfect duties by considering the division of these duties in *The Metaphysics of Morals*, we can therefore use it to inform the nature of the contradictions said to be apparent in maxims contrary to duty on universalization. In particular, if the distinctive moral problem associated with actions contrary to perfect duties is that they tend to compromise our ability to function as rational agents governing ourselves on the basis of reason rather than self-love, and if this is what universalizing a maxim is meant to call to our attention, this suggests that maxims that fail to meet this requirement are those that are seen on universalization to be incompatible with some necessary condition(s) of rational self-government. So, for example, drunkenness is regarded as contrary to perfect duty on the grounds that this vice tends to compromise the agent's capacity for the kind of reflection and deliberation basic to rational self-control. Thinking of this as the aspect of the maxim called to our attention in universalization, a maxim of using intoxicating substances to the point of stupefaction merely for enjoyment fails to meet this requirement because it involves willing the general practice of conduct incompatible with the exercise of rational agency.

If this is what universalization is meant to emphasize, this suggests that the CI requires maxims be universalizable in the sense that they can be regarded as universal laws consistent with the integrity of rational agency (and hence capable of serving as laws governing the choice of all rational agents in pursuing their ends). A maxim that fails to meet this requirement will be contradictory if it is seen on universalization to be incompatible with some condition of rational agency, and the contradiction at issue here will relevantly be one *in conception* because, as rational agents, we cannot conceive of a world in which such a maxim obtains as a universal law. So, for example, finite and dependent human beings whose rational agency depends upon the proper functioning of their physical bodies cannot, as rational beings, conceive of a world in which this functioning is routinely compromised, and hence cannot conceive of universal laws enjoining intoxication for pleasure, suicide, self-mutilation, etc.

A similar line of argument suggests a particular understanding of the kind of contradiction in the will associated with maxims contrary to imperfect duties. Imperfect duties are understood here as duties concerned with ensuring our moral prosperity and therefore with promoting the flourishing (as opposed to the possibility) of rational agency. The wrong associated with maxims contrary to these duties thus consists in the tendency of certain policies to compromise the effective exercise of rational agency and, as above, we ought to expect this aspect of the maxim to be made manifest on universalization. Interpreting the sense of the universalization constraint applied in this case in terms of this account, a maxim contrary to imperfect duties fails on universalization on the grounds that it is incompatible with the effective

exercise of rational agency. Conversely, a maxim we can consistently will as a universal law is one that can be seen to be compatible with the conditions under which individuals can effectively identify and pursue ends enjoined by reason.

Here the contradiction is therefore one "in the will" in the sense that although we can conceive of a world in which the effective exercise of rational agency is thwarted, we cannot, as rational agents, consistently will that such a world exist. Thus, for example, while rational agency would still be possible in a world in which non-beneficence was the norm, such a world would be one in which the effective exercise of rational agency would be compromised, given the fact of our natural limitations and our resulting dependence on others for aid in securing our ends.

III

I propose, therefore, that the contradictions in conception and in the will associated with the CI's universalization requirement should be understood as contradictions arising in virtue of our willing maxims incompatible with the *integrity* and *effective exercise* of rational agency. Thus understood, willing in accordance with the integrity and effective exercise of rational agency is regarded as constitutive of properly rational willing. Kant thinks of rational agents as beings with wills capable of originating and initiating choices of action. In order for this to be possible, he claims that the will must be regarded as governed by laws of its own making, and, since these laws cannot be informed or constrained by any end or authority external to the will, there is nothing other than the form of law itself —universalization— to which they can be answerable. The CI is therefore represented as the supreme law or determining ground of a rational will in virtue of which it can be regarded as a "first cause" of events in the world, and, according to the analysis I propose, this principle requires that the maxims on which agents act be consistent with the possibility and effective exercise of this kind of self-directed agency in general.

This fundamental principle of practical reason will always govern the willing of a purely rational agent but, familiarly, becomes a contingent commitment in the case of finite and dependent beings subject also to the influence of interests of self-love. Such beings have individual needs manifested through the operation of some mode of sensibility that are therefore experienced as *particular to them*. Influenced by desires associated with these needs, and capable of understanding and prioritizing them, dependent rational beings have a conception of their own good. Although they are still subject to the constraints on practical reasoning associated with the CI, dependent rational agents like this may therefore choose to act for the sake of self-love at the expense of the requirements of pure practical reason. Importantly, this is not to be understood as a contingent fact about their agency. Finite and de-

pendent rational agents cannot transcend the limits of their nature and come to will as a purely rational agent does. Instead, rational but also natural beings are always subject to competing influences on their choices of ends and therefore realize their rational nature in the exercise of conscious self-constraint (or inner freedom) through which they choose to govern these other interests under the auspices of their reason.

This fact of their nature and their vulnerability to various kinds of injury and impairment will then be reflected in the duties associated with this view. These duties will encompass the kind of concern with the integrity and proper functioning of our natural organism seen in the prohibitions on drunkenness and gluttony considered above, and will extend also to a range of duties prohibiting suicide, murder, and other acts of violence against persons on similar grounds. In addition, given that finite and dependent rational agents realize their rational nature in conscious acts of self-constraint, this analysis will yield duties concerned with ensuring the capacity of individuals to manifest and exercise this kind of rational self-control. Included here, I have argued elsewhere, will be the duties prohibiting the vices of servility, miserly avarice, and ridicule that Kant discusses in the *Metaphysics of Morals,* and a central and important requirement of honesty both with oneself and with others.

In particular, a concern with the integrity of others' rational agency will include a basic and robust concern with what we might think of as their natural autonomy. We do not realize our rational nature if all we do is act in conformity with the requirements of reason on the basis of self-love (because we desire the approval of others, seek some other end, etc.). We do so only through a self-conscious choice to govern ourselves on the basis of our reason — one that can be said to reflect our own choices of values and priorities and a personal commitment to taking the fact that reason requires something of us to give us sufficient reason to choose and act accordingly. A concern with the integrity of others' rational agency will therefore require a broad concern with their general freedom to adopt, pursue, and revise, ends and interests on the basis of their own conception of their value (such that they may come, in time, to make this kind of commitment for themselves), and a series of concerns with preserving a social environment conducive to this kind of self-development.

There are a number of other significant implications of this analysis, some of which I want to mention before evaluating it further. Firstly, notice that when the CI is understood in this way, the terms of appropriate maxim description will be set by the underlying concern associated with the universalization constraint. So, for example, it will not be relevant that one proposes to get drunk on a Tuesday, or, ordinarily, that by doing so one may gain some weight. Instead, the relevant considerations are identified through the background concern here with the integrity of rational agency. An appropriate description of the maxim in this case will thus reflect the fact that one

intends to use a debilitating narcotic substance in the pursuit of pleasure (or release from cares, etc.), and there will be an important distinction between a maxim like this and one enjoining the use of similar substances as anesthetics or to alleviate debilitating pain. Similarly, given the kind of concern required with the natural autonomy of others described above, the act of mugging someone will be appropriately understood as an act involving the attempt to forcibly appropriate another's property for one's own use, and not, say, as an act meant merely to alleviate them of the weight of their wallet.

For similar reasons, it is not ordinarily going to be the particular desire or end of self-love that the agent seeks to realize that will give rise to a contradiction in their maxim. If, for example, false promising is wrong, this will be because a universal maxim of false promising is seen to be incompatible with the integrity of rational agency. This will be a matter of the role of the practice of promising in ensuring the possibility of rational agency, and, if a case is made that this is a necessary institution, its forms and limits will be understood in relation to the role it plays in doing so. Any excusing conditions that are allowed will reflect this underlying concern, but the impermissibility of false promising will be understood in these general terms. As such, false promising will be regarded as presumptively wrong, without this being seen as a function of the particular ends of self-interest that agents might have in mind in seeking to exploit the practice for their own benefit.

Notice, therefore, that this analysis tends to suggest the kind of role for the CI envisaged by Barbara Herman in *The Practice of Moral Judgment*.[3] Herman argues that the CI ought to be understood as yielding *principles of deliberative presumption* intended to guide routine moral judgment and to inform the terms of occasional moral deliberation. Thus understood, the CI is applied to generic maxims of self-interest that recommend certain action types in pursuit of interests of self-love, and the outcome of its application are standing presumptions that serve routinely to limit the pursuit of these interests. To the extent, then, that concerns with ensuring the integrity and effective exercise of rational agency will also focus on action types and on the pursuit of ends of self-interest in general, the present analysis suggests this kind of role rather than the role associated with the moral deliberation model (according to which the CI is to be used directly to test the permissibility of the particular maxims on which agents take themselves to be acting).

The deflationary approach to the problem of maxim description taken on this account and the different, more general, role assigned to the CI, should not, however, be thought of as objectionably marginalizing the role of this fundamental principle. Despite the volume of secondary literature devoted to discussion of a CI *procedure*, there is little textual evidence to support the view that Kant meant the discussion of *Groundwork* examples to herald the introduction of a distinctive procedure intended to directly assess individuals' maxims. The second section of the *Groundwork* continues the

project of identifying the fundamental principle of morality begun in the first, and, as others have noted, the discussion of examples here can plausibly be understood as intended simply to lend support to the claim that the CI is this principle (by showing how it underlies a series of canonical moral obligations). More generally, a distinctive procedure of the sort envisaged by the moral deliberation model is conspicuous only by its absence in Kant's subsequent discussions of duties in the *Metaphysics of Morals* and elsewhere. What is clear, and what he does directly claim, is that the CI is meant to represent the most formal expression of the moral law and that it is the ground of all moral obligation. This claim, however, is compatible with the account I propose.

Alternatively, it might seem that the present analysis marginalizes the centrality of the CI in another way. It might be thought that universalization is unnecessary on this account and that we could just forgo this stage and think more directly about harms to individuals' agency. After all, we do not seem to require any elaborate procedure or formal notion of universalization to realize that drunkenness interferes with the capacity for rational deliberation. It is important, therefore, to emphasize the formal nature of the required concern with ensuring the integrity and effective exercise of rational agency. The suggestion here is not that drunkenness or false promising are wrong insofar as they can be shown to have some actual or probable debilitating effects on identifiable individuals in given cases. The claim is that maxims involving these sorts of actions cannot serve as universal laws for rational agents pursuing ends, because, as such laws, they would undermine the integrity and effective exercise of the rational agency they assume as purposive and autonomous acts of will.

In this sense, then, it is the form of our maxims that is at issue, rather than whether or not the harm associated with particular prohibitions occurs. Maxims like these are wrong because they are seen to be incompatible with universal laws governing the rational pursuit of ends (and hence in virtue of their form), without regard to whether the harm associated with them will actually accrue to any given violation. Although we can think of the CI's universalization constraint informally in terms of the harms associated with it, then, universalization remains the definitive measure of permissibility.

Notice, finally, that the proposed analysis affords us a straightforward account of Kant's claim that the various different formulations of this principle are equivalent. As finite and dependent beings, human agents are subject to the competing influence of interests of self-love and act on the basis of a conception of the value of different ends and hence on subjective principles (maxims) that may or may not be universalizable. Representing universalization as an unconditional constraint requiring that choices of maxims be capable of serving as universal laws consistent with the possibility and effective

exercise of rational agency, the CI represents rational agency as having a *non-relative value* we ought to respect in choice in action.

Reflecting this aspect of the CI, the formula of humanity represents rational agency as an "end in itself," requiring that choices of maxims do not "conflict" and actively "harmonize" with this end in the sense the universalization constraint requires. Similarly, thinking of this constraint from the different perspective of an idealized community of rational agents acting under the auspices of practical principles endorsed by the CI, we can think of it in terms of a "kingdom of ends": a "whole of ends in systematic conjunction" in which self-legislating autonomous agents pursue diverse ends in community with others under laws allowing the freedom and agency of all. Thus understood, the subsequent formulations of the CI then serve, as Kant suggests, to bring its more formal universalization requirement "closer to intuition […] and thereby to feeling" [GW, 4:436] for sensuously affected human agents acting from a conception of the value of ends.

IV

I argued elsewhere that the analysis of perfect and imperfect duties associated with this view is a promising and subtle one. What I propose to do in the remainder of this paper is focus on the adequacy of this account as measured in terms of four criteria suggested by Christine Korsgaard in her defense of the practical contradiction account in *Creating the Kingdom of Ends*. Korsgaard notes that the textual evidence for any model of the CI is ambiguous and suggests that since "no interpretation can be based on textual evidence alone," we should evaluate different accounts of the CI in terms of four questions:

> (i) [W]hat kinds of cases it can handle, (ii) whether it can meet some standard objections, (iii) what sort of distinction between the contradiction in conception and the contradiction in the will test is implied by it, and, most importantly, (iv) what presuppositions about rationality it makes and so what kind of case it will allow Kant to make when he turns to the critical project of showing that morality *is* pure rationality [Korsgaard (1996), pp 80-1].

She argues that her interpretation is better able to deal with these questions than the rival logical and teleological interpretations she considers, and suggests that it ought to be preferred on this basis. She claims that the practical contradiction interpretation is superior to the logical and teleological accounts in that it can deal with at least some cases of natural actions involving the use of violence to secure one's ends. Continuing the same general theme, she argues that the practical contradiction account is better able to avoid some

standard objections and that it has the advantage of offering a unified account of the contradictions associated with the CI. Lastly, she claims that the practical contradiction account is readily associated with the kind of rational contradiction familiar to us in maxims that violate the hypothetical imperative and will therefore serve us well when it comes to understanding the later arguments in the third section of the *Groundwork*. Using these same criteria, I will argue that the account I suggest can claim similar virtues and has a number of important advantages over the practical contradiction interpretation seen, in particular, in the kind of examples it can accommodate and the way in which it connects the ideas of contradictions in conception and in the will.

Korsgaard regards the inability of accounts such as the logical contradiction account to deal appropriately with examples of natural actions as their central failing and claims it is an important advantage of her account that it can accommodate some of these cases. According to the practical contradiction account, the idea of a contradiction in conception is to be understood in terms of the practical efficacy of maxims when considered as universal laws governing the pursuit of chosen ends. A maxim whose form is such that the agent would no longer be able to secure her desired end in the world of universalization is said to be practically self-defeating. So, for example, if we imagine an agent proposing to make a false promise in order to secure a loan and apply the universalization test to their maxim, we see that the agent would be willing a world in which everyone desiring money is free to make promises like this. In such a world, no one would take the act of promising as a sign of the future performance of some action, and promising would therefore cease to be efficacious as a means to the agent's original end. A maxim like this, it is claimed, cannot be conceived of as a universal law and is therefore to be rejected as prohibited by perfect duty.

Applying this model to the case of natural actions, Korsgaard considers an example in which an agent is thinking about killing a competitor in order to get a job. She argues that the maxim in this case is self-defeating in the relevant sense because one would be willing a world in which a success condition necessary to obtain one's end (that the individual stay alive to get the job) would not be assured [Korsgaard (1996), p. 41] and that her account can therefore prohibit at least some natural actions of this sort.

Having claimed this important virtue, however, she immediately seems to undermine it. She criticizes the logical contradiction account on the grounds that it cannot classify prohibitions on natural actions as perfect duties, and is therefore left in the embarrassing position of classifying them as perfect duties *of virtue* which cannot be prohibited by law [Korsgaard (1996), pp. 83-4]. She concedes, however, that her own account also cannot deal with a broad range of acts of violence where the end sought is one that can be secured directly through the action without requiring the continued existence of the agent undertaking it (e.g., cases of killing for pleasure, revenge, etc.). It

seems, therefore, that the best we could hope for in a range of common cases like these is classification as some other kind of duty, with the same problems this brings for the logical contradiction account.

She also explicitly acknowledges that her account will not include a perfect prohibition on suicide for the same general reason. Though not an uncommon omission in accounts of the CI, this should, I think, be seen as a serious problem. Unless we reject completely the idea that we have obligations owed to ourselves, it is surely not implausible to think that we might have a duty to refrain from self-murder or mutilation, at least in a broad range of normal situations.[4] It is also plausible to suggest that the duty prohibiting suicide is meant to rule out an instance of precisely the kind of natural action under discussion here. There is no reason why Kant would think of self-murder as categorically different from the murder of another. If killing a rational agent in one's own person in order to reduce discomfort is morally wrong, surely killing another person for similar reasons must be wrong also, *and be wrong for the same general reason(s)*.

These considerations suggest that the suicide example ought to be seen as a relevant case, and that an understanding of it should afford us some insight into Kant's view about the more general case.[5] Moreover, the suicide example is one of only two paradigm examples of perfect duties that Kant cites in the *Groundwork.* It is also one that he clearly intends to be taken seriously, since he revisits it later in the *Metaphysics of Morals* and continues to identify it specifically as a perfect duty. It seems, therefore, that any account of the CI intended as an interpretation of *Kant's* view ought to take this duty seriously, and, if it cannot be accommodated, ought to offer some explanation of why Kant could have been mistaken in applying his own universalization constraint in this example.

The account I propose avoids all of these problems by affording us a plausible and direct understanding of this case and treating it, appropriately, as a particular instance of a more general moral issue. Making the value of rational agency central, the suggestion here is that self-murder and self-mutilation are contrary to perfect duties owed to oneself because acts of this sort compromise the integrity and proper functioning of the physical organism in which our rational agency is instantiated. As such, the same will be true of acts of murder and violence perpetrated against others, which will be seen as wrong for the same general reasons.[6]

Unlike the practical contradiction interpretation, then, the account I suggest can accommodate the kind of perfect prohibition on suicide that Kant seems to have in mind and treats it, appropriately, as a particular instance of a broader range of problematic actions. Notice, in particular, that when we understand the CI's universalization constraint in this way, our various appetites, practices, etc., can be regarded for the purposes of moral assessment as functionally organized around the preservation and promotion of rational

agency. This allows us to understand both Kant's references to the natural purposes of self-love in the *Groundwork* discussion of suicide, and, more generally, his occasional qualification of the CI's universalization constraint as one requiring that a maxim be fit to be a universal law *of nature*.

In the case of finite agents with various bodily needs that must be met, the purpose of a natural interest in their own well-being (the feeling of self-love) can be understood in terms of its role in serving to prompt a care for the integrity and health of our natural being (and hence, as Kant puts it, "the furtherance of life"). Thus understood, the problem with the suicide's maxim of willing "from self-love" to make it her principle to end her life whenever "its longer duration threatens more troubles than it promises agreeableness" is then that, when universalized, it involves willing that we would have a basic nature "whose law it would be to destroy life by means of the same feeling whose destination is to impel towards the furtherance of life" [GW, 6:422]. This would be a basic nature in which rational agency, so constituted, would be compromised by the routine neglect of our bodily needs in favor of the pursuit of pleasure and hence one that we cannot conceive of as a universal law in the relevant sense. Notice also that the duty prohibiting suicide will then be broader on this account, in just the way that Kant envisages in *The Metaphysics of Morals*' discussion of this vice. Here suicide is described as encompassing concerns with "partial" suicide and requires us to refrain from "mutilating" ourselves "materially" (by "*depriving* oneself of certain integral, organic *parts*") or "formally" (by "*depriving* oneself (permanently or temporarily) of one's *capacity* for the natural (and so indirectly for the moral) *use* of one's powers") [MM, 6:421].

Continuing her evaluation of the practical contradiction approach, Korsgaard considers two standard objections originally associated with Hegel, but which are, she notes, by now relatively familiar in the literature on Kant in one form or another. According to the first of these objections, the universalization requirement imposed by the CI is empty because, while there might be contradiction apparent in willing both that we be able make a false promise and that there could be no such thing as promising, there is no contradiction in willing that there be no such thing as promising. As such, it is suggested, Kant simply assumes that everyone will desire to be able to make promises, and it is this background assumption that makes the promising example, and others like it, work. The second objection claims, instead, that the universalization test is too strong. If it is contradictory to will a maxim of false promising as a universal law, it must, it is suggested, also be contradictory to will a universal law requiring us to succor the poor (since if everyone did so, there would be no poor to succor).

Korsgaard argues that both the logical and practical contradiction accounts can avoid these problems relatively easily. The proponent of the logical contradiction interpretation can reply by arguing that the contradiction in

the first case is found in the fact that the agent proposes to make a promise under conditions in which there could be no such thing as promising (and not, therefore, merely in the idea of a system that lacks an institution of promising). In the second case, she suggests, they can respond by arguing that the maxim involved is misstated. It should be understood as a principle enjoining us to succor those in need, and there is nothing inconceivable about this maxim when we imagine a universalized world in which everyone aids others, and, as a result, no one is in need.

Her own account, she continues, also easily avoids these objections and does so in a more direct manner. In the first case, she responds by emphasizing the practical nature of the contradiction involved here, and noting that the agent's commitment to willing there be a practice of promising is *already directly assumed* in their willing to make use of this institution for their own purposes. In the case of the second example, she claims that her account responds "both readily, and in an obvious way, correctly" [Korsgaard (1996), p. 95]. Understanding the agent's purpose as being to bring relief to the poor, there is, she claims obviously no practical contradiction in a world in which this purpose has been fully realized and hence no contradiction in universalizing a maxim enjoining us to aid others on this basis.

The account I propose can also deal with these problem cases, but does so in an importantly different way in the case of the first objection. As limited and finite beings, our knowledge of the world is always fallible and often partial or mistaken, and we cannot rely on being able to secure our various needs and interests by our own efforts. These limitations are not contingent; they constitute the natural context of our agency within which we must learn to function and have to be appreciated and allowed for in identifying and pursuing a conception of one's good and the various particular ends it recommends. The interest in the capacity of others to order and control their own lives that I have argued will be integral to a concern with the integrity of their rational agency will therefore extend to a concern with the capacity of agents to control for these limitations as much as is possible by, among other things, being able to enter into the kind of cooperative agreements with others characteristic of institutions like promising, secure in the expectation that agreements like this are ordinarily reliable.

On this kind of account, then, the purpose and norms of the practice of promising will be understood in relation to the need for cooperation with others, and ordinarily reliable access to this convention will be regarded as necessary on the grounds of concerns with the integrity of rational agency in finite and dependent human agents. Thus understood, the problem with willing a universalized maxim of false promising in pursuit of one's interests is not then that it would make the realization of the agent's (or any) particular end impossible. Instead, the suggestion here is that a universalized maxim of

false promising in pursuit of our ends would vitiate the morally necessary practice of promising and that it is contradictory for this reason.

Thus understood, this account responds to the first kind of example by arguing that the contradiction seen in the world of a universalized maxim of false promising *is* to be found precisely in the idea of willing the conditions under which the practice would cease to exist. The commitment to the practice, however, is not merely covertly assumed or thought of simply as a shared desire. Instead, promising is seen as an institution required on the basis of concerns with the integrity of rational agency in the case of limited and finite beings like ourselves. In the case of the second kind of example, the reply will be similarly direct. A commitment to providing for the basic needs of others and for their happiness in general will be directly required on the basis of concerns with ensuring the integrity and effective exercise of rational agency. The commitment to aid others is therefore again not tied to there actually being identifiable others in need of aid, and hence there is no contradiction in a world of universalization in which no one actually needs aid.

Korsgaard goes on to argue that her account is superior to the two others considered in that the contradictions in conception and the will associated with it can be seen as connected with one another and with a broader account of practical rationality. She suggests that there are certain ends that "belong essentially to the will" and that it is contradictory to will the conditions under which these ends cannot be realized. In particular, she suggests that as agents acting for ends, we must will our "general effectiveness" in securing ends and our freedom to adopt and pursue new ends. Taking this to be the kind of contradiction in the will evident in willing maxims contrary to imperfect duties, she suggests that we cannot consistently will universal laws enjoining indifference to our natural talents or the needs of others. Thus understood, she claims, this second sense of contradiction can be seen as similarly practical and hence as connected with the idea of a (practical) contradiction in conception.

However, while her account allows us to view the contradictions at issue as being connected with one another, the similarity she cites makes it difficult to regard the distinction between these classes of duties as a *substantive one* marking a difference between more and less serious duties. That Kant intends this distinction in this way is suggested both by his depiction of it as involving a distinction between actions "opposed to strict or narrower (unremitting) duty" and those "only to wide (meritorious) duty" [GW, 4:424], and by the kind of duties he includes in each class (suicide and false promising in one, and self-development and beneficence in the other). It seems plausible, therefore, to suggest that an account of the kind of contradictions associated with maxims contrary to these different classes of duty ought to reflect this difference in significance and this does not seem to be the case in Korsgaard's treatment. Contradictions in conception are associated with the failure of a maxim to be efficacious in securing the agent's particular end when

universalized and those in the will with a more general failure of agency associated with the thwarting of some essential purpose. Thus understood, it is difficult to see why one class of practical contradiction would be more problematic than the other, and hence difficult to see this distinction as tracking the kind of difference in moral significance between classes of duty that Kant has in mind.

In contrast, in the account I propose, contradictions in conception and in the will are connected with one another in a way that suggests an obvious difference in seriousness between these classes of duty with which they are associated. Contradictions in conception arise in relation to our willing maxims incompatible with the *integrity* of rational agency and those in conception when we will maxims that would undermine the *effective exercise* of this kind of agency. Thus understood, acts of violence, etc., are prohibited as contrary to perfect duty, and, understanding perfect duties as duties concerned with ensuring the *possibility* of rational agency and imperfect duties as obligations concerned with promoting the *flourishing* of this agency, it is plausible to view the former as more stringent.

The last issue that Korsgaard raises is much harder to evaluate. The suggestion in this case is that interpretations of the CI need to be responsive to the problems posed by Kant's attempt to motivate the idea that universalization is a requirement of reason (and hence to establish the CI as the fundamental principle of morality) in the third section of the *Groundwork*. Korsgaard approaches this issue by first connecting the practical contradiction she describes with the kind of contradiction associated with choices that would violate the hypothetical imperative. Thinking of the will as a kind of causality, she suggests that an agent willing an end sees herself as initiating a causal chain that will bring it about. In this sense, "willing the end contains, or insofar as you are rational is already, willing the means" [Korsgaard (1996), p. 94], and it is therefore contradictory to will the realization of an end but not the means necessary to realize it.

Using the false promising example, she then argues that the same kind of contradiction is seen in the practical contradiction account when a maxim is said to yield a contradiction in conception:

> In the world of the universalized maxim, the hypothetical imperative from which the false promiser constructs his maxim is no longer true. It was "if you want some ready cash, you ought to make a false promise." But at the same time that he employs this hypothetical imperative in constructing his maxim, he wills its falsification, by willing a state of affairs (the world of the universalized maxim) in which it will be false. In that world, false promising is not a means to getting ready cash. Kant, therefore, not only has a specifically practical sense of "contradiction," but should be seen as employing it in his contradiction tests [Korsgaard (1996), p. 94].

She claims, therefore, that her account has the advantage of drawing on a familiar kind of contradiction associated with the hypothetical imperative and proceeds to argue that the model of practical rationality associated with it promises to be fruitful when it comes to trying to understand the third section of the *Groundwork*. Drawing a parallel between theoretical and practical reasoning, she begins with the idea that, "the rational will models its conception of a practical law on that of a causal law" [Korsgaard (1996), p. 102]. A rational being may take two events to be causally connected, but it is only connections like this that hold universally that qualify as causal laws. Similarly, a rational will adopting and pursuing ends, takes the connection between an end and the means to realizing it to give it a reason to act in some way. If, however, this connection is to have the requisite form of a practical law, it must hold universally, and this, she continues, is exactly what the practical contradiction interpretation tests. In universalizing our maxim we imagine a world in which everyone who seeks our end(s) acts in the manner we propose. If the means we intend to employ can be adopted universally without defeating our original purpose, our maxim can then be said to have the requisite form of a universal practical law.

Reasoning in this way, Korsgaard thinks that we can connect autonomous rational willing and willing in conformity with the CI through a relatively familiar and uncontroversial conception of practical rationality, and, by doing so, hope to make some headway in understanding and motivating the argument Kant makes in the last section of the *Groundwork*. The central problem here stems, however, from the admitted limitations of the practical contradiction account. As we saw in the discussion of natural actions, this account cannot deal with a broad range of cases in which the ends sought by violence to oneself or others can be realized without requiring the continued existence of the perpetrator. Korsgaard suggests that cases like these are readily dealt with by the formula of humanity and attributes the problems they pose here to the fact that Kant tends to emphasize more everyday cases of moral temptation (making it difficult to decide how to approach cases where the agent has a "diseased purpose" or those in which she has to respond to the evil actions of others). The explicit admission that the practical contradiction account cannot handle important cases like these must, however, surely be regarded as significantly undercutting the advantage claimed for this account above. It will avail us little if we are able to understand the claimed connections between rationality, autonomy, and the CI, only by limiting the scope of this supposedly fundamental principle in a way that entails it cannot directly accommodate a broad range of familiar and obviously immoral actions.

In contrast, the account I propose is able to capture the full range of cases of natural actions and does not limit the scope of the CI in this way. Moreover, when we recall Korsgaard's treatment of contradictions in the will, the claim to simplicity and metaphysical economy in the account of

practical rationality seems overstated, suggesting that her account ought not to be preferred simply on this basis. According to her account, a maxim yields a contradiction in the will when it is found on universalization to be incompatible with some necessary ends that rational agents are committed to willing. Thus understood, the practical contradiction in this case is one step removed from the more maxim-specific contradiction seen in a contradiction in conception. Instead of being directly self-defeating relative to the choice of some specific end, a contradiction in the will arises when we will conditions under which we would not be free to adopt and revise ends or those in which our effectiveness in realizing them would be compromised. As such, this account is also committed to an account of practical reason that regards willing in conformity with the conditions under which we are capable of exercising agency in general and of doing so effectively as constitutive of fully rational willing. That being the case, there seems little to distinguish the two accounts in terms of their respective metaphysical commitments and the other advantages accruing to the integrity of agency account seem to favour it.

V

The analysis of the CI I propose therefore has a number of advantages when assessed in terms of Korsgaard's criteria. Like the logical and practical contradiction accounts, it can deal with the problematic Hegelian examples, but, unlike these other accounts, it accommodates the prohibition on suicide and other cases of natural actions involving violence done to others. In doing so, it classifies natural actions like these as being prohibited by perfect duties and hence distinguishes appropriately between important prohibitions like these and intuitively less serious duties requiring us to develop our talents, show gratitude to others, etc. Like the practical contradiction account, it regards contradictions in conception and the will as being connected with one another and with a broader account of practical rationality. Unlike Korsgaard's account, however, the analysis I suggest distinguishes between these different senses of contradiction in a way that allows us to regard them as reflecting a difference in the casuistical significance of the classes of duty associated with them. Lastly, while an adequate discussion of the arguments Kant develops in the last section of the *Groundwork* is clearly beyond the scope of the present paper, Korsgaard's claim that the practical contradiction account affords us a natural reading of the central connections between reason, autonomy, and the CI that Kant seeks to establish in this section seem overstated and there is no reason to think that her account is obviously to be preferred on the basis of metaphysical economy.

In addition, although Korsgaard rightly emphasizes the difficulties of defending an interpretation on textual grounds, there are nonetheless some

textual considerations that seem relevant here that I want to mention in closing. I have already suggested that my account affords us a straightforward reading of the claim that the different versions of the CI are equivalent and an explanation of Kant's references to the purposes of our practices and natural drives. It is also a much more inclusive account, capable of capturing the range of duties that Kant discusses and of explaining the relationship between the different categories that he identifies in the *Groundwork* and *The Metaphysics of Morals*.

When comparing different models in terms of the kinds of cases they can accommodate, Korsgaard focuses on the issue of general plausibility and hence on the failure of the logical contradiction account to deal appropriately with natural actions. It is, however, also reasonable to suggest that an account of the CI's contradictions in conception and in the will ought to be able to accommodate at least some central cases of perfect and imperfect duties that Kant explicitly mentions, and here the more standard accounts considered again seem wanting. As we have seen, both are unable to include the *Groundwork* prohibition on suicide in the class of perfect duties, and, more generally, are clearly not going to extend to the broader range of perfect and imperfect duties that Kant discusses in the *Metaphysics of Morals* prohibiting vices like self-deception, miserly avarice, and servility. In contrast, the analysis I propose begins with this broader account of perfect duties and hence includes them, and, as we have seen, naturally encompasses both of the *Groundwork* examples.

More generally, the analysis I propose also affords us a unified account of the different classes of duty that Kant identifies. Having warned us in the *Groundwork* that he intends to "reserve the division of duties entirely for a future *Metaphysics of Morals*" [GW, 4:421] Kant divides this later work into two main sections: the Doctrine of Right concerned with external duties of right and with the proper form and limits of civil society; and the Doctrine of Virtue emphasizing the development of good moral character and duties of virtue. In doing so, however, he explicitly retains the older division (classifying duties of Right in their ethical aspect and some duties of virtue as also perfect duties), making it hard to understand what the relationship between these various classes of duty is supposed to be. The present analysis allows us to see these different classifications as connected and explains the order and emphasis of Kant's discussion of duties in the *Metaphysics of Morals*.

The "sole aim" of the *Groundwork* is "to seek out and establish the supreme principle of morality" which Kant tells us must "be looked for not in the nature of man nor in the circumstances of the world in which he is placed, but solely a priori in the concepts of pure reason" [GW, 4:389]. The division of duties into perfect and imperfect duties in this work thus represents a formal division of duties associated with the two ways in which maxims may fail to conform to the requirements this fundamental principle imposes. In

contrast, the later *Metaphysics of Morals* assumes the authority of the CI and is concerned, instead, directly with its practical application to the particular circumstances and nature of human agency.

Understanding the CI in the manner I suggest, the order and emphasis of this later work can then be seen simply as a practical application of the *Groundwork's* formal concerns with the integrity and effective exercise of rational agency. Concerned with the integrity of rational agency in human beings who realize this agency in leading lives of conscious self-constraint, *The Metaphysics of Morals* begins with a discussion of duties of right concerned with ensuring the external freedom necessary if we are to be able to come to make this kind of commitment for ourselves, and the form and limits of a rightful civil condition that makes this self-development possible. Similarly, turning later to a consideration of specifically ethical obligation in the Doctrine of Virtue, the discussion is again oriented around practical concerns with ensuring the integrity and effective exercise of rational agency arising in virtue of the particular nature of human agency. Given our susceptibility to the influence of interests of self-love, determining ourselves on the basis of our reason requires commitment and effort on our part and Kant therefore singles out duties of virtue concerned with our developing and sustaining the strength of character necessary if we are to resist temptation and overcome our natural frailties and limitations.

In conclusion then, I have argued that the CI should be regarded as a formal principle requiring that maxims be compatible on universalization with the integrity and effective exercise of rational agency. Thus understood, willing in conformity with this requirement is regarded as internal to fully rational agency but the precise nature of the duties required of us will take into account basic facts about the nature of human agency and existence. This account, I have claimed, has a number of advantages over Korsgaard's practical contradiction account both in terms of the criteria of assessment she uses and in those of broader textual considerations. While there is obviously a great deal more that needs to be said here, a case has, I think, therefore been made that this analysis merits further consideration.

Department of Philosohy
Kansas State University
201 Dickens Hall
Manhattan, KS 66506, USA
E-mail:dywilson@ksu.edu

NOTES

[*] All references to Kant are to *Kants gesammelte Schriften, herausgegeben von der Deutschen Akademie der Wissenschaften* (1902), Berlin: Walter de Gruyter. Translations and abbreviations used are cited in the reference section.

[1] *Kant on Moral Health and Moral Prosperity*, unpublished manuscript.

[2] For an informative and detailed discussion of inner freedom and its central role in Kant's treatment of virtue, see Stephen Engstrom's *The Inner Freedom of Virtue* [Timmons (2002), pp. 289-316].

[3] It should also be noted that although the proposed analysis of contradictions in conception and in the will resembles Herman's account of these notions, it is not simply equivalent to it. According to my interpretation, concerns with the integrity of rational agency encompass a much broader range of concerns than Herman's more formal notion of separateness with its emphasis on acting on reasons that go "all the way down."

[4] It would seem, in particular, very odd to recognize a duty to develop one's talents without also thinking we have a duty to refrain from self-injury and that this duty is the more basic of the two.

[5] This is not to suggest that Kant thinks of self-murder as the central case from which the wrong of killing others is *derived*. Only that, however the account is understood, there is reason to think that the wrong involved in both cases ought to be similar in character.

[6] Notice also that understood in terms of concerns with the integrity of rational agency, prohibitions on acts of violence will not be restricted only to cases of physical injuries that tend to actually impair our capacity to use our powers purposively in choice and action (mutilation, serious head injuries, etc.). Given normal human psychology, being the victim of (or being threatened with) violence by another is likely to be accompanied by a range of more subtle psychological harms, causing us, for example, to shy away from contact with others or inducing in us a disproportionate or morbid fear of further injury or attack that will figure prominently in future deliberation and choice. Insofar as these kinds of effects are likely to affect our capacity for the kind of impartial deliberation required for inner freedom, actions of this sort will therefore also fail on universalization on this account.

REFERENCES

HERMAN, B. (1993), *The Practice of Moral Judgment*, Massachusetts, Harvard University Press.

KANT (1974), *Anthropology from a Pragmatic Point of View*, trans. Mary J. Gregor, The Hague, Martinus Nijhoff. Abbreviated as PA.

— (1991), *Groundwork of the Metaphysics of Morals*, trans. Mary J. Gregor, Cambridge, Cambridge University Press. Abbreviated as GW.

— (1991), *The Metaphysics of Morals*, trans. Mary J. Gregor, Cambridge, Cambridge University Press. Abbreviated as MM.

— (1996), "Religion Within the Boundaries of Mere Reason," *Religion and Rational Theology*, trans. Allen W. Wood and George Di Giovanni, Cambridge, Cambridge University Press. Abbreviated as RR.

KORSGAARD, C. (1996), *Creating The Kingdom of Ends*, Cambridge, Cambridge University Press.
TIMMONS (ed.) (2002), *Kant's Metaphysics of Morals: Interpretive Essays,* Oxford: Oxford University Press.