

# Kant's Self-Legislation Procedure Reconsidered<sup>1</sup>

Adrian M. S. Piper, APRA Foundation Berlin

Most published discussions in contemporary metaethics include some textual exegesis of the relevant contemporary authors, but little or none of the historical authors who provide the underpinnings of their general approach. The latter is usually relegated to the historical, or dismissed as expository. Sometimes this can be a useful division of labor. But it can also lead to grave confusion about the views under discussion, and even about whose views are, in fact, under discussion. Elijah Millgram's article, "Does the Categorical Imperative Give Rise to a Contradiction in the Will?"<sup>2</sup> is a case in point. In it, he takes the New Kantians<sup>3</sup> to task for various flaws in their interpretation of Kant's moral theory, to be detailed shortly. He concludes with a question and a suggestion. In order to properly dissect the first, "universal law" formulation of the Categorical Imperative, he argues, we first need to understand "why an agent wills the universalization of his max-

---

<sup>1</sup> This essay is excerpted from a longer discussion, *Kant's Metaethics: First Critique Foundations* (in progress). Earlier versions were delivered to the *Morality of Law: Kantian Perspectives* conference at the University of Keele (November 2010), and the Philosophy Department at King's College, University of London (February 2011). I am grateful to both audiences for questions and discussion. Katerina Deligiorgi provided very useful comments at the conference, as did Henry Allison in correspondence. I also learned much from discussing the issues with Garrath Williams and Sorin Baiasu. Two anonymous readers for *Kant Studies Online* have greatly improved the final version of this paper.

<sup>2</sup> *The Philosophical Review* 112, 4 (October 2003), 525-560. References to this article are paginated in the text.

<sup>3</sup> By which Millgram means most prominently John Rawls, Onora O'Neill, Christine Korsgaard, and Barbara Herman. See Millgram 552, fn. 1 for a sensitive taxonomy of this philosophical territory.

im” (549). He also suggests that in order to answer this question, we must recur to what Kant himself actually says (550). His question is a good one, and his advice on how to go about answering it is sound. But to take Millgram’s advice is to call this division of labor into question, at least for this case. For it demands close and sustained exegesis, not only of his argument against the New Kantians, but also – in order to assess whether and where they go wrong – of Kant’s text itself.

However, close attention to what Kant actually says yields thirty-eight “universal law” formulations of the categorical “imperative.” Sections II and III below call attention to thirty-one of these that can be divided into two subsets. The larger subset describes a form of rational deliberation that Section IV argues is an application of Kant’s Regulative Principle of Pure Reason from the Dialectic of the first *Critique* to the special case of action. These formulations depict the particular manner of willing through which a rational being specifies what action is morally permissible under given circumstances. Universal law formulations in the smaller subset, by contrast, exhort us to cultivate this manner of willing as an intentional goal of deliberation. Sections V and VI contend that the empirical method we are to follow in order to achieve this goal is an application of Kant’s Hypothetical Employment of Reason, also from the Dialectic of the first *Critique*. This second type of formulation provides an heuristic for improving our ability to deliberate about action as a rational being would. Each kind of formulation thus puts to work a different aspect of Kant’s extended first *Critique* analysis of reason in general, on a different aspect of reasoning about human action in particular.

It is not possible to answer Millgram’s question without examining these close connections between Kant’s comprehensive account of reason in the first *Critique* and his “universal law” formulations of the categorical “imperative” in the *Groundwork*. Unfortunately, we see in Section VII that these connections undermine rather than support Kant’s as-

piration to derive particular duties from universal moral law. They therefore subvert the New Kantian project of refining this aspiration into a viable decision procedure that might usefully claim the mantle of Kant's authority. To this extent, Millgram's objections to the New Kantians are well taken. However, Section VIII maintains that these first *Critique* connections also imply a weaker and therefore more powerful role for the two contradiction tests originally developed by Onora O'Neill. In this form, the tests successfully refute Hobbes' Foole – and therefore, in Section IX, reveal Millgram's attack on the New Kantians as ultimately self-defeating.

Nevertheless, this refutation of Millgram should not be thought to imply any claim to membership in the kingdom of New Kantian ends in themselves on whom Millgram focuses his objections. Nor does it undertake to defend the New Kantian interpretation of Kant. That interpretation, which Millgram credits Rawls with disseminating to his students (525), is one that Rawls himself was very resistant to publishing until late in his life and after considerable lobbying from some of them. In the mid-1970s, Rawls's lectures on Kant were just that: lectures in his course on moral psychology that explored Kant's texts in terms that illuminated, supported and amplified Rawls's own views. It would have been difficult to understand Rawls's approach to Kant, or even the terminology in which he discussed Kant's views, without knowing what his own views were. Those views were strongly influenced, in turn, by the Kantian model of moral deliberation presented in O'Neill's groundbreaking *Acting on Principle*.<sup>4</sup> Rawls also had great respect for the exegesis-based Kant scholarship from which he distanced his own writing,<sup>5</sup> and he urged his students to

---

<sup>4</sup> Onora Nell (née O'Neill), *Acting on Principle: An Essay on Kantian Ethics* (New York: Columbia University Press, 1975).

<sup>5</sup> Millgram describes this exegesis-based scholarship in Kant's moral philosophy as having occupied a "marginalized position ... a little over a quarter-century ago (525)." Perhaps the more accurate term would be "specialized." Since then, another useful division of labor has appeared that takes its cue from the first: Scholars of Kant's texts have continued to do the specialized exegetical work of paying close, analytical attention to what Kant

study the primary texts on which that scholarship was based.<sup>6</sup>

So the discussion that follows aims to take Rawls's counsel, too, to heart in addressing Millgram's objections to the New Kantians; and to limn an alternative, *Ur-Kantian* interpretation that avoids them. But because this alternative goes very much against the grain of received New Kantian interpretation of Kant's moral philosophy, much preliminary exegetical weeding and trimming, and a replanting of Kant's arguments in the fertile soil of the first *Critique* is necessary even in order to state it.<sup>7</sup> The end result retains the structural core of O'Neill's original analysis, but rejects both Millgram's arguments and also the New Kantian arguments that he criticizes. This end result does not appear until the very end, however, once all the New Kantian thicket has been cleared away.

---

himself says. At the same time, New Kantians have made Kant's basic approach and themes more accessible to a wider philosophical audience. Just as one need not be a Bentham or Sidgwick scholar in order to be a Utilitarian, one no longer need be a Kant scholar in order to be a Kantian. This has enabled the development of views that are Kantian in inspiration and sensibility without needing to strictly adhere to Kant's actual texts; I count my own work in metaethics among them. But as we see from close review of Millgram's critique, where the New Kantian approach purports to represent Kant's own views, it has also created certain problems.

<sup>6</sup> Indeed, it could be said that he occasionally did more than this: Rawls convinced some of the students who had come to him planning to write dissertations in contemporary moral philosophy to write in the history of ethics instead.

<sup>7</sup> The fundamentals of this alternative pre-date Millgram's article. See "Kant on the Objectivity of the Moral Law," in Andrews Reath, Christine M. Korsgaard and Barbara Herman, Eds. *Reclaiming the History of Ethics: Essays for John Rawls*, (New York: Cambridge University Press, 1997), 240-269; also at: [http://adrianpiper.com/docs/WebsiteKantOnMorObjRwlsFstschrft\(1994\).pdf](http://adrianpiper.com/docs/WebsiteKantOnMorObjRwlsFstschrft(1994).pdf).

## I. Millgram's Critique of the New Kantians

Millgram's objections are based on his rendering of what New Kantians call the *CI-procedure* (526-531). This consists essentially in O'Neill's contradiction in conception and contradiction in the will universalization tests of the moral validity of an action, plus more recent emendations to O'Neill's original analysis. Millgram's very generalized account of it is accurate in outline, and he is careful to footnote the details and qualifications by which various New Kantians have supplemented it. So newcomers may safely consult Millgram's gloss directly for a reliable generalized introduction to the CI-procedure (526-531). Suffice it to say here that for Millgram, the trouble begins when one attempts to apply what he calls the *test of reflexive application* – the application of one's theory to itself – to the CI-procedure. Millgram observes that “governing one's activity by the CI-procedure . . . is, anyway when one is ‘acting from the moral law,’ itself an action or plan of action;” and this, he suggests, enjoins on one what he calls the *CI-maxim*:

- (1) When I am making up my mind what to do, I will act only on maxims that pass the CI-procedure, so as to make (morally or rationally) permissible decisions (531).

However, he goes on to argue, it is not possible to guide all of one's actions according to the CI-maxim, because the CI-maxim requires that one understand one's actions as governed by lawlike, which is to say exceptionless policies; whereas successful agency requires exceptions from such policies as a matter of principle (532). Millgram concludes that

- (2) [. . .] the CI-maxim gives rise to a contradiction in the will: one's stake in one's own agency is such that one cannot endorse

having it undermined by being deprived of the exceptions that are its precondition . . .  
 [T]hen it is forbidden to act on the CI-maxim,  
 and Kantian moral theory is (at least in its  
 New Kantian rendition, and stating the  
 conclusion informally) self-refuting (533).

Millgram supports this conclusion with decision-theoretic considerations: As agents, we need the cooperation of other agents in order to achieve our ends. We must try to achieve these largely in ignorance of the personal rules, policies and ends governing other agents' actions, as well as of future changes to our own ends. These shared conditions of risk and uncertainty, both about others' ends and about our own future ends, hold for all agents; and thus reduce the ease and likelihood of spontaneous interpersonal coordination. They therefore increase the need for agents to cooperate in achieving one another's ends, at least to the extent of permitting exceptions to those rules when the costs to oneself are minimal (534-536). In its presupposition that interpersonally coordinated lawlike practices can be the outcome but not the precondition of self-interested rational action, Millgram's critique implicitly relies on the type of game-theoretic account of the origins of interpersonal coordination defended by such authors as David Lewis,<sup>8</sup> Allan Gibbard<sup>9</sup> and Edna Ullman-Margalit.<sup>10</sup>

New Kantians concede the necessity of at least seeming to make exceptions to lawlike rules and policies. However, Millgram argues that the New Kantian CI-procedure forecloses all of the modifications New Kantians have proposed in order to accommodate them. Modifying the rule by build-

---

<sup>8</sup> David Lewis, *Convention: A Philosophical Study* (Cambridge, Mass.: Harvard University Press, 1969).

<sup>9</sup> Allan Gibbard, "Utilitarianisms and Coordinations" (Ph.D. diss., Harvard University, 1971)

<sup>10</sup> Edna Ullman-Margalit, *The Emergence of Norms* (Oxford: Clarendon Press, 1977).

ing the exception into it will not do, because that modified rule itself must now satisfy the CI-procedure – often with the unhappy result that universalizing the exception sabotages the point of the original rule: Allowing a promising student to skip class and graduate anyway, for example, would, when made a universal rule, undermine academic performance among all students. “[B]ecause rational agents adjust their plans to take account of changes in the rules ..., institutionalizing exceptions ... amounts to what is, by Kantian lights, a self-frustrating plan of action (541).” Indeed, a *Revision Maxim* stipulating that

- (3) When I run into a case that my (first-order) maxim does not handle properly, I will revise my (first-order) maxim to incorporate the exception, so as to be able to act on rules that I understand to be “provisionally universal” ... (542)

fails the CI-procedure; for such rules will be too complicated and qualified by exceptions, first, to ascertain whether or not they are universalizable; and second, to know what actions they require (542-3). Nor will building in generic exception clauses rescue these modified rules, because our lack of information about their specific nature, number and frequency makes it impossible to formulate the universalized counterfactual conception the internal consistency of which the contradiction in conception test is supposed to evaluate (544). Nor can the New Kantians deny the necessity of observing the CI-maxim, because this would be to violate one of the central tenets of Kant’s moral theory, that one act not only in accord with the moral law, but also out of respect for it (546). Nor can such exceptions be excluded from the class of *bona fide* maxims on account of insufficient generality, without depriving the CI-procedure of the power to guide specific actions. Nor can what counts as a maxim be left to the judgment of the agent, for this gives the CI-procedure the structure of an imperfect duty, in which “agents [have]

discretion to decide when to invoke the CI-procedure as a constraint on their actions” (547). Millgram rather wickedly concludes that

- (4) [...] if contradiction-in-the-will arguments establish imperfect duties, it looks as though we have exhibited a Kantian imperfect duty to *violate* the Categorical Imperative, and so that part of the position will have to go (549).

Millgram’s critique of the New Kantian CI-procedure thus has the form of a *reductio ad absurdum*. Although it would be possible to find fault with some of its details, this would be mere epicycle-spinning that missed Millgram’s main point: Something has gone seriously wrong with the CI-procedure, and it might be useful to recur to what Kant actually says, in order to ascertain both what it is and also whether Kant’s own analysis is subject to the same attack.

## II. Millgram's Question

To address the question of why an agent wills the universalization of her maxim, it is first necessary to look more closely at Millgram’s advice as to how to go about answering it. To what extent, and how closely, is any writer obligated to consult or adhere to Kant’s text, when purporting to represent Kant’s own view, rather than merely a Kantian view? Of course there can be no determinate or conclusive answer to this question; and there is now a very respectable tradition of representing what purport to be Kant’s own views on moral philosophy in terms that contain a notable paucity of citations to or direct exegesis of Kant’s text. Nevertheless, just as Millgram implicitly poses this question to the New Kantians, whom he chastises for not attending closely enough to Kant’s text, one can also pose it to Millgram himself. For his critique of the New Kantian approach is based not on his own, independent



reading of Kant's text, but rather on what he takes the New Kantians to say about it (525), which in his view is insufficient. Because Millgram's own comments are thus also several removes from what Kant actually says, a certain amount of underbrush must be cut away just in order to determine whether and to what extent those comments hit the mark.

For example, Millgram counsels us to ask why an agent wills the universalization of his maxim. His question is based on what both he (526) and the New Kantians describe as Kant's "first" formulation of the categorical "imperative":

- (5) There is only one categorical imperative, and it is this: (a) act only in accordance with that maxim (b) through which you can at the same time will (c) that it become a universal law (GMS Ak. 04:421.08-09).<sup>11</sup>

But first, passage (5) is not Kant's first formulation of the categorical imperative, but rather the fourth.<sup>12</sup> Second, (5) is not cast in the imperative mood, but rather has the form of a

---

<sup>11</sup> Immanuel Kant, *Grundlegung zur Metaphysik der Sitten*, Herausg. Karl Vorländer (Hamburg: Felix Meiner Verlag, 1965), 42. Henceforth references to this work are paginated in the text according to the German Academy edition. The German original runs as follows:

Der kategorische Imperativ ist also ein einziger und zwar dieser: (a) handle nur nach derjenigen Maxime, (b) durch die du zugleich wollen kannst, (c) daß sie ein allgemeines Gesetz werde. (GMS Ak. 04:421.08-09)

All translations into English are mine unless otherwise noted.

<sup>12</sup> See GMS Ak. 04:402.09-11 for the first, his question to the reader at GMS Ak. 04:403.26-27 for the second and GMS 421.03-06 for the third. All forty-seven formulations of the so-called categorical "imperative" in the text of the *Groundwork* are listed in Appendix Table 1; I claimed thirty-two in "Kant on the Objectivity of the Moral Law," *op. cit.* Note 7, but that was wrong. Appendix Table 1 numbers each formulation in order of its sequential appearance in the text of the *Groundwork*. But in the event that even this attempt at comprehensiveness overlooks some, I simply cite the pagination for each as needed for purposes of this discussion. Either alternative seems to me more useful than the now-standard New Kantian convention of referring to the First, Second or Third; or to *the* Formula of Universal Law, *the* Formula of Autonomy, and so on.

command.<sup>13</sup> Third, (5) enjoins, not that one will the universalization of one's maxim, but rather that one will its universalizability, i.e. that one *be able to* will its universalization. Fourth, (5) does not actually enjoin one to will the universalizability of one's maxim, but rather to will *that* one's maxim be universalizable. That is, passage (5) makes clause (c) the forward-looking intentional object of a sentential rather than a subsentential or adverbial attitude expressed in clause (b). It enjoins the will to adopt as a goal that one's maxim become a universal law.

Passage (5) is not the only formulation of the categorical "imperative" that enjoins us to adopt the universalizability of our maxim as a forward-looking intentional object of the will. All of these passages are listed in Appendix Table 4, of which GMS Ak. 04:447.05-08 is discussed below in Section VII as passage (26). But all of these *intentional-object* formulations raise the same question. For on the face of it, this feature of Passage (5) sets it in direct conflict with the passage that immediately precedes it:

- (6) But when I think a categorical imperative, I know immediately what it contains. For because the imperative, aside from the law, contains only the necessity that the maxim conform to the law, and the law contains no condition that limits it, nothing remains to which the maxim of the action shall conform except the universality of a law in general; and which conformity alone the imperative actually represents as necessary (GMS Ak. 04:420.31-33, 421.01-06).<sup>14</sup>

---

<sup>13</sup> I offer an analysis of the distinction between imperatives and commands in *Rationality and the Structure of the Self, Volume II: A Kantian Conception* (Berlin:

<http://adrianpiper.com/rss/docs/Rationality%20and%20Structure%20of%20the%20Self,%20Volume%20II-%20A%20Kantian%20Conception.pdf>

APRA Foundation Berlin, 2008), Chapter IX. "'Ought'."

<sup>14</sup> Denke ich mir aber einen kategorischen Imperativ, so weiß ich sofort, was er enthalte. Denn da der Imperativ außer dem Gesetze nur die Notwendigkeit der Maxime enthält, diesem Gesetze gemäß zu sein, das Gesetz aber keine

Kant's claim here is that the categorical imperative contains *only* the necessity that the maxim of one's action conform *only* to the universality of a law in general. It repeats his earlier argument in *Groundwork* Chapter I, that the categorical imperative is what remains after we remove from the will *every* incentive – i.e. every forward-looking goal or end – that could arise from following *any* law, namely its conformity to law in general; and that this must be its *only* principle (GMS Ak. 04:402.05-09).<sup>15</sup> In both places, Kant's argument is that the categorical imperative abstracts from all conditions and incentives of the will, leaving only what he describes as the “bare conformity to law in general (GMS Ak. 04:402.11-12)”<sup>16</sup> of its maxims. His official definition of the categorical imperative as

- (7) [...] concerned not with the matter of the action and what is supposed to result from it, but rather with the form and principle from which it itself follows (GMS Ak. 04:416.12-15)<sup>17</sup>

reinforces this emphasis on the bare formal structure of one's maxims, and unconditional disregard of all of its incentives.

But if the moral law consists merely in the conformity to law of the will's maxims, without regard to *any* of its

---

Bedingung enthält, auf die es eingeschränkt war, so bleibt nichts als die Allgemeinheit eines Gesetzes überhaupt übrig, welchem die Maxime der Handlung gemäß sein soll, und welche Gemäßheit allein der Imperativ eigentlich als notwendig vorstellt (GMS Ak. 04:420.31-33, 421.01-06).

<sup>15</sup> Da ich den Willen aller Antriebe beraubt habe, die ihm aus der Befolgung irgend eines Gesetzes entspringen könnten, so bleibt nichts als die allgemeine Gesetzmäßigkeit der Handlungen überhaupt übrig, welche allein dem Willen zum Prinzip dienen soll, d.h. [...] (GMS Ak. 04:402.05-09).

<sup>16</sup> [...] die bloße Gesetzmäßigkeit überhaupt [...] (GMS Ak. 04:402.9-11).

<sup>17</sup> Er betrifft nicht die Materie der Handlung und das, was aus ihr erfolgen soll, sondern die Form und das Prinzip, woraus sie selbst folgt, [...] (GMS Ak. 04:416.12-15).

conditions, results or incentives, the imperative formulation of that law should not then ascribe to the will the incentive of universalizing its maxims. By making (c) the forward-looking, goal-directed intentional object of (b) in the formulation at passage (5) above (and others like it), Kant does exactly this: He exhorts me to act such that I can will that my maxim become universal; and thereby ascribes to my will as its incentive my maxim's becoming universal – right after insisting that the categorical imperative “contains *only* the necessity that the maxim conform to the law [italics added];” and that the categorical imperative “contains ... *nothing* ... to which the maxim of the action shall conform except the universality of a law in general [italics added].” None of these passages that describe what the categorical imperative is or how it is to be recognized suggest that it must also include itself as an incentive of the will that conforms to it. Indeed, in rejecting *any and all* incentives as irrelevant to the identification of the categorical imperative, these passages explicitly rule this out.

So in passages (5) and (6), taking Millgram's advice to attend to what Kant actually says, rather than to the New Kantian account that abstracts from it, has two consequences. First, it raises the possibility that Millgram's question as he formulates it is a *non sequitur* that does not interrogate any actual claim Kant makes in the text. Second, it exposes an apparent inconsistency in Kant's view that could have been ignored by declining to scrutinize his actual words so closely. Why does Kant first insist that conformity of a maxim to the moral law excludes all incentives, and then ascribe to the will conformity of one's maxim to the moral law as its incentive? Section VII below proposes an explanation of this seeming inconsistency, and, in so doing, calls attention to another way in which both Millgram and some of the New Kantians have misread Kant's self-legislation procedure. But in order to set the stage for that explanation, a bit more pruning and ground-clearing is necessary. Above all, Millgram's question must be reframed

in terms that – as he recommends – heed what Kant actually says.

### III. Millgram's Question Reframed

So, then: Are there any passages in which Kant does make something like the claim Millgram ascribes to him, that an agent wills the universalization of her maxim? There are several. In all of them, Kant prescribes not that I am to *will that* my maxim could become a universal law; but rather that I am to *will a maxim as* universal law, i.e. that I am to will a universal or universalizable maxim. To consider a paradigmatic one, Kant characterizes the principle of autonomous self-legislation as that “quality of the will through which it is a law unto itself (independently of any quality of the objects of willing).” He continues,

- (8) (a) The principle of autonomy is thus: not to choose otherwise than so that (b) the maxims of its choice are at the same time (c) comprehended as universal law (d) in the same willing (GMS Ak. 04: 440.20-23).<sup>18</sup>

Clauses (b) and (d) make the important point that a self-legislating will does not perform two acts of willing – the first described by the maxim of our action, the second described by the universalization of that maxim; but rather just one: we are to will the maxim *as itself* universal law. The form of the maxim itself must be universal. So passage (8) tells us not *what* intentional object we are to will, but

---

<sup>18</sup> Autonomie des Willens ist die Beschaffenheit des Willens, dadurch derselbe ihm selbst (unabhängig von aller Beschaffenheit der Gegenstände des Wollens) ein Gesetz ist. Das Prinzip der Autonomie ist also: nicht anders zu wählen, als so, daß die Maximen seiner Wahl in demselben Willen zugleich als allgemeines Gesetz mit begriffen sein (GMS Ak. 04:440.18-23).

To identify an autonomy formulation as itself a universal law formulation would be heresy in some circles, but not in this one.

rather *how* we are to will. We are to will in such a way as to understand our maxims as universal law. Here Kant merely elaborates on a point he has already made, that in rational willing

- (9) [. . .] the concept of the action in itself already contains a law for me (GMS Ak. 04:402.40-41).<sup>19</sup>

Passage (8) is only one of several adverbial formulations of the categorical “imperative” that enjoin us to act on that maxim through which we can will it *as* a universal law; all twenty-four are collected in Appendix Table 5. Most occur later in the text than the intentional-object formulations listed in Appendix Table 4. A maxim *through which we can will our action* in any way at all is a motivationally effective maxim. A maxim through which we can will our action *as* something is one that represents the action we are motivated to perform in a certain way, or under a certain interpretation or description. A maxim through which we can will our action *as a universal law* is one that represents the action we are motivated to perform as itself being, containing, enacting, or expressing a universal law. We need not choose between these possibilities right now, because Kant’s other formulations of the moral law develop each of them in depth. For present purposes, the important point is Kant’s insistence that rational maxims represent their actions as having this intrinsic structural connection to universal law, i.e. as being lawlike; call this the *representation thesis*. Kant’s representation thesis as stated does not specify exactly what kind of connection that must be. Nor does it imply that it is precisely this representation of the action that motivates us to perform it; that argument must be made

---

<sup>19</sup> Nun ist es doch etwas ganz anderes, aus Pflicht wahrhaft zu sein, als aus Besorgnis der nachteiligen Folgen: indem im ersten Falle der Begriff der Handlung an sich selbst schon ein Gesetz für mich enthält [...] (GMS Ak. 04:402.38-41).

separately.<sup>20</sup> Nor does it follow from the representation thesis, however, that the maxim itself must include our thus representing it in the formulation of its goal or purpose. That argument, too, would have to be made separately, and Section VII below takes it up in greater depth.

What the representation thesis does imply is that a rational maxim requires, not any specific forward-looking intentional object of the will, but rather a specific *manner* of willing, whatever that object might be; i.e. *in such a way* that it is lawlike. In this case, Millgram's question should be, not "why an agent wills the universalization of his maxim (549)," but rather why an agent wills *a universalizable maxim*; that is, why what an agent wills has the form of a universalizable maxim. This reading does not require a universalized maxim as the goal-directed intentional object of the agent's will. It requires only that his will have a certain form and potential, irrespective of that intentional object.

So we have sorted the "universal law" subset of Kant's formulations of the categorical "imperative" into two groups: The first, intentional-object group comprising seven formulations (Appendix Table 4) treats the universalizability of one's maxim as a forward-looking intentional object of the will. These are the ones that deploy the "will that" locution. By contrast, the second group comprising twenty-four formulations (Appendix Table 5) treats this idea adverbially, as a *manner of willing* that is lawlike; the formulations in this manner-of-willing group often deploy the "will as" locution. Why Kant alternates between these two constructions requires explanation, and Section VII will propose one. But on the face of it, only the manner-of-willing formulations are consistent with those passages –

---

<sup>20</sup> I offer such an argument in "Practical Action: First *Critique* Foundations," in *Kant und die Philosophie in Weltbürgerlicher Absicht: Akten des XI. Internationalen Kant-Kongresses*, Herausg. Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca, und Margit Ruffing (Berlin: Walter de Gruyter, forthcoming).

(6), (7), and (9) above – that identify those features of the categorical imperative that differentiate it from other principles of action, namely its lawlike form, independent of any and every incentive they might include.

But in thus obeying Kant's directive to abstract from the will *every* incentive it might have to follow *any* law, we also abstract from the will the incentive to pass the CI-procedure, and thus to achieve moral validity for our actions. Kant's directive seems to imply a disavowal of the very foundation of the New Kantian project. So we arrive at another juncture at which we must choose between following Millgram's advice to attend to what Kant actually says, and accepting what the New Kantians – and Millgram – have constructed from it. Suppose we now choose again to follow Millgram's advice; i.e. to accept the consequences of what Kant actually says in the passages discussed above, despite the seeming inconsistency unearthed there. Then neither the New Kantian CI-procedure, nor Millgram's CI-maxim, nor his Revision Maxim offer a satisfactory account of most of Kant's "universal law" formulations of the categorical imperative. Millgram and the New Kantians urge us to will a particular state of affairs that ensures the moral validity of our maxims. By contrast, in passage (8) and the twenty-three other manner-of-willing formulations like it, Kant is urging us, rather, to will *in a manner* that ensures the moral validity of our maxims. Therefore, it would seem, the New Kantian CI-procedure, Millgram's CI-maxim, and his Revision Maxim should be abandoned.

Now Millgram maintains that his CI-maxim cannot be abandoned, without violating Kant's stricture that one act not only in accordance with the moral law, but also out of respect for it (546). However, passage (8) is not vulnerable to this argument. The German term usually translated as "respect" is *Achtung*. This word is ordinarily used in situations in which our respectful *attention* is being commanded; when we are being asked to heed some background condition or state of affairs in formulating and carrying out our actions (for example, as the sign *Achtung Stufe!* instructs us



to mind the step as we walk through the door). It thus expresses a *backward-looking* rather than a forward-looking intentional attitude, and takes as an object a backward-looking motive cause of action, or (to use Kant's term) a *ground*, rather than its forward-looking goal, incentive, or purpose.<sup>21</sup> *Achtung* causes me to heed the universality requirement of the moral law by conforming the maxims of my actions to it; that is, by willing universalizable maxims such that "the concept of the action in itself already contains a law for me."

This manner of willing expresses respect for the moral law by enacting a certain quality of respectful attention to the way in which I formulate my maxim. First, I formulate it in a way that is mindful of its extended temporal applicability beyond the particular action that I am contemplating at this moment. Second, my formulation is mindful of its extended spatial applicability to other possible agents contemplating the same action at this or other times. Third, my formulation is mindful of the feasibility of such application to other agents at this or other times, whether actual or possible. It is sensitive to the nomological scope of my maxim's spatiotemporal applicability. These three conditions are individually necessary and jointly sufficient for satisfying the representation thesis. Conjointly they explain how a maxim of action expresses respect for the moral law.

None of these indices of respectful attention can be plausibly incorporated into the forward-looking intentional object of the maxim, not even as "side-constraints" we must strive to observe. Picture the classroom, packed with smart and impertinent students ready to pounce on one's every word, not only in order to absorb and quote them back on the next exam, but also to disruptively dispute them if clumsily expressed. Consider the impact of one's awareness of these students on the care with which one formulates a

---

<sup>21</sup> I defend the distinction between backward-looking motives and forward-looking purposes in "Moral Theory and Moral Alienation," *The Journal of Philosophy* LXXXIV, 2 (February 1987), 102-118.

difficult point in lecture. Certainly one strives to get the point right. But is it therefore plausible to describe the rapt and vigilant students as “side-constraints” on one’s formulation of this point that one also strives to observe? Of course not. That one’s students are intentional objects of one’s awareness does not make them forward-looking goals of one’s intentions. There are other kinds of intentional objects besides intentions. And these may coexist alongside or in the background of intentions themselves.

Now substitute for that packed classroom the extended possible world packed with past, present and future counterfactual agents to whom one’s maxim rationally must apply. One’s mindfulness of their presence no more involves a forward-looking intentional object in this latter case than in the former. Rather, one’s attention and care to the lawlike form of one’s maxim, independent of its content or goal, expresses that backward-looking intentional attitude whose object is an invariant ground of one’s willing; namely, “concer[n] [...] with the form and principle from which it itself follows.” A backward-looking intentional attitude does not require, in addition, that one also will conformity to its ground or motive as a forward-looking purpose. So willing a universalizable maxim expresses respect for the moral law without requiring Millgram’s trouble-making CI-maxim.

Abandoning Millgram’s CI-maxim does not imply that we do not care about the moral validity of our actions. But it does imply that the moral validity of our actions does not require our striving to pass a test of moral validity. Rather, it requires our strivings to be in principle lawlike; that is, to satisfy the representation thesis. Kant’s claim is that lawlike willings confer moral validity on the actions they effect. So it appears that Millgram’s sound advice, to attend to Kant’s text, has led us to a rather different question than the one he himself originally formulated: Why, according to Kant, is a universalizable maxim the form and principle of an agent’s will?

#### IV. Maxims and the Regulative Principle of Pure Reason

The texts of the first *Critique* and *Groundwork* answer that the agent has no choice; that is the way we are cognitively constructed.<sup>22</sup> An agent must will a universalizable maxim because of the way maxims fit into Kant's broader account of reason. Kant's model of reason is always one and the same, regardless of context or application, namely that which he develops in detail in the Dialectic of the first *Critique*; and he emphasizes this repeatedly (cf. KrV, A 815/B 843.28-29;<sup>23</sup> GMS Ak. 04:391.29-34; KpV, 05:89.15-16 and 05:121.02-06). Reason consists in the familiar law-governed operations of logical analysis, generalization, deductive and inductive inference, hypothesis formation and application. These enable us to organize and unify the data of experience under higher-order concepts, principles, ideas, and theories through judgment.

Judgment, for Kant, is the cognitive function by which the laws of reason are applied to particular cases, and to the special case of action specifically, in order to achieve this unity. Judgment collects relevantly similar representations already subsumed under a less abstract concept in turn under a more abstract one (KrV, A 79.09-19/B 104.26-B 105.12), by ascribing a predicate to a subject in a categorical indicative judgment  $Fx$ . This is also the form of judgment through which we ascribe various predicates to ourselves as subjects. In Kant's Table of Judgments, the categorical indicative form of judgment is the first under the rubric of Relation, and enters into both hypothetical and disjunctive judgments

<sup>22</sup> This and the following section recapitulate and elaborate some material from "Kant on the Objectivity of the Moral Law," *op. cit.* Note 7. More recently, see also "Practical Action: First *Critique* Foundations," *op. cit.* Note 20.

<sup>23</sup> Immanuel Kant, *Kritik der reinen Vernunft*, Herausg. Raymund Schmidt (Hamburg: Felix Meiner Verlag, 1976). Henceforth references to this work are paginated in the text according to the standard A/B pagination. All translations from the German are mine.

as subsentential or sentential components (KrV, A 70/B 95.13-27). The logical function of all three forms of judgment is to collect representations under more general and abstract ones at every level, and finally all of them under the concept "I think" in a single unified consciousness (KrV, B 143.08-16). Kant directs us to sort and organize the variegated, manifold contents of our minds – "sensation, consciousness, memory, wit, power of discrimination, inclination [*Lust*], desire, etc." – in such a way that we can apply the "logical maxim through which we reduce, so far as possible, this seeming diversity, so that through comparison one might discover hidden identity (KrV, A 649/B 677.01-09)."<sup>24</sup> In so doing, judgment functions to increase theoretical coherence, therefore cognitive unity, and thereby the unity of the self under this maxim.

Kant also tells us in the first *Critique* what a maxim is:

- (10) [I]t is simply a subjective law for inventorying the supplies of our understanding, through comparison of its concepts, in order to bring their general use to the smallest possible number (KrV, A 306.04-08/B 362.21-25).<sup>25</sup>

---

<sup>24</sup> [...] wie in dem menschlichen Gemüte die Empfindung, Bewußtsein, Einbildung, Erinnerung, Witz, Unterscheidungskraft, Lust, Begierde usw. Anfänglich gebietet eine logische Maxime, diese anscheinende Verschiedenheit soviel als möglich dadurch zu verringern, daß man durch Vergleichung die versteckte Identität entdecke, [...] (KrV, A 649/B 677.01-09).

<sup>25</sup> [...] sondern ist bloß ein subjektives Gesetz der Haushaltung mit dem Vorrat unseres Verstandes, durch Vergleichung seiner Begriffe, den allgemeinen Gebrauch derselben zu bringen [...] (KrV, A 306.04-08/B 362.21-25).

Patricia Kitcher's "What is a Maxim?" *Philosophical Topics* 31, 1 & 2 (April 2005) provides a careful, scholarly and very thorough survey of the literature on this topic that is confined to Kant's claims about maxims of action, and does not address the first *Critique* background of those claims. I think this may partly explain why, on her analysis, only maxims of self-love, and not those with moral content, are to be tested by the CI-procedure. I don't see how this can be right, because it is the test itself that determines the moral status of a maxim's content – required, permitted, or prohibited – in the first place. But then, as is clear from the discussion, my animadversions against the CI-procedure go even deeper than this.

A *maxim* is a principle of judgment that guides the subject in collecting, sorting and organizing the experiential and conceptual contents of lower-order judgments. The concept of a maxim, for Kant, derives from that of a *maximum* as the rational idea of what is "greatest and absolutely complete ... in the division and unification of the knowledge of the understanding under one principle" (KrV, A 665/B 693.09-11)<sup>26</sup>. The rational idea of a maximum is, in effect, the idea of a Theory of Everything. On Kant's view, it is complete and comprehensive knowledge itself that reason is driven to maximize, rather than any particular desire-satisfaction that might be gained by putting it to use. A *maxim* approximates the *maximum* to the extent to which the maxim reduces the number, while increasing the scope, of the organizing principles by which knowledge is managed.

In addition, Kant tells us what a maxim *of reason* is:

- (11) I call all subjective principles that are derived not from the constitution of the object, but rather from reason's interest in considering a certain possible perfection of the knowledge of this object, *maxims of reason* (KrV, A 666/B 694.05-09).<sup>27</sup>

Whereas a maxim in general is a kind of judgment that simply collects the contents of experience under the smallest possible number of concepts that organize them, a maxim of reason is a judgment that does this with the specific purpose of completing our knowledge of the object thus organized – that is, with the rational idea in mind of a maximum as the valued end-state. A maxim of reason is a certain kind of maxim: it does not merely simplify and streamline our con-

<sup>26</sup> [...] das Größte und absolut Vollständige [...] der Vereinigung der Verstandeserkenntnis in einem Prinzip [...] (KrV, A 665/B 693.09-11).

<sup>27</sup> Ich nenne all subjektiven Grundsätze, die nicht von der Beschaffenheit des Objekts, sondern dem Interesse der Vernunft, in Ansehung einer gewissen möglichen Vollkommenheit der Erkenntnis dieses Objekts, hergenommen sind, Maximen der Vernunft (KrV, A 666/B 694.05-09).

ceptual scheme, but also, thereby, aims to maximally enhance its scope and depth. Whereas a maxim in general could be little more than a shallow and dismissive generalization about experience, such as “It’s a jungle out there,” or “Got to look out for Number One,” and the like, a maxim of reason strives for the status of explanatory first principle in a sophisticated, fully unified theory of experience. Whereas a maxim in general need not be guided by the rational idea of a maximum, a maxim of reason must be.

Kant’s most important maxim of reason is his *Regulative Principle of Pure Reason*:

- (12) . . . to find, for the conditioned knowledge of the understanding, the unconditioned whereby its unity is completed.

But this logical maxim can only be a principle of *pure reason* through our assuming that if the conditioned is given, so is the entire series of conditions, subordinated to one another and itself unconditioned, also given (KrV, A 307.22-28, A 308.01/B 364.14-21).<sup>28</sup>

The Regulative Principle of Pure Reason is a maxim of reason that guides our search, given any particular, for that most abstract, universal and all-inclusive first principle, which we also take to be given, from which the particular can be deduced (KrV, A 652/B 680.04-12).<sup>29</sup> An action, for

---

<sup>28</sup> [...] der eigentümliche Grundsatz der Vernunft überhaupt (im logischen Gebrauche) sei: zu dem bedingten Erkenntnis des Verstandes das Unbedingte zu finden, womit die Einheit desselben vollendet wird.

Diese logische Maxim kann aber nicht anders ein Prinzipium der *reinen Vernunft* werden, als dadurch, daß man annimmt: wenn das Bedingte gegeben ist, so sei auch die ganze Reihe einander untergeordneter Bedingungen, die mithin selbst unbedingte ist, gegeben [...] (KrV, A 307.22-28, A 308.01/B 364.14-21).

<sup>29</sup> [T]hat a certain systematic unity of all possible empirical concepts so far as they can be derived from higher and universal ones must be sought is elementary, a logical principle, without which no employment of reason would occur, since we can conclude from the universal to the particular only so far as universal properties of the thing are presupposed, under which the particulars stand (KrV, A 652/B 680.04-12).

example, is an empirical concrete particular event; and the maxim of that action aspires to the status of a maxim of reason that attempts to unify all of the data of our experience, all of our conscious knowledge and information at a certain time, under a highest-order principle that maximally organizes it according to the lawlike concept of the action we resolve to take at that time. In formulating my maxim, I “pull it all together,” and I act!

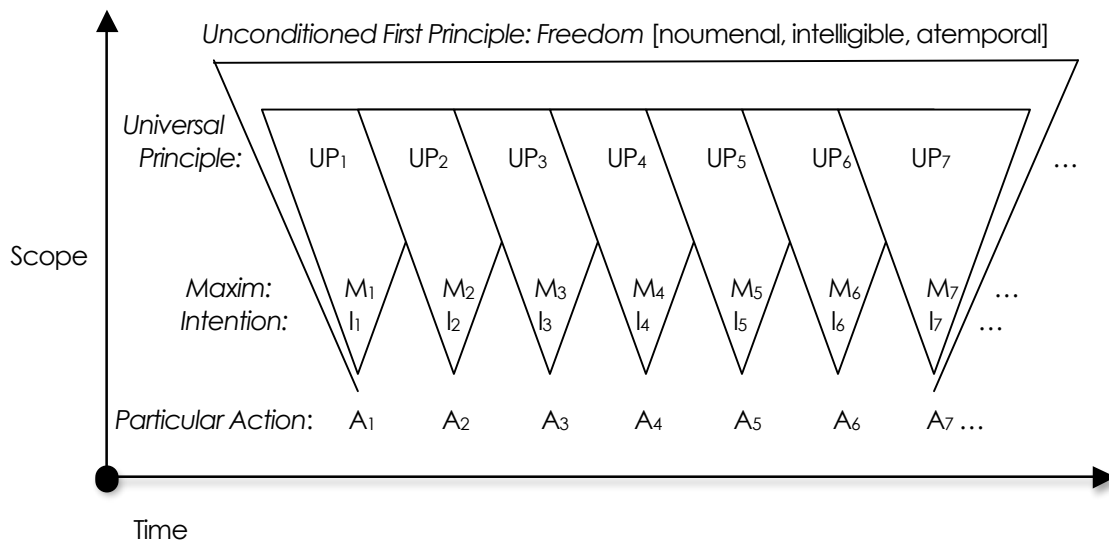
We can see Kant’s Regulative Principle of Pure Reason at work as a guide to inquiry when we perform an action whose maxim fails to achieve the status of a maxim of reason, and instead strikes us as opaque, arbitrary, or self-defeating: Our instinctive interrogation of our own motives for having performed it seeks an explanatory first principle that will render it maximally intelligible to ourselves. This search for transparency does not always succeed, for not all maxims of action are, in fact, maxims of reason. But Kant assumes that if an event really is an action rather than mere behavior, its ultimate explanatory first principle is always the same, namely one’s freedom to determine one’s own actions in light of reason’s demands at that time and place – whether this is clear to us or not. These particularizing demands specify the range of intermediate universal principles under which the action can, in fact, be subsumed. By contrast to a mere maxim, a maxim of reason consciously and deliberately determines our behavior in direct and transparent response to those demands. Thus Kant’s point is that *rational* action already *is* maximally intelligible to us, *in advance of performing it*. It is rendered transparent by the maxim of reason that governs it and the higher-order universal principles that subsume it. And it is motivated by

---

In the German original: [...] daß also eine gewisse systematische Einheit aller möglichen empirischen Begriffe, sofern sie von höheren und allgemeineren abgeleitet werden können, gesucht werden müsse; ist ein Schulregel oder logisches Prinzip, ohne welches kein Gebrauch der Vernunft stattfinden, wie wir nur sofern vom Allgemeinen aufs Besondere schließen können, als allgemeine Eigenschaften der Dinge zu Grunde gelegt werden, unter denen die besonderen stehen (KrV, A 652/B 680.04-12).

respectful attention to its link as a particular with that universal first principle that is its ultimate ground, irrespective of its end, goal or purpose. In this case, the Regulative Principle of Pure Reason is functioning, not as a guide to inquiry, but rather as a guide to action.

Thus Kant's manner-of-willing formulations of the categorical "imperative" in the *Groundwork* are merely an application of the first *Critique's* Regulative Principle of Pure Reason to the special case of action: We are to will only that action for which we can find the universal first principle from which its maxim can be deduced. That is, we are to will a universalizable maxim:



In this diagram, the vertex of each triangle identifies a particular action in a chronological series of such actions. Each vertex fans out in scope to encompass the intention under which the particular action can be subsumed, the maxim under which the intention can be subsumed, and finally to the most universal principle under which the maxim can be subsumed at its inverted base. But each such universalizable principle itself can be subsumed under the unconditioned first principle of freedom. This encompasses an agent's entire chronological series of actions and so has the broadest scope and widest base of all. An action from



whose universal first principle its maxim can be deduced is one that ultimately instantiates the principle of freedom.

For any particular action within this series, its universalizable maxim is a first-person categorical indicative judgment, ascribing the action as a causal property to oneself as agent, of the form, “I will do A.” This judgment collects, sorts, orders and ranks all of the cognitive and conative data that constitute one’s consciousness of oneself, one’s motives, one’s purposes and one’s circumstances at a particular time and place. As a maxim of reason, it thereby unifies the contents of self-knowledge, enabling one to collect *oneself*, in preparation for action. To describe it as universalizable is to ascribe to it the implicit force of law – not as a prescription but rather as a principle of organization by which these contents are recognized. When Kant later in the first *Critique* describes maxims as “practical laws, in so far as they become at the same time subjective grounds of actions, i.e. subjective principles (KrV A 812/B 840.18-20),”<sup>30</sup> he is calling attention to the conative power of such a judgment to spur action by organizing the various contents of one’s present experience under it. Through the regulative act of judgment itself, then, reason becomes motivationally effective:

- (13) [Reason’s laws] admittedly still require a power of judgment sharpened by experience, in part to distinguish to which cases they apply, in part to procure for them entry to the human will and impact [*Nachdruck*] on practice (GMS Ak. 04:389.35-39; cf. first *Critique*, KrV, B 150.16-19, B 153.19-24, B 154.01-02).<sup>31</sup>

---

<sup>30</sup> Praktische Gesetze, sofern sie zugleich subjektive Gründe der Handlungen, d.i. subjective Grundsätze werden, heißen *Maximen* (KrV A 812/B 840.18-20).

<sup>31</sup> [...] Gesetze *a priori*, die freilich noch durch Erfahrung geschärfte Urteilstkraft erfordern, um teils zu unterscheiden, in welchen Fällen sie ihre Anwendung haben, teils ihnen Eingang in den Willen des Menschen und Nachdruck zur Ausübung zu verschaffen, [...] (GMS Ak. 04:389.35-39)

These laws require application in the universalizable maxims that instantiate them. The judgment that applies to my own cognitive and conative state a lawlike descriptive maxim that subsumes and organizes its manifold mental contents thereby causally effects the action it ascribes to me.

So in order for a maxim of action to be rational – i.e. to enter into any process of rational deliberation, whether actual, subliminal, or rationally reconstructed, it must be lawlike in exactly the same two ways that any such cognitive judgment is lawlike. First, it must subsume lower-order particulars under it as instances – specifically, in this case, the particulars of the action I intend to perform. Second, it itself must be subsumable under such a higher-order judgment. The principle of universality that Kant describes as a “standard [*Richtmaße*] of judgment” (GMS Ak. 04:404.01) is the standard that any rational judgment with any content whatsoever must meet. “The concept of [my] action in itself already contains a law for me” in both of these ways. To answer Millgram’s reframed question, then, a universalizable maxim is the form and principle of an agent’s will because its universalizable form is what makes it rational.

This brief and very incomplete textual summary of what a maxim is may begin to shed light on why Kant says at GMS Ak. 04:432.35-36 that the human will is a universal law-making will – not that it should be, or could be under ideal conditions, but that it *is*, in so far as it is rational. If a maxim of action is, as he has said, simply the formal principle of a certain kind of lawlike judgment that expresses our intention to act, then it is not difficult to see why he thinks this. For unless a judgment contains proper names or rigged definite descriptions, it is, in fact, universal in its scope of denotation.<sup>32</sup> Kant argues that the human will makes

---

<sup>32</sup> Thus I agree with O’Neill, against Hare, that universality can be predicated of the content of some judgments, but not of acts of judgment (see Nell, *op. cit.* Note 4, 17). However, I would add to O’Neill that universality can be predicated of the form of some judgments and not only of their content.

universal law through its maxims because maxims are judgments, and other things equal, this is the way nonindexical judgments ordinarily work.

This holds for judgments expressing the intention to act no less than for any other kind. All of the textbook counterexamples to the CI-procedure Millgram considers fail it:<sup>33</sup>

(14) I will buy clockwork trains but not sell them (548).<sup>34</sup>

(15) I will always be first through the door (548).

(16) I will turn up at Times Square tomorrow (548).

---

Rawls distinguishes between the *generality* of a principle – that it should contain no proper names or rigged definite descriptions, and its *universality* – that it should apply to “everyone in virtue of their being moral persons,” rather than merely to a restricted subset of such individuals (*A Theory of Justice* (Cambridge, Mass.: Harvard University, 1971), 131-132). I conflate this distinction, first, because Kant often does – he frequently uses *allgemein* and *universal* interchangeably; second, because if a principle is universal in Rawls’s sense then *a fortiori* it is general; third, because Rawls’s rendering of a universal principle fails to distinguish between its denotation and its application; and fourth, because “moral persons” denotes precisely the kind of restricted subset of individuals his account of universality is supposed to exclude, thereby begging all of the interesting questions; see my *Rationality and the Structure of the Self, Volume 1: The Humean Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20Vol%20the%20Structure%20of%20the%20Self%20Volume%20I-%20The%20Humean%20Conception.pdf>, APRA Foundation Berlin, 2008), Chapter X. “Rawl’s Instrumentalism” for a fuller treatment. Any judgment restricts its scope of application to that subset of individuals picked out by its referring terms. I call a judgment or principle *universal in its denotation* if, in virtue of containing no proper names or rigged definite descriptions, it refers indifferently to all such individuals.

<sup>33</sup> I offer some additional mischief-making maxims that the CI-procedure cannot so easily dispose of in *Rationality and the Structure of the Self, Volume II: A Kantian Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20and%20the%20Structure%20of%20the%20Self,%20Volume%20II-%20A%20Kantian%20Conception.pdf>, APRA Foundation Berlin, 2008), Chapter V. “How Reason Causes Action.”

<sup>34</sup> Nell, *op. cit.* Note 4, 76.

Millgram rightly deplores the New Kantian tendency to dispose of all such counterexamples by declassifying them, waiving the CI-procedure, or “setting them aside” (547 and fn. 44). This is to make the tail wag the dog, and draws the wrong conclusion from these “puzzle cases.” The correct conclusion is not that such puzzle cases are to be ignored or dismissed; but rather that they demonstrate something different from what New Kantians claim for the CI-procedure. What they show is that not all maxims of action are maxims of reason. Instead, some are arbitrary, idiosyncratic, or otherwise opaque to explanatory first principles in precisely the manner of (14), (15) or (16). Among those, only a subset corresponds to our intuitive notions of unethical behavior. The inscrutability of evil is merely a special case of the explanatory opacity of incompletely rational maxims of action more generally; and an expanded Davidsonian principle of charity (i.e. that failures of rationality imply that the agent is either stupid, crazy or evil) would counsel special treatment for all of them.

## V. Kant's Maxim Rationality Test

So far I have argued that Kant's twenty-four manner-of-willing formulations of the categorical imperative in the *Groundwork* are the result of applying the first *Critique's* Regulative Principle of Pure Reason to the special case of action: we are to will only that action for which we can find the universal first principle from which its maxim can be deduced. But how are we to ascertain which maxims bear this logical relationship to the universal first principles that govern them? In the first *Critique*, Kant also describes the general procedure by which we in fact identify such a universalizable maxim. This he calls the *Hypothetical* – actually the hypothesis-forming – *Employment of Reason*:

(17) If reason is a faculty of deducing the

Adrian M. S. Piper, Kant's Self-Legislation Procedure  
Reconsidered, KSO 2012: 203–277

Posted October 20, 2012 [www.kantstudiesonline.net](http://www.kantstudiesonline.net)  
© 2012 Adrian M. S. Piper & Kant Studies Online Ltd.

particular from the universal, then either the universal is already certain in itself and given, and so requires only judgment for subsumption, and the particular is thereby necessarily determined; ... or else the universal is given only problematically, and is a mere idea. Here the particular is certain, but the universality of the rule from which it follows is still a problem. Then more particular instances, each of which are certain, try on the rule for size to see whether they follow from it. In this case, if it appears that all particular instances assigned do follow from it, then we conclude to the universality of the rule, and thence to all instances, even those not themselves given (KrV, A 646/B 674.18-29, A 647/B 675.01-05).<sup>35</sup>

Passage (17) describes how we fix the relationship between particular and universal that enter into rational deduction and induction. “Particulars” may be either concrete particulars or lower-order concepts. The latter in turn include particular principles. All are possible candidates for subsumption under a universal, or for derivation from it. Here Kant is offering not only a description of how reason works, but also a method for testing the nomological scope of a rule relative to such candidates.

---

<sup>35</sup> Wenn die Vernunft ein Vermögen ist, das Besondere aus dem Allgemeinen abzuleiten, so ist entweder das Allgemeine schon an sich gewiß und gegeben, und alsdann erfordert es nur Urteilskraft zur Subsumtion, und das Besondere wird dadurch notwendig bestimmt. [...] Oder das Allgemeine wird nur problematisch angenommen, und ist eine bloße Idee, das Besondere ist gewiß, aber die Allgemeinheit der Regel zu dieser Folge ist noch ein Problem; so werden mehrere besondere Fälle, die insgesamt gewiß sind, an der Regel versucht, ob sie daraus fließen, und in diesem Falle, wenn es den Anschein hat, daß all anzugebenden besonderen Fälle daraus abfolgen, wird auf die Allgemeinheit der Regel, aus dieser aber nachher auf all Fälle, die auch an sich nicht gegeben sind, geschlossen (KrV, A 646/B 674.18-29, A 647/B 675.01-05).

There are two alternatives. In the first, the universality of the rule is certain and given by definition or analysis. In this case, judgment need only deduce the particular from it. In the second, the universal rule is merely stipulated hypothetically, whereas it is the particular that is certain and given. This latter case is the one in which we are reasoning about empirical conditions, or trying to explain empirical events and states of affairs. In this case, the method determines the scope of empirical application of the rule, by assessing its ability to subsume other, equally certain particulars as instances. If these can, indeed, be derived from the rule, we conclude to its universality and nomological necessity. In this case, the relation between particular and rule satisfies Kant's Regulative Principle of Pure Reason.

The procedure described by Kant's Hypothetical Employment of Reason ascertains the rational status of any concrete empirical particular, event, or state of affairs, as well as of any lower-order empirical concept or principle that might be invoked to describe it. By implication, therefore, it also ascertains the rational status of particular maxims of action. That is, the Hypothetical Employment of Pure Reason in general implies a Maxim Rationality Test for actions in particular:

(18) *The Maxim Rationality Test*

(1)	We begin with two elements:	
	(1.1)	A particular, given intention to act; this is what Kant calls an <i>instance</i> .
	(1.2)	A particular maxim that expresses that intention sententially, a provisionally universal categorical indicative judgment ascribing the action as a causal property to the agent; this is what Kant calls a <i>rule</i> .
(2)	Next we consider whether the maxim is or is not in fact universalizable, in answer to the following questions:	
	(2.1)	Does the particular intention accurately instantiate it?
		(2.1.1) If not, either the intention or the maxim must be revised.
		(2.1.2) If so, we conclude that the maxim veridically conceptualizes the agent's actual intention and therefore functions as a genuine lower-order rule that governs and subsumes that intention.
	(2.2)	Does the maxim itself in turn instantiate a higher-order rule that applies to other agents who intend the same action?
		(2.2.1) If not, we conclude to the maxim's violation of the universalizability requirement, and therefore that it fails the Regulative Principle of Pure Reason.
		(2.2.2) If so, we conclude that the maxim is, indeed, universalizable, and therefore satisfies the Regulative Principle of Pure Reason.

The Maxim Rationality Test derives from Kant's Hypothetical Employment of Reason in general a specific method for determining whether a maxim of action satisfies Kant's Regulative Principle of Pure Reason, and is therefore a maxim of reason. It departs from one particular ambiguity in Kant's text, by making a significant terminological substitution. Whereas passage (17) speaks of whether or not particular instances *follow from* the rule being tested, the Maxim Rationality Test instead speaks only of whether or not intentions or maxims *instantiate* the rule being tested. It excludes that reading of Kant's text that would require a deductive or materially implicative relationship between instance and rule, because no empirical rule logically or even materially implies all of its instances. The Maxim Rationality Test thereby avoids committing Kant to a form of Deductivism that is, as is argued in Section VII below, much too strong. Instead it offers a weaker and more plausible application of the reasoning procedure Kant describes at passage (17) that avoids this commitment.<sup>36</sup>

The Maxim Rationality Test tells us how to meet the requirements on action of Kant's Regulative Principle of Pure Reason. In order to ascertain what actions to will, we must first ascertain what maxims of action instantiate higher-order universal principles and finally the first principle of freedom;

---

<sup>36</sup> This procedure, and passage (17) more generally, bear a more than passing resemblance to Rawls's Reflective Equilibrium procedure, a method for achieving consistency between one's moral intuitions about particular cases and the overarching moral theory that is supposed to explain them. In both cases, one is required to both formulate the general rule in light of the particular instance, and also be prepared to jettison the particular instance should it fail subsumption under a plausibly formulated general rule. In both cases, a mismatch between them forces a choice about which to modify or abandon. And in both cases, the intended outcome is the same: that transparent deliberative link between the particular case in question and the universal first principle that is its ultimate explanatory ground. Up through his supervision of Susan Neiman's 1986 dissertation (later published as *The Unity of Reason: Rereading Kant* (New York: Oxford University Press, 1994) Rawls was very familiar with the Dialectic of the first *Critique*, so it is unlikely that this resemblance is accidental. See Rawls, *op. cit.* Footnote 32, especially his introductory description of the process of reaching reflective equilibrium at page 20.



i.e. what maxims of action are also maxims of reason. In order to ascertain that, we must first ascertain the logical relationship between the particular maxims our intentions instantiate, and the more general higher-order rules we take those maxims to instantiate in turn. That is, correct reasoning about action veridically subsumes intentions under maxims, and maxims in turn under higher-order universal principles, or maxims of reason. The maxim universalizability criterion Kant introduces in the manner-of-willing formulations of the categorical imperative in the *Groundwork* is in fact a highly condensed recapitulation of the much more detailed and complex account of reasoning about action that he offers in the Dialectic of the first *Critique*.

## VI. Exceptions

**H**aving ventured an answer to Millgram's modified question as to why a universalizable maxim is the form and principle of an agent's will, we are now in a better position to address his argument that successful agency requires making exceptions to such maxims as a matter of principle. Let us again take his advice, by looking in depth at what Kant actually says in the *Groundwork* about maxims that fail the Maxim Rationality Test (18) at step (2.2.1) above:

- (19) (a) If we now attend to ourselves whenever we transgress a duty, (b) we find that we do not really will such that our maxim should become a universal law, (c) because that is impossible for us; (d) but rather the opposite is itself in reality to remain universally a law. (e) We only take the liberty of making an *exception* to it for ourselves (or only just for this once) to the advantage of our inclination. (f) Consequently, if we were to consider everything from one from one and the same

standpoint, namely reason, (g) we would come across a contradiction in our own will, (h) namely that a particular principle should be objectively necessary as universal law, (i) and yet subjectively not be universally valid, but rather should allow exceptions. (j) But as we first consider our action from the standpoint of a will wholly in accord with reason, (k) but then exactly the same action from the standpoint of a will affected by inclination, (l) actually there is no contradiction here, (m) but rather an opposition of inclination to the prescription of reason (*antagonismus*), (n) through which the universality of the principle (*universalitas*) is turned into a mere generality (*generalitas*), (o) in such a way that the practical principle of reason is supposed to join up halfway with the maxim (GMS Ak. 04:424.18-39).<sup>37</sup>

Passage (19) applies the Maxim Rationality Test from the Dialectic of the first *Critique* to Kant's analysis of what goes

---

<sup>37</sup> (19) (a) Wenn wir nun auf uns selbst bei jeder Übertretung einer Pflicht Acht haben, (b) so finden wir, daß wir wirklich nicht wollen, es solle unsere Maxime ein allgemeines Gesetz werden, (c) denn das ist uns unmöglich, (d) sondern das Gegenteil derselben soll vielmehr allgemein ein Gesetz bleiben; (e) nur nehmen wir uns die Freiheit, für uns (oder auch nur für diesesmal) zum Vorteil unserer Neigung davon eine *Ausnahme* zu machen. (f) Folglich, wenn wir alles aus einem und demselben Gesichtspunkte, nämlich der Vernunft, erwägen, (g) so würden wir einen Widerspruch in unserem eigenen Willen antreffen, (h) nämlich daß ein gewisses Prinzip objektiv als allgemeines Gesetz notwendig sei (i) und doch subjektive nicht allgemein gelten, sondern Ausnahmen verstatten sollte. (j) Da wir aber einmal unsere Handlung aus dem Gesichtspunkte eines ganz der Vernunft gemäßen, (k) dann aber auch ebendieselbe Handlung aus dem Gesichtspunkte eines durch Neigung affizierten Willens betrachten, (l) so ist wirklich hier kein Widerspruch, (m) wohl aber ein Widerstand der Neigung gegen die Vorschrift der Vernunft (*antagonismus*), (n) wodurch die Allgemeinheit des Prinzips (*universalitas*) in eine bloße Gemeingültigkeit (*generalitas*) verwandelt wird, (o) dadurch das praktische Vernunftprinzip mit der Maxime auf dem halben Wege zusammenkommen soll (GMS Ak. 04:424.18-39).

wrong with a morally derelict maxim. Step (1) presents us with the two elements: (1.1) a particular, derelict intention to act as the *instance* in clause (a); and (1.2) its maxim as the *rule* in clause (b). In step (2), we try to fix the scope of the maxim – in (2.1.2), by acknowledging our derelict intention as instantiating the maxim we formulate to express it; and in (2.2), by trying to extend that maxim as a rule that applies to other, similarly situated agents by provisionally ascribing the intended action to them. In Step (2.2.1), “we find that we do not really will such that our maxim should become a universal law” [clause (b)], because we see that “this is impossible for us” [clause (c)]. That is, our derelict intention itself causes a loss of respectful attention to securing the lawlike form of its maxim (2.2), and thereby causes us to *lose the will* to formulate the maxim representing our intention (2.1.1) at all. Instead, we resituate the derelict intention (1.1) as an exception to a pre-existing universal principle that it violates [clauses (d)-(e)], “to the advantage of our inclination.” That derelict intention now stands alone as a causally efficacious inclinational event; a concrete particular instance without a rule or principle, i.e. without a maxim that subsumes it. Kant is here describing the morally fraught case in which we act deliberately and intentionally, but also deliberately relinquish the attempt to conceptualize to ourselves what it is exactly that we are doing. By temporarily disconnecting the operations of reason from their application to our own behavior, we circumvent its motivational power to direct or inhibit that behavior.

Now, from the unified perspective of reason, we both will the continuing universality of the pre-existing principle [clause (h)], and also will that it should not be universal but rather accommodate the derelict intention (1.1) as an exception [clause (i)]. When we will in the first case, we are approaching action from the standpoint of a rational, lawlike will [clause (j)]; when we will in the second case, we are approaching action from the standpoint of “a will affected by inclination” [clause (k)]. In both cases, the backward-looking intentional object of our willings is the contested

scope of the pre-existing principle, not the maximless derelict intention that conflicts with it. It is the muting gravitational pull on the will of that unconceptualized intention that creates the “will affected by inclination,” i.e. a *weakened* willing and failure of resolve that progressively diminishes the scope of one’s willing to the concrete particular present moment of the inclinational event. Conjointly, then, these two contradictory willings reduce the universal scope of the pre-existing principle to “a mere generality” [clause (n)], with which the derelict intention (1.1) is consistent as an exception.

In accordance with the argument of Section III above, these two contradictory willings do not constitute a contradiction embedded within a forward-looking intentional object, in which I will both the universality of a maxim and simultaneously an exception to it. This would be to will a self-contradictory object, and can be represented symbolically using a notation that integrates the language of preference used by Ramsey-Savage decision theory into the classical predicate logic on which Kant relied.<sup>38</sup> Reading  $P_w$  as “ $w$  wills [or strictly prefers] ...” (i.e. as a predicate rather than an operator), we represent the case in which I will a self-contradictory object that includes both the universal maxim and the exception to it as

$$(20) P_w[(\exists a)(x)Fx . \sim Fa].$$

Here the contradiction is not in my will *per se*, as Kant’s text requires, but rather in the object of my will. Nor, similarly, do these two contradictory willings constitute a contradiction between two contradictory intentional objects; for

---

<sup>38</sup> I develop this quantification-based notation for the logic of preference in depth in *Rationality and the Structure of the Self, Volume II: A Kantian Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20and%20the%20Structure%20of%20the%20Self,%20Volume%20II-%20A%20Kantian%20Conception.pdf>, APRA Foundation Berlin, 2008), Chapter III. “The Concept of a Genuine Preference.”

(21)  $Pw[(x)Fx] . Pw[\sim Fa]$

is not a contradiction at all. Passage (19) contains no suggestion that the “contradiction in the will” described in clause (g) depends on any particular forward-looking intentional object of the will – much less on any such pair of objects that contradict or conflict with each other. It thus does not invite an interpretation of Kant’s discussions of cultivating one’s talents or helping others that elaborates those analyses in terms of wanting both  $x$  and not- $x$ , or both  $x$  and  $y$  when these are conceived as being in conflict, or in terms of what we would want or not want, were certain counterfactual conditions to apply. When Kant says about the man contemplating whether or not to neglect the cultivation of his natural gifts,

(22) He sees here, that a system of nature could always indeed exist under such a universal law ...; only he cannot possibly *will* that this become a universal law of nature ... (GMS Ak. 04:423.09-16),<sup>39</sup>

or about the man deliberating about whether or not to help others in need,

(23) But although it is possible that a universal law of nature could exist according to this maxim, it is nevertheless impossible to *will* that such a principle should hold everywhere as a law of nature (GMS Ak. 04:423.33-37),<sup>40</sup>

---

<sup>39</sup> Da sieht er nun, daß zwar eine Natur nach einem solchen allgemeinen Gesetze immer noch bestehen könne, ...; allein er kann unmöglich *wollen*, daß dieses ein allgemeines Naturgesetz werde ... (GMS Ak. 04:423.09-16).

<sup>40</sup> Aber obgleich es möglich ist, daß nach jener Maxime ein allgemeines Naturgesetz wohl bestehen könnte, so ist doch unmöglich, zu *wollen*, daß ein solches Prinzip als Naturgesetz allenthalben gelte (GMS Ak. 04:423.33-37).

his argument is exactly the same in both cases. Someone with a derelict intention seeks to justify it by willing an action he tries unsuccessfully to represent as a law that neither he nor anyone else is to cultivate their talents or help others. He also “does not really will” this, because he knows he can indulge his own passive neglect only if others do not indulge theirs. Such a will would, indeed, “be in conflict with itself, since many cases can arise in which he needs [the advantage of others’ efforts] ...” (GMS Ak. 04:423.37-38).<sup>41</sup> So he intends, rather, that “the opposite [of his derelict intention] is itself in reality to remain universally a law” [clause (d) of passage (19)], so as to continue to enjoy the advantages of others’ efforts. At the same time that his derelict intention subverts his respectful attention to the lawlike form of his maxim, the lawlike form of his reason more generally subverts his success in ignoring it.

Thus clauses (f) – (i) of passage (19) call our attention to a straightforward logical contradiction of the old-fashioned kind, between two incompatible backward-looking intentional *attitudes*, in which I both do and “do not really will” [clause (b)] a universalizable maxim, such that

$$(24) Pw[(x)Fx] . \sim Pw[(x)Fx].$$

As Kant says, there is no inconsistency involved in willing as a universal law the practice of neglecting one’s talents or refusing to render aid, and the corresponding derelict intention motivates one to do so. Yet it also discourages one from doing so at the same time, on pain of thwarting one’s advantage in enacting it. What is “impossible for us” is not willing an exception *to* a universal law; but rather to “really will” the exception *as itself* a universal law.

---

<sup>41</sup> (a) Denn ein Wille, der dieses beschlösse, (b) würde sich selbst widerstreiten, (c) indem der Fälle sich doch manche ereignen können, (d) wo er anderer Liebe und Teilnehmung bedarf, ... (GMS Ak. 04:423.37-38). Notice Kant’s use of the indicative voice in clauses (c) and (d). He clearly means here to be describing the empirical reality that explains why such a contradiction in the will would be at best a counterfactual conditional supposition.

This conclusion holds by virtue of Kant's conception of what an exception is. An exception, for Kant, is not simply a state of affairs that departs from some particular rule. It is a state of affairs that we choose to exempt from *any* rule, i.e. that we decline to universalize at all. Under these circumstances, it is just not logically possible to have it both ways: to claim both exceptional status for that state of affairs, and also its universalizability in its own right. This is a perfectly general point that holds independently of the particular content of that state of affairs. *A fortiori*, it holds for states of affairs that are derelict intentions, independently of their particular content. *Whatever that content is*, we cannot both exempt the derelict intention from the ongoing lawlike operations of reason on the one hand, and ascribe to it a lawlike rational status on the other. Therefore, it really is not possible for us to both enact our unconceptualized derelict intention, and also to conceptualize or represent it as one that everyone might have at this moment. That would be to forego all the collective advantages of reason and system that enable us to indulge it. Maxims of action that try to subsume derelict intentions fail to achieve the status of a maxim of reason because one's formulation of it fails to "pull it all together." That is, it fails to pull the exception together with the other particulars it subsumes. Because this is the case independently of the particular content of that derelict intention, it therefore holds independently of its particular goal. Thus the contradiction in the will that Kant describes in clause (g) of passage (19) is between two incompatible *manners* of willing, not in the intentional *object* that we will.

Now the core of Millgram's critique of the CI-procedure was the New Kantians' problematic treatment of exceptions. Kant's own treatment of exceptions in passage (19) is rather less problematic. Just as he earlier described in the first *Critique* the way reason must function in order for us to have unified experience and take unified action, here in the *Groundwork* he similarly describes how we must deal with the sorts of exceptions that Millgram's critique targets. Kant does not think we can successfully modify a universal

principle by building an exception into it – not even provisionally. Nor does he think we can build in generic, *ceteris paribus* exception clauses into such principles. Nor does he ever deny the necessity of our trying to will the universality of all of our maxims. Nor, therefore, can derelict maxims be excluded from the class of *bona fide* maxims. Nor does the agent have discretion to decide what is and what is not a maxim. On Kant’s model of reason, our innate cognitive structure is such that all of these *ad hoc* New Kantian moves are foreclosed to us, and Millgram is quite right to reject them. Because maxims express intentions to act, and because we are moved to will universalizable maxims, we have only one recourse for dealing with maxims that violate this requirement. That is, as clauses (n) – (o) describe, to pretend it is not really a requirement at all. This, for Kant, is where bad faith – evil, actually – begins.

## VII. Kant’s Self-Legislation Procedure Reconsidered

So far I have tried to show, first, that most of Kant’s thirty-eight “universal law” formulations of the categorical “imperative” (all are listed in Appendix Table 3) can be sorted into two subsets: those which enjoin the universalizability of our maxim as a *goal or incentive* of the will, and those which enjoin it as a lawlike *manner* of willing. Second, I have tried to show that this second set of formulations, and passage (8) as a particular example, is an application of his Regulative Principle of Pure Reason from the first *Critique* to the special case of action. Third, I have argued that Kant’s Hypothetical Employment of Reason in the first *Critique* offers a method for determining whether or not a particular satisfies the lawlike requirements of the Regulative Principle of Pure Reason. Fourth, I have derived from Kant’s description of how the Hypothetical Employment of Reason functions a Maxim Rationality Test for determining whether or not a particular maxim of action satis-



fies the lawlike requirements of the Regulative Principle of Pure Reason.

However, I have not argued that Kant's Maxim Rationality Test enables us to derive our moral duties from the categorical imperative. I think Kant should not have claimed this,<sup>42</sup> and that the New Kantians make a mistake in trying to defend it. As I have tried to show both here and elsewhere, their assumption that a maxim, and indeed an intention – is a prescriptive judgment in the imperative mood about what I ought to do<sup>43</sup> does not square well with the textual evidence, either in the first *Critique* or in the *Groundwork*. A maxim is, indeed, a *regulative* principle that guides and directs our reasoning in such a way as to extend it past the limits of present empirical experience (KrV, A 666/B 694.13-22) – most importantly, into future action. But that does not make it *prescriptive* in form. A maxim is a descriptive judgment about what my intention actually is, not a prescriptive judgment about what I ought to do.

Correspondingly, the law that universalizes that maxim is similarly a descriptive principle in the indicative mood that denotes the ideal behavior of a fully rational agent, not an imperative that exhorts us to live up to it. The reason why

---

<sup>42</sup> And Kant definitely does claim this, at GMS Ak. 04:421.10-12: “Wenn nun aus diesem einigen Imperative alle Imperativen der Pflicht, als aus ihrem Prinzip, abgeleitet werden können, so werden wir, ob wir es gleich unausgemacht lassen, ob nicht überhaupt das, was man Pflicht nennt, ein leerer Begriff sei, doch wenigstens anzeigen können, was wir dadurch denken und was dieser Begriff sagen wolle.“ At GMS Ak.04:423.43-44, 424.01-02, the German Academy edition runs, “Dieses sind nun einige von den vielen wirklichen oder wenigstens von uns dafür gehaltenen Pflichten, deren Abtheilung aus dem einigen angeführten Prinzip klar in die Augen fällt.“ Timmerman's translation is true to this text, whereas I would substitute „Ableitung“ [derivation] for „Abtheilung“ [division]. The reason is the relation between this passage and the earlier one at GMS Ak. 04:421.10, where Kant uses and clearly means to use the verb „abgeleitet.“ If we take his pronouncement there at face value, then his use of „Abtheilung“ at GMS Ak. 04:423.43 may have been a slip of the pen.

<sup>43</sup> Rawls embraces this assumption in his “Themes in Kant's Moral Philosophy,” in Eckart Förster, Ed., *Kant's Transcendental Deductions: The Three 'Critiques' and the 'Opus Postumum'* (Stanford: Stanford University Press, 1989), 83. O'Neill almost escapes its infelicitous consequences by distinguishing between a universalized maxim, which she characterizes in the imperative mood, and its *universalized typified counterpart* (UTC), which she frames in the indicative. *Op. cit.* Note 4, 61-62.

the *Groundwork* contains only four formulations of Kant's basic principle of morality in the strict imperative mood<sup>44</sup> is because Kant was more interested in analyzing the complex relation of imperfectly rational human beings to the regulative, descriptive rational ideal that they naturally generate to guide their behavior. So although Millgram is right that successful agency under conditions of risk or uncertainty requires occasional exceptions to such rules, these conditions do not apply to the circumstances that the categorical "imperative" envisions. Rather, these ideally rational circumstances assume full and public information that enables each such agent to easily coordinate her behavior with others'. It is only because imperfectly rational human agents are compelled by their own cognitive structure to guide their actual behavior by principles designed for fully rational ones that they reflexively excuse their derelictions by demoting those principles to "mere generalities."

Now about O'Neill's contradiction tests Millgram says, and most of the New Kantians agree,<sup>45</sup> that

(25) [A]lthough I have described the procedure as something you pause to execute before going ahead with an action you have in mind, the Kantian requirement is of course not that you stand around muttering to yourself before you do anything. The procedure is meant as a rational reconstruction of the deliberative background to a decision properly arrived at  
.... (527)

---

<sup>44</sup> These four are listed in Appendix Table 2.1. Table 2.2 lists the only two formulations that uses the Subjunctive I form of the imperative to modify a *wollen-* or *handeln-*type verb. There are other formulations, listed in Appendix Table 1, that use the Subjunctive I form of the imperative in subsidiary positions, otherwise than to exhort the reader to will or act in a certain way. I leave their identification as an exercise for the reader.

<sup>45</sup> See John Rawls, "Themes in Kant's Moral Philosophy," *op. cit.* Note 43, 82. Also see his *Lectures on the History of Moral Philosophy*, Ed. Barbara Herman (Cambridge: Harvard University Press, 2000), 218. O'Neill, however, does not share this view. See *Acting on Principle*, *op. cit.* Note 4, 71-72, 74.

Actually that is exactly *not* what Kant himself meant. It is clear from the language he uses when he is introducing or commenting upon the four examples, both in Chapters I and II of the *Groundwork*, that he intends it precisely as “something you pause to execute before going ahead with an action you have in mind.” Thus he says, in introducing the false promising case, “In order to instruct myself in the quickest and surest way whether a lying promise accords with duty . . . , I ask myself: etc.” (GMS Ak. 04:403.03-06).<sup>46</sup> In his concluding remarks about that case, he says, “Inexperienced in considering the ways of the world, unable to prepare myself for all the incidents that occur in it, I simply ask myself: etc.” (GMS Ak. 04:403.24-26).<sup>47</sup> In setting up the four examples in Chapter II, he describes the potential suicide as a man who “is still so far in possession of his reason that he can ask himself, etc.” (GMS Ak. 04:422.02-03).<sup>48</sup> He revisits the false promising example by describing the agent as someone who “has a mind to make such a [false] promise; but still has sufficient conscience to ask himself: etc.” (GMS Ak. 04:422.22-23).<sup>49</sup> And he describes the man of uncultivated talents in the third example as someone who “sees himself in comfortable circumstances . . . [y]et asks further, etc.” (GMS Ak. 04: 423.02-03, 05).<sup>50</sup>

In all of these cases, Kant clearly means to be describing an actual empirical event, i.e. the process of interrogative

---

<sup>46</sup> Um indessen mich in Ansehung der Beantwortung dieser Aufgabe, ob ein lügenhaftes Versprechen pflichtmäßig sei, auf die aller kürzeste und doch untrügliche Art zu belehren, so frage ich mich selbst: usw. (GMS Ak. 04:403.03-06).

<sup>47</sup> Unerfahren in Ansehung des Weltlaufs, unfähig auf alle sich ereignenden Vorfälle desselben gefaßt zu sein, frage ich mich nur: usw. (GMS Ak. 04:403.24-26).

<sup>48</sup> Einer ... ist noch soweit im Besitze seiner Vernunft, daß er sich selbst fragen kann, usw. (GMS Ak. 04:422.02-03).

<sup>49</sup> Er hat Lust, ein solches Versprechen zu tun; aber noch hat er so viel Gewissen, sich zu fragen: usw. (GMS Ak. 04:422.22-23).

<sup>50</sup> Er sieht sich aber in bequemen Umständen [...]. Noch fragt er aber, usw. (GMS Ak. 04:423.02-03, 05).

reasoning one undertakes in order to reach a concrete decision about whether or not to embark on some specific, envisaged course of action. That process is captured by his description of the Hypothetical Employment of Reason discussed above. It is clearly a method we are supposed to actually *employ* when reasoning, and we do. The contradiction in the will test is Kant's philosophical elaboration of the common and very commonsensical counterfactual question, "How would you like it if someone did that to you?" And the contradiction in conception test is Kant's philosophical elaboration of the equally common and commonsensical counterfactual question, "What if everybody did that?" We do, in fact, ask ourselves these very questions when debating a course of action, whether in solitude or in society. O'Neill's two contradiction tests reconstruct and refine what Kant takes to be the fundamental mode of moral reasoning actual human agents undertake when they "stand around muttering to [themselves] before [they] do anything." And Kant believes that answering these questions is sufficient to conclude it. All of Kant's examples describe an agent who is viscerally attempting to answer them; and for whom, consequently, this attempt is the object of a forward-looking, goal-directed intentional attitude of inquiry.

Kant thought to answer them by applying to actual, forward-looking, goal-directed deliberation the very same backward-looking formal requirement by which he had previously defined the capacity to reason in general, namely that our thoughts be organized in a lawlike manner. If, as he earlier argued (KrV, A 652/B 680.04-12), his Regulative Principle of Pure Reason leads us to seek that most abstract, universal and all-inclusive law from which everything else can be deduced, then *a fortiori* "the concept of [our] action in itself" also must "already contain a law for [us]," from which that particular action similarly can be deduced. His Hypothetical Employment of Reason aims to describe the empirical procedure by which we ensure that this deductive relationship between law and instance is sound, whatever the instance might be.

This is the reason for the inconsistency noted in Section II, between Kant's purely formal analysis of the moral law, and his ascription of it to the will as a forward-looking incentive in his third formulation of the categorical "imperative" at passage (5) above. For this consists in the injunction to apply the form of reason as itself a method of reasoning. Here is the later formulation of the categorical "imperative" (the thirty-ninth, according to Appendix Table 1) that expresses this injunction most explicitly:

- (26) But the sentence: the will itself is a law in all actions, indicates only the principle to act according to no other maxim, *than that which can also have for an object itself as universal law* (GMS 04:447.05-08; italics added).<sup>51</sup>

Here, as in passage (5) and others like it, Kant is speaking to the reader who is actively seeking a conscious decision-making method for determining the right action. He is urging us to adopt *as a forward-looking empirical goal* the backward-looking intentional attitude of "concern not with the matter of the action and what is supposed to result from it, but rather with the form and principle from which it itself follows." When he exhorts us to will as a forward-looking, goal-directed intentional object *that* our maxims be universalizable, he is telling us to adopt the formal criterion of rationality – bare conformity to law – as itself a concrete empirical procedure for making a decision; whereas when he exhorts us simply to will a universalizable maxim, he is telling us what that formal criterion is.

All seven of the "universal law" formulations of the categorical "imperative" listed in Appendix Table 4 that exhort us to will *that* our maxims be universalizable urge us

---

<sup>51</sup> Der Satz aber: der Wille ist in allen Handlungen sich selbst ein Gesetz, bezeichnet nur das Prinzip, nach keiner anderen Maxime zu handeln, als die sich selbst auch als ein allgemeines Gesetz zum Gegenstande haben kann (GMS 04:447.05-08).

to strive to cultivate the backward-looking intentional attitude of respectful attention to the lawlike form of our maxims; and to strive to satisfy the three conditions of mindfulness that constitute it in every maxim we formulate, as itself a procedure for determining what to do. They enjoin on us a way of *thinking*, namely thinking in a lawlike manner, from which the right action for us to take at this place and time can be rationally derived. Kant sees that we already have the capacity for this mode of thought, and that we already employ it when reasoning about abstract theoretical subjects. He is simply urging us to employ *exactly the same mode of reasoning* about ourselves; to treat ourselves and our actions not as exceptions to the universal rules we invariably generate, but rather as instances of them.

So the point of the project in these formulations is not to submit our actions to a test of moral validity, or to cast their maxims in such a manner as to pass it, or to protect or enhance our “stake in our own agency.” Nor are we to strive to conform the maxims of our actions to the CI-procedure, the CI-Maxim, the Revision Maxim, or anything similar. Rather, the point is to strive to adopt the same rational mode of thinking about ourselves and our actions as we do about any other object of inquiry; indeed, to subordinate the provincial, constricting demands of ego and inclination to the higher, liberating demands of this more distanced and comprehensive mode of thought. Once we become skilled and comfortable at regarding ourselves and our actions as concrete particulars that succeed or fail *gradatim* the objective demands of reason, Kant thinks, the answer to our moral questions as to what to do will follow naturally. These formulations of the categorical “imperative” recommend that we use the same Hypothetical Employment of Reason in answering the moral question as in answering any other.

Kant failed at this project for three reasons. First, it really is not possible to derive substantive moral directives from a purely formal and functional analysis of reason; this is the proverbial “garbage in-garbage out” problem that Deduc-

tivists always encounter<sup>52</sup> and that logicians know as Mill's Paradox of Inference.<sup>53</sup> Second, as we have seen, Kant in the *Groundwork* did not distinguish clearly or consistently between backward-looking and forward-looking intentional attitudes; nor, therefore, between the backward-looking formal conditions a maxim of action must satisfy in order to be rational on the one hand, and the pressing forward-looking empirical task of deciding whether or not to perform it, which its satisfaction of those conditions is supposed to determine, on the other.

Third, Kant's taxonomy of the moral duties that the contradiction tests are supposed to imply is faulty. If, as the

---

<sup>52</sup> I discuss the problem of Deductivism in *Rationality and the Structure of the Self, Volume I: The Humean Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20and%20the%20Structure%20of%20the%20Self%20Volume%20I-%20The%20Humean%20Conception.pdf>, APRA Foundation Berlin, 2008), Chapter IX. "The Problem of Moral Justification."

<sup>53</sup> "The maxim . . . , That whatever can be affirmed (or denied) of a class, may be affirmed (or denied) of everything in the class . . . , supposed to be the basis of the syllogistic theory, is termed by logicians the *dictum de omni et nullo* (234). . . . Now, however, when it is known that a class, an universal, a genus or species is not an entity per se, but neither more nor less than the individual substances themselves which are placed in the class, and that there is nothing real in the matter except those objects, a common name given to them, and common attributes indicated by the name; what, I should be glad to know, do we learn by being told, that whatever can be affirmed of a class, may be affirmed of every object contained in the class? The class is nothing but the objects contained in it: and the *dictum de omni* merely amounts to the identical proposition, that whatever is true of certain objects, is true of each of those objects. . . . To give any real meaning to the *dictum de omni*, we must consider it not as an axiom but as a definition; we must look upon it as intended to explain, in a circuitous and paraphrastic manner, the meaning of the word *class* (235-6)." John Stuart Mill, *A System of Logic, Ratiocinative and Inductive, Volume I* (London: John W. Parker, 1843). Cf. also 217-222, on the variety of cases in which "the proposition ostensibly inferred from another, appears on analysis to be merely a repetition of the same, or part of the same, assertion, which was contained in the first (217). . . . In all these cases there is not really any inference; there is in the conclusion no new truth, nothing but what was already asserted in the premises, and obvious to whoever apprehends them. The fact asserted in the conclusion is either the very same fact, or part of the fact, asserted in the original proposition (219)." For a contemporary treatment, see Jaakko Hintikka, *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic* (New York: Oxford University Press, 1973 and 2002), 222 on the "scandal of deduction."

first formulation of the categorical imperative claims,<sup>54</sup> the moral law directs us in moral deliberation *always* to will a universalizable maxim, and our derelict maxim violates the moral law, then it produces a contradiction in the will: I both want the moral practice to remain in force, and also do not want this for my own case. *This holds as well for lying as for loafing.* However, if the contradiction in the will test identifies imperfect duties that we have latitude and discretion about when and how to fulfill, then obedience to the moral law is an imperfect duty, and we have that same latitude and discretion about when and how to obey it. This contradicts Kant's earlier claim that we are always and only to act on those maxims that do obey it. The contradiction thus forces a choice between Kant's analysis of the moral law as instantiating a law of reason on the one hand, and his taxonomy of the moral duties purportedly derived from it on the other. I choose the former: Kant's warning that we are not to short-circuit the ongoing deductive and inductive operations of reason for purposes of self-indulgence or self-aggrandizement expresses the deeper insight into the essence of vice and virtue.

### VIII. O'Neill's Contradiction Tests

Nevertheless, Kant's formalism is far from empty. His two contradiction tests as O'Neill dissects them still yield important conclusions about the limits and potentials of rational agency. We see this by scrutinizing the passage at GMS Ak. 04:424 directly preceding passage (19) above, which O'Neill's analysis takes as its starting point:<sup>55</sup>

---

<sup>54</sup> ... (a) I am never to proceed otherwise than (b) such that I could also will (c) that my maxim should become a universal law (GMS Ak. 04:402.09-11).” In the German original, [...] ich soll niemals anders verfahren als so, daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden (GMS Ak. 04:402.09-11).“

<sup>55</sup> Nell, *op. cit.* Note 4, 60.



- (27) (a) One must be able to will that a maxim of our action become a universal law: (b) this is the canon of moral judgment of action in general. (c) Some actions are so constituted that their maxims cannot even be *thought* without contradiction as a universal law of nature, (d) much less *willed* such that they *should* become one. (e) With others this inner impossibility is not to be found, (f) but it is impossible to *will* that their maxim should be elevated to the universality of a law of nature, (g) because such a will would contradict itself. (GMS Ak. 04:424.02-12)<sup>56</sup>

Clauses (a) and (b) identify the canon of moral judgment in general with the formulation of the moral law at passages (5) and (26) above: Our maxim of action should represent the action as becoming universal law, and we must be able to will this as a forward-looking intentional object. This means that when seeking consciously to answer the question as to what to do, we are to apply the backward-looking rational requirement, that the maxims of our will have a lawlike character, to the forward-looking intentional object of inquiry, namely the matter of how to settle on what to do. We are to strive to formulate the maxim of our envisioned action in a lawlike manner, such that it exhibits respectful attention to the “form and principle from which it itself follows,” and therefore passes the Maxim Rationality Test.

---

<sup>56</sup> (a) Man muß *wollen können*, daß eine Maxime unserer Handlung ein allgemeines Gesetz werde: (b) dies ist der Kanon der moralischen Beurteilung derselben überhaupt. (c) Einige Handlungen sind so beschaffen, daß ihre Maxime ohne Widerspruch nicht einmal als allgemeines Naturgesetz *gedacht* werden kann; (d) weit gefehlt, daß man noch *wollen* könne, sie *sollte* ein solches werden. (e) Bei anderen ist zwar jene innere Unmöglichkeit nicht anzutreffen, (f) aber es ist doch unmöglich zu *wollen*, daß ihre Maxime zur Allgemeinheit eines Naturgesetzes erhoben werde, (g) weil ein solcher Wille sich selbst widersprechen würde (GMS Ak. 04:424.02-12).

Thus the canon of moral judgment in general enjoins us to strive to formulate our intention in acting as a maxim of reason; this, Kant implies, will suffice to answer the question as to what to do.

Clauses (f) and (g) introduce the analysis in passage (19) of in what a contradiction in the will consists. We have already seen that our attempt to conform the maxim of our derelict intention to the rationality requirements of the Regulative Principle of Pure Reason creates a contradiction between this manner of willing and that in which “we do not really will” its universalizability. I have argued above in Section VI that this is not a contradiction embedded within a forward-looking intentional object, such that  $P_w[(\exists a)(x)Fx \cdot \sim Fa]$ , nor a conjunction of two incompatible intentional objects, such that  $P_w[(x)Fx] \cdot P_w[\sim Fa]$ ; but rather a straightforward logical contradiction between two contradictory backward-looking intentional attitudes, such that  $P_w[(x)Fx] \cdot \sim P_w[(x)Fx]$ . The implication is that when Kant maintains that “it is impossible to will that their maxim should be [universal] ... because such a will would contradict itself” [clauses (f) and (g) above], that is exactly what he means: not that the *object* of the will would contradict itself, but rather that the *will itself* would contradict itself.

However, when we consciously apply the canon of moral judgment in order to ascertain the moral acceptability of a derelict maxim such as

(28)  $\sim Fa$ : I will not help others in need,

we generate the intentional representation of this contradiction as itself the forward-looking object of our will:

(29)  $P_w\{P_w[(x)Fx] \cdot \sim P_w[(x)Fx]\}$ .

Here I want both to will the universalizability of the negation of my derelict maxim, namely the pre-existing moral principle – so as to enjoy the benefits of others’ help; and also not to will its universalizability – so as to enjoy the

benefits of not reciprocating it. This is, in effect, to will a self-contradictory intentional object, namely that contradiction in the will itself. When Kant then says in clause (f) that it is “impossible” to will in this manner, and in clause (g) that this is “because such a will would contradict itself,” he means simply to point out that it is *logically, conceptually* impossible to will a will that contradicts itself. Thus is the irrationality of “our own will” [passage (19), clauses (g)-(i)] made visible to us in conscious deliberation.

Now for a closer look at clauses (c) and (d) of passage (27). These describe a second kind of case, to which Kant’s analysis at passage (19) applies equally well: I both will the pre-existing moral principle and, in order to exempt my own action, “do not really will” it. But in this case, my application of the canon of moral judgment in general to the derelict maxim multiplies the contradiction fourfold. For now, not only do I will a conjunction of two contradictory willings as a forward-looking intentional object, as in formula (29). Now the forward-looking intentional objects of my two contradictory willings themselves also contradict themselves: The very “concept of [my] action in itself [as] already contain[ing] a law for me” is internally contradictory. There can be no manner of willing this maxim itself for which a backward-looking intentional attitude of “concer[n] ... with the form and principle from which it itself follows” could be cultivated, because the maxim as universalized has no coherent form and there is no consistent principle from which it follows. The universalization of this kind of derelict maxim is literally inconceivable.

The best example of this kind of action is the one Kant discusses directly following his first formulation of the categorical imperative at GMS Ak. 04:402.09-11, which he describes as a “lying promise” (GMS Ak. 04:403.04), and elaborates further at GMS Ak. 04:422.18-44. I am going to abide by Kant’s initial description of this as a *false* promise, in order to sift out the complexities that attach to his analyses

of lying and dishonesty more generally.<sup>57</sup> Rather, I want to focus specifically on the case in which, as Kant describes it, I will “make a promise with the intention of not keeping it” (GMS Ak. 04:402.20). When I try to universalize this maxim in the counterfactual thought experiment in which everyone makes promises with the intention of not keeping them, I find that the envisioned universal practice is not one of promising of any kind, but instead consists in a meaningless utterance without social effect. As Kant says,

(30) [I]n accordance with such a universal law

---

<sup>57</sup> The main problem I find with Kant’s arguments against lying is that it is easy to imagine a world in which everybody lies all the time, so long as not everybody admits it. A world in which everyone in fact hates themselves, their lives and one another, yet systematically avows and otherwise represents exactly the opposite, in order to lure in the unsuspecting to share the burden of their misery would illustrate such a possibility. False promising is different because the behavioral breach itself usually reveals the dereliction; whereas the failure to tell the truth may never come to light, so long as everyone cooperates in concealing it – for example, in order not to look foolish for having been duped in the first place, or to justify with each new recruit having been duped themselves. Thus a universal practice of lying would pass the contradiction in conception test, although a specific instance of it – false promising – would not. Thus I take issue here with O’Neill’s argument that:

falsehood in communication could not serve as a universal principle for communications among a plurality of rational beings ... [because] comprehension itself would cease, and so also the possibility of communication. ... [T]he very possibility of recognizing what is said in such contexts as falsehood presupposes comprehensibility, and thus also that standards of truth telling obtain more generally in those communities (*Constructions of Reason: Explorations of Kant’s Practical Philosophy* (Cambridge: Cambridge University Press, 1989), Chapter 2. “The Public Use of Reason,” 45).

Systematic falsehood – of the sort we practice on children in order to shield them from ugly realities, or on those we mean to use in ways to which they would not consent if consulted, for example – would not, so far as I can see, cause comprehension or communication to cease. As to whether anyone *recognizes* what is said under such circumstances as falsehood, this is beside the point. “Sugaring the pill,” “spinning the bad news,” “saving face,” “doing damage control,” etc. are only a few of the favored alternative practices that do just as well. See my “Kant’s Two Solutions to the Free Rider Problem,” *The Kant Yearbook 4/2012: Kant and Contemporary Moral Philosophy*, 113–142 for an extended discussion.

Adrian M. S. Piper, Kant's Self-Legislation Procedure  
Reconsidered, KSO 2012: 203–277

Posted October 20, 2012 [www.kantstudiesonline.net](http://www.kantstudiesonline.net)  
© 2012 Adrian M. S. Piper & Kant Studies Online Ltd.

there actually would be no promising at all, for it would be futile to profess my will with regard to my future actions to others who would not believe me ... (GMS Ak. 04:403.14-18); ... [it] would make promising and the very point of promising itself impossible, as no one would believe they were promised anything ... (GMS Ak. 04:423.40-44).<sup>58</sup>

In this case, when I consciously apply the canon of moral judgment to my derelict maxim as a forward-looking intentional object, I find I am willing both a universal practice of false promising *and thereby* its negation:

$$(31) Pw[(x)(\sim Fx \rightarrow \sim Fx)].$$

But remember that I am both willing this practice and also not willing it:

$$(32) Pw[(x)(\sim Fx \rightarrow \sim Fx)] . \\ \sim Pw[(x)(\sim Fx \rightarrow \sim Fx)],$$

and, in addition, consciously applying the canon of moral judgment in order to determine its moral acceptability:

$$(33) Pw\{Pw[(x)(\sim Fx \rightarrow \sim Fx)] . \\ \sim Pw[(x)(\sim Fx \rightarrow \sim Fx)]\}.$$

What a mess!

---

<sup>58</sup> [...] denn nach einem solchen würde es eigentlich gar kein Versprechen geben, weil es vergeblich wäre, meinen Willen in Ansehung meiner künftigen Handlungen anderen vorzugeben, die diesem Vorgeben doch nicht glauben [...] würden [...] (GMS Ak. 04:403.14-18) [...] [es] würde das Versprechen und den Zweck, den man damit haben mag, selbst unmöglich machen, indem niemand glauben würde, daß ihm was versprochen sei [...] (GMS Ak. 04:423.40-44).

## IX. Millgram's Test of Reflexive Application Reconsidered

**K**ant's argument that false promising cannot be universalized proposes a resolution to the Free Rider problem. Hobbes' Foole<sup>59</sup> was the first to reason that if self-interest justifies exchanging the state of nature for the Social Contract, then self-interest also justifies violating the Social Contract for personal gain. The Free Rider, then, makes a really quite important promise, namely to obey the rules as consensually agreed, with the intention of not keeping it when that is to her benefit. She takes advantage of others' sacrifice of immediate self-interest in order to advance her own. Kant was fully aware of the coordination problem posed by Hobbes' Foole, and offered a two-pronged solution to it. I discuss Kant's application of the solution he offers at KrV, A 751/B 779.30-40 and A 752/B 780.01-14 to passage (19) at greater length elsewhere.<sup>60</sup> As we have just seen in Section VIII, Kant's argument in passages (19) and (27) conjointly is that the Free Rider induces in herself not merely a contradiction in her will, but also a straightforwardly logical contradiction in her conception of what it is she is intending to do: Formulas (31) – (33) express the case in which she both does and also does not will both to break a promise and not to break it, therefore neither to break it nor to keep it; both to presuppose and to deny a consensual practice of promise-keeping; both to invoke and to reject this practice to rationalize her derelict intention. She thereby short-circuits the very powers of reasoning she meant to harness in the service of self-interest. Thus Kant critiques the Free Rider's reasoning by showing that false promising is *individually* irrational in this double sense: It fails to advance her self-interest, because it actively undermines the cognitive coherence of the self whose in-

<sup>59</sup> Cf. Thomas Hobbes, *Leviathan*, Ed. Michael Oakeshott (New York: Macmillan/Collier Books, 1977), 115-117.

<sup>60</sup> "Kant's Two Solutions to the Free Rider Problem," *op. cit.* Note 57.

terest it is; and this is a direct consequence of the conflict between her attempt to particularize her situation on the one hand, and her attempt to rationally justify it by universalizing its principles on the other. Were all such agents in the state of nature to reason similarly, no Social Contract could be established because, as Kant observes, “actually there then would be no promising” at all (GMS Ak. 04:403.15) and hence no shared rules to obey.

So just as O’Neill’s contradiction in the will test tracks a criterion of maxim rationality rather than one of maxim morality, similarly her contradiction in conception test tracks a criterion of regulative coherence rather than one of justice. It does not show that false promising is always unjust. Rather, it shows that false promising cannot be an intelligible object of regulation, whether of self or of society. In order for the concept of a false promise to be meaningful, it must presuppose a shared linguistic convention in which the concept of a true promise denotes a widespread, motivationally effective social convention. Sabotaging this convention in turn sabotages the linguistic convention, and the performative concept of a promise loses its meaning. Hence the force of the failure of universalizability in this case is its demonstration that, *prior to and independent of information availability constraints on everyone’s goals, rules and policies and any interpersonal coordination problems that might result from them*, the very form of theoretical reason itself excludes such conjoint and simultaneous self-interested behavior among the members of a community. O’Neill’s contradiction in conception test, applied to the example of false promising, refutes the presuppositions of the game-theoretic account of the origins of interpersonal coordination on which Millgram’s critique relies. Surely this is achievement enough.

Recall that Millgram’s critique took its cue from the New Kantian line, that the CI-procedure rejects as morally unacceptable those actions that are “practically inconsistent,” i.e. that thwart “one’s stake in one’s own agency (529-533).” One’s stake in one’s own agency is of course not merely any

old self-interest, but perhaps the most profound interest in one's self one can take. Millgram's response was to show that, on the contrary, it is the CI-procedure itself that thwarts this most basic self-interest. He began with the *prima facie* palatable thought that, as self-interested agents, we each need the cooperation of other self-interested agents in order to achieve our ends, and must try to achieve them largely in ignorance of others' ends, policies and rules – thus requiring exceptions to those rules *as a matter of principle*. But Kant's analysis of the false promising case at passage (30) shows that such exceptions cannot be a matter of principle, not even New Kantian principle. Millgram's argument, that self-interest justifies breaking the rules when necessary to advance one's ends, is a simple variation on the Free Rider principle, that self-interest justifies breaking the rules for personal gain. It implies that self-interest takes rational priority over, and is conceptually prior to, following the rules; and that such rules are defeasible as a matter of principle by considerations of personal advantage. O'Neill's contradiction in conception test tells us, by contrast, that the Free Rider's reasoning – and so the reasoning on which Millgram's critique depends – is, when elevated to the status of principle, logically incoherent. It tells us that self-interest cannot be conceptually prior to following the rules, and therefore cannot provide principled guidelines for when they are to be followed and when they are to be broken.

Now Millgram concludes his discussion by chiding the New Kantians for failing to observe his test of reflexive application:

- (34) When you are working up your philosophical theory, always stop to check what happens when you apply that theory to itself. Sometimes that operation will not so much as make sense. Sometimes the result will be fast and reassuring. But sometimes, as we have just seen, it will not, and so the text of reflective application is not one that you can



afford to neglect. (550)

But the application of Millgram's own theory to itself is anything but "fast and reassuring." Part of that theory was a quite compelling argument that he made against the New Kantians, that exceptions cannot be built into the principle they violate, on pain of self-defeat, either specifically or generically, because this thwarted the universalization of the exempted action's maxim. Of course Millgram does not have to universalize his maxims, at least on his view. In that case, he is free to allow as many promising students as he likes to skip class and graduate anyway. But if he means to allow any such exception *on principle*, then the principle must be precisely that which we now see is the Free Rider principle. And this is the principle that O'Neill's contradiction in conception test reveals as logically incoherent. The test shows that if particular exceptions cannot be built into the particular rules they break, then *a fortiori* the generic concept of an exception cannot be built into the generic concept of a rule it breaks, either. It shows that if everyone reasoned as Millgram does, there would be no rules to break.

## APPENDIX

Table 1:

47 Formulations of the Moral Law in Kant's  
*Groundwork of the Metaphysic of Morals*

This table contains all formulations in the order in which they appear in the text, whether in the form of commands, imperatives, questions or assertions. The list aims to be exhaustive. Subsequent tables cull further categories from this list: (2) formulations in the imperative; (3) universal law formulations; (4) intentional-object universal law formulations; (5) manner-of-willing universal law formulations; and (6) miscellaneous universal law formulations.

No.	German	English	Citation GMS Ak. 04:
1	[...] [I]ch soll niemals anders verfahren als so, daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden.	I am never to proceed otherwise than such that I could also will that my maxim should become a universal law.	402.09-11
2	Kannst du auch wollen, daß deine Maxime ein allgemeines Gesetz werde?	Can you also will that your maxim become a universal law?	403.26-27
3	[S]o bleibt nichts als die Allgemeinheit eines Gesetzes überhaupt übrig, welchem die Maxime der Handlung gemäß sein soll, und welche Gemäßheit allein der Imperative eigentlich als notwendig vorstellt.	[N]othing remains to which the maxim of the action shall conform except the universality of a law in general; and which conformity alone the imperative actually represents as necessary.	421.03-06
4	[H]andle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde.	Act only in accordance with that maxim through which you can at the same time will that it become a universal law.	421.08-09
5	[H]andle so, als ob die Maxime deiner Handlung durch deinen Willen zu m allgemeinen Naturgesetze werden sollte.	So act as though the maxims of your action should become UNIVERSAL LAWS OF NATURE through your will.	421.21-23

6	Man muß <i>wollen können</i> , daß eine Maxime unserer Handlung ein allgemeines Gesetz werde [...].	One must <i>be able to will</i> that a maxim of our action become a universal law . . . .	424.02-03
7	[...] daß wir wirklich nicht wollen, es solle unsere Maxime ein allgemeines Gesetz werden, [...]	. . . that we do not really will such that our maxim should become a universal law, . . .	424.19-20
8	[D]er Mensch und überhaupt jedes vernünftige Wesen <i>existiert</i> als Zweck an sich selbst, <i>nicht bloß als Mittel</i> zum beliebigen Gebrauche für diesen oder jenen Willen, sondern muß in allen seinen sowohl auf sich selbst als auch auf anderen vernünftige Wesen gerichteten Handlungen jederzeit <i>zugleich als Zweck</i> betrachtet werden.	A human being and generally every rational being <i>exists</i> as an end in itself, not <i>merely as a means</i> for the arbitrary use of this or that will. Rather, in all its actions, whether directed towards itself or also towards other rational beings, it must always be considered <i>at the same time as an end</i> .	428.09-14
9	<i>Handle so, daß du die Menschheit, sowohl in deiner Person als in der Person eines jeden anderen, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchst.</i>	<i>So act, that you use humanity, in your own person as well as in the person of anyone else, always at the same time as an end, never merely as a means.</i>	429.12-14
10	[...] die Idee des Willens jedes vernünftigen Wesens als eines allgemein gesetzgebenden Willens.	. . . the idea of <i>the will of every rational being as a universal lawmaking will</i> .	431.20-22
11	[...] die Idee des Willens eines jeden vernünftigen Wesens als <i>allgemeingesetzgebenden Willens</i> .	. . . the idea of the will of each and every rational being <i>as a universal lawmaking will</i> .	432.03-04
12	[...] das <i>Prinzip</i> eines jeden menschlichen Willens als <i>eines durch alle seine Maximen allgemein gesetzgebenden Willens</i> [...].	. . . the principle of each and every human will as <i>a will that makes universal law through all of its maxims</i> . . . .	432.14-16
13	[...] daß es, eben um der Idee der allgemeinen Gesetzgebung willen, <i>sich auf kein Interesse gründet</i> und als unter allen möglichen Imperativen allein <i>unbedingt</i> sein kann [...]	. . . precisely for the sake of universal legislation, <i>it is not grounded on any interest</i> and among all possible imperatives alone can be <i>unconditioned</i> . . . .	432.17-20
14	[...] wenn es einen kategorischen Imperativ gibt (d.i. ein Gesetz für jeden Willen eines vernünftigen Wesens) [...]	. . . [i]f there is a categorical imperative (i.e. a law for every will of a rational being) . . . .	432.21-23
15	[...] so kann er nur gebieten, alles aus der Maxime seines Willens als eines solchen zu tun, der zugleich sich selbst als allgemein gesetzgebend zum Gegenstand haben könnte [...]	. . . it can only command to do everything from the maxims of its will as one that at the same time could have as its object itself as universally lawmaking . . . .	432.23-25

16	[...] daß er <i>nur seiner eigenen</i> und dennoch allgemeinen Gesetzgebung unterworfen sei [...]	... that [the human being] is subject <i>only to her own</i> and yet universal legislation . . . .	432.33-34
17	[...] daß er nur verbunden sei, seinem eigenen, dem Naturzwecke nach aber allgemein gesetzgebenden Willen gemäß zu handeln.	... that he is only bound to act according to natural ends, but in conformity with a universally lawmaking will.	432.34-36
18	Der Begriff eines jeden vernünftigen Wesens, das sich durch alle Maximen seines Willens als allgemein gesetzgebend betrachten muß, [...].	The concept of each and every rational being that must consider itself as universally lawmaking through all of the maxims of its will . . . .	433.15-17
19	[V]ernünftige Wesen stehen alle unter dem Gesetz, daß jedes derselben sich selbst und alles anderen <i>niemals bloß als Mittel</i> , sondern jederzeit <i>zugleich als Zweck an sich selbst</i> behandeln solle.	All rational beings stand under the law that each of them should treat itself and every other <i>never simply as means</i> , but rather always <i>at the same time as an end in itself</i> .	433.31-34
20	Das vernünftige Wesen muß sich jederzeit als gesetzgebend in einem durch Freiheit des Willens möglichen Reiche der Zwecke betrachten [...].	A rational being must always consider itself as lawmaking in a realm of ends possible through the freedom of the will . . . .	434.01-04
21	[...] keine Handlung nach einer anderen Maxime zu tun als so, daß es auch mit ihr bestehen könne, daß sie ein allgemeines Gesetz sei, [...]	... to perform no action according to any other maxim than such that it could be consistent with its being a universal law, . . . .	434.12-14
22	[...] und also nur so, daß der Wille durch seine Maxime sich selbst <i>zugleich als allgemein gesetzgebend betrachten könne</i> .	... and thus only such <i>that the will through its maxim could consider itself at the same time as universally lawmaking</i> .	434.14-16
23	[...] daß die Maximen so müssen gewählt werden, als ob sie wie allgemeine Naturgesetze gelten sollten [...].	... that the maxims must be so chosen as though they should be valid as universal laws of nature . . . .	436.19-21
24	[...] daß das vernünftige Wesen als Zweck seiner Natur nach, mithin als Zweck an sich selbst, jeder Maxime zur einschränkenden Bedingung aller bloß relativen und willkürlichen Zwecke dienen müsse [...].	... that a rational being in accordance with its nature as an end, and consequently as an end in itself, must serve every maxim as the limiting condition of all merely relative and arbitrary ends . . . .	436.23-26
25	[...] daß alle Maximen aus eigener Gesetzgebung zu einem möglichen Reiche der Zwecke, als einem Reiche der Natur, zusammenstimmen sollen.	... that all maxims out of one's own legislation are to harmonize in a possible realm of ends, as a realm of nature.	436.28-30
26	[H]andle nach der Maxime, die sich selbst <i>zugleich zum allgemeinen Gesetze machen kann</i> .	<i>Act according to the maxim that can make itself at the same time into a universal law.</i>	437.01-02

27	[H]andle jederzeit nach derjenigen Maxime, deren Allgemeinheit als Gesetzes du zugleich wollen kannst [...].	Always act according to that maxim whose universality as law you can at the same time will . . . .	437.12-14
28	<i>Handle nach Maximen, die sich selbst zugleich als allgemeine Naturgesetze zum Gegenstand haben können.</i>	Act according to maxims that can at the same time have as their object themselves as universal laws of nature	437.21-23
29	[H]andle in Beziehung auf ein jedes vernünftige Wesen (auf dich selbst und andere) so, daß es in deiner Maxime zugleich als Zweck an sich selbst gelte [...]	Act in relation to a rational being (yourself and others) such that in your maxim it would at the same time count as an end in itself . . . .	437.40-41
30	[H]andle nach einer Maxime, die ihre eigene allgemeine Gültigkeit für jedes vernünftige Wesen zugleich in sich enthält [...].	Act according to a maxim that at the same time contains within itself its own universal validity for every rational being . . . .	438.02-04
31	[...] daß ich meine Maxime im Gebrauche der Mittel zu jedem Zwecke auf die Bedingung ihrer Allgemeingültigkeit als eines Gesetzes für jedes Subjekt einschränken soll [...]	. . . that in their use as means to any ends, I am to limit my maxims to the condition of their universal validity as a law for every subject . . . .	438.04-07
32	[D]as Subjekt der Zwecke d.i. das vernünftige Wesen selbst muß niemals bloß als Mittel, sondern als oberste einschränkende Bedingung im Gebrauche aller Mittel, d.i. jederzeit zugleich als Zweck, allen Maximen der Handlungen zum Grunde gelegt werden.	[T]he subject of ends, i.e. the rational being itself, must be made the foundation of all maxims of actions, never simply as means, but rather as the supreme limiting condition of the use of all means, i.e. always at the same time as an end.	438.07-11
33	[...] daß jedes vernünftige Wesen als Zweck an sich selbst sich in Ansehung aller Gesetze, denen es nur immer unterworfen sein mag, zugleich als allgemein gesetzgebend müsse ansehen können [...]	. . . that every rational being, as an end in itself, must only be able to see itself as at the same time universally lawmaking, with regard to any laws whatever to which it may be subject . . . .	438.12-15
34	Demnach muß ein jedes vernünftige Wesen so handeln, als ob es durch seine Maximen jederzeit ein gesetzgebendes Glied im allgemeinen Reiche der Zwecke wäre.	Accordingly, each and every rational being must thus act as though it were, through its maxims, always a lawmaking member in a universal realm of ends.	438.25-27
35	Das Prinzip der Autonomie ist also: nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Willen zugleich als allgemeines Gesetz mit begriffen seien.	The principle of autonomy is thus: not to choose otherwise than so that the maxims of its choice are at the same time comprehended as universal law in the same willing.	440.20-23
36	Dagegen sagt der moralische, mithin kategorische Imperativ: ich soll so oder so handeln, ob ich gleich nichts anders wollte.	By contrast, the moral, and consequently categorical imperative says: I am to act thus and so, whether or not I wanted anything else.	441.12-14

37	[D]ie Tauglichkeit der Maxime eines jeden guten Willens, sich selbst zum allgemeinen Gesetze zu machen, ist selbst das alleinige Gesetz, das sich der Wille eines jeden vernünftigen Wesens selbst auferlegt [...].	The suitability of the maxims of every good will to make itself into universal law is itself the only law that the will of every rational being imposes on itself . . . .	444.35-38
38	[D]er Wille ist in allen Handlungen sich selbst ein Gesetz [...]	[T]he will itself is a law in all actions . . . .	447.04-05
39	[...] das Prinzip, nach keiner anderen Maxime zu handeln als die sich selbst auch als ein allgemeines Gesetz zum Gegenstande haben kann.	. . . the principle to act according to no other maxim, than that which can also have for an object itself as universal law.	447.05-08
40	[E]in schlechterdings guter Wille ist derjenige, dessen Maxime jederzeit sich selbst, als allgemeines Gesetz betrachtet, in sich enthalten kann [...].	[A]n absolutely good will is one whose maxims always can always contain itself considered as universal law . . . .	447.14-16
41	[...] daß die subjektiven Grundsätze der Handlungen, d.i. Maximen jederzeit so genommen werden müssen, daß sie auch objektiv, d.i. allgemein als Grundsätze gelten, mithin zu unserer eigenen allgemeinen Gesetzgebung dienen können.	. . . that the subjective principles of actions, i.e. maxims, must always be taken in such a way that they also count as objective, i.e. universal principles, and consequently can serve for our own universal legislation.	449.10-14
42	<i>Weil aber die Verstandeswelt</i> [...] also in Ansehung meines Willens [...] unmittelbar gesetzgebend ist und also auch als solche gedacht werden muß, so werde ich mich als Intelligenz, [...] folglich die Gesetze der Verstandeswelt für mich als Imperativen und die diesem Prinzip gemäßen Handlungen als Pflichten ansehen müssen.	<i>Because the world of understanding</i> . . . thus in regard to my will . . . is directly lawmaking and therefore also must be thought as such, thus I will have to regard myself as intelligence, and hence regard the laws of the world of understanding as imperatives for me and the actions according to this principle as obligations.	453.37, 39-40; 454.01, 05-08
43	[...] so zu handeln, daß das Prinzip der Handlungen der wesentlich Beschaffenheit einer Vernunftursache, d.i. der Bedingung der Allgemeingültigkeit der Maxime als eines Gesetzes gemäß sei.	. . . to act in such a way that the principle of actions is in accordance with the essential quality of a rational cause, i.e. with the condition of universal validity of the maxim as a law.	458.15-18
44	[...] macht den Begriff einer intelligibelen Welt [...] notwendig, [...] als bloß ihrer <i>formalen</i> Bedingung nach, d.i. der Allgemeinheit der Maxim des Willens als Gesetzes, mithin der Autonomie des letzteren [...].	. . . makes necessary the concept of an intelligible world, . . . as simply according to its <i>formal</i> condition, i.e. the universality of the maxim of the will as law, consequently the autonomy of the latter . . . .	458.32-33, 34, 35-36
45	[S]o ist die Erklärung, wie und warum uns die <i>Allgemeinheit der Maxime als Gesetzes</i> , mithin die Sittlichkeit interessiere, uns Menschen gänzlich unmöglich.	[T]hus the explanation of how and why the <i>universality of the maxim as law</i> , and consequently morality, would interest us, is entirely impossible for us human beings.	460.26-29

46	<i>[W]ie das bloße Prinzip der Allgemeingültigkeit aller ihrer Maximen als Gesetze [...] für sich selbst eine Triebfeder abgeben [...].</i>	<i>[H]ow the simple principle of the universal validity of all of [human reason's] maxims as law ... by itself yields an incentive ... .</i>	461.31-32, 35-36
47	Von der reinen Vernunft, [...] bleibt nach Absonderung aller Materie, d.i. Erkenntniß der Objecte mir nichts als die Form übrig, nämlich das praktische Gesetz der Allgemeingültigkeit der Maximen [...].	After culling all matter from pure reason, i.e. cognition of objects, nothing remains to me other than its form, namely the practical law of the universal validity of maxims ... .	462.15-16, 16-17, 17-19

Table 2:

These two tables contain all formulations that use forms of *sollen* to modify *wollen*- or *handeln*-type verbs.

### 2.1. Four Strict Imperative (*sollen*) Formulations of the Moral Law in Kant's *Groundwork of the Metaphysic of Morals*

No.	German	English	Citation GMS Ak. 04:
1	[I]ch soll niemals anders verfahren als so, daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden.	I am never to proceed otherwise than such that I could also will that my maxim should become a universal law.	402.09-11
25	[...] daß alle Maximen aus eigener Gesetzgebung zu einem möglichen Reiche der Zwecke, als einem Reiche der Natur, zusammenstimmen sollen.	... that all maxims out of one's own legislation are to harmonize in a possible realm of ends, as a realm of nature.	436.28-30
31	[...] daß ich meine Maxime im Gebrauche der Mittel zu jedem Zwecke auf die Bedingung ihrer Allgemeingültigkeit als eines Gesetzes für jedes Subjekt einschränken soll [...]	... that in their use as means to any ends, I am to limit my maxims to the condition of their universal validity as a law for every subject ... .	438.04-07
36	Dagegen sagt der moralische, mithin kategorische Imperativ: ich soll so oder so handeln, ob ich gleich nichts anders wollte.	By contrast, the moral, and consequently categorical imperative says: I am to act thus and so, whether or not I wanted anything else.	441.12-14

2.2. Two Subjunctive I Imperative (*solle, sollten*)  
Formulations of the Moral Law in Kant's *Groundwork of  
the Metaphysic of Morals*

No.	German	English	Citation GMS Ak. 04:
7	[...] daß wir wirklich nicht wollen, es solle unsere Maxime ein allgemeines Gesetz werden, [...]	... that we do not really will such that our maxim should become a universal law, . . .	424.19-20
19	[V]ernünftige Wesen stehen alle unter dem Gesetz, daß jedes derselben sich selbst und alles anderen <i>niemals bloß als Mittel</i> , sondern jederzeit <i>zugleich als Zweck an sich selbst</i> behandeln solle.	All rational beings stand under the law that each of them should treat itself and every other <i>never simply as means</i> , but rather <i>always at the same time as an end in itself</i> .	433.31-34



Table 3:

38 Universal Law Formulations of the Moral Law in Kant's  
*Groundwork of the Metaphysic of Morals*

This table contains all formulations that characterize the moral law as a universal law, whether embedded in passages discussing universalizability, autonomy, ends in themselves, the kingdom of ends, or other topics.

No.	German	English	Citation GMS Ak. 04:
1	[I]ch soll niemals anders verfahren als so, <i>daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden.</i>	I am never to proceed otherwise than such <i>that I could also will that my maxim should become a universal law.</i>	402.09-11
2	Kannst du auch wollen, daß deine Maxime ein allgemeines Gesetz werde?	Can you also will, that your maxim become a universal law?	403.26-27
3	[S]o bleibt nichts als die Allgemeinheit eines Gesetzes überhaupt übrig, welchem die Maxime der Handlung gemäß sein soll, und welche Gemäßheit allein der Imperative eigentlich als notwendig vorstellt.	[N]othing remains to which the maxim of the action shall conform except the universality of a law in general; and which conformity alone the imperative actually represents as necessary.	421.03-06
4	[H]andle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde.	<i>Act only in accordance with that maxim through which you can at the same time will that it become a universal law.</i>	421.08-09
5	[H]andle so, als ob die Maxime deiner Handlung durch deinen Willen zu <i>m</i> allgemeinen Naturgesetze werden sollte.	<i>So act as though the maxims of your action should become UNIVERSAL LAWS OF NATURE through your will.</i>	421.21-23
6	Man muß <i>wollen können</i> , daß eine Maxime unserer Handlung ein allgemeines Gesetz werde [...].	One must <i>be able to will</i> that a maxim of our action become a universal law . . . .	424.02-03
7	[...] daß wir wirklich nicht wollen, es solle unsere Maxime ein allgemeines Gesetz werden, [...]	. . . that we do not really will such that our maxim should become a universal law, . . .	424.19-20

Adrian M. S. Piper, Kant's Self-Legislation Procedure  
Reconsidered, KSO 2012: 203—277

Posted October 20, 2012 [www.kantstudiesonline.net](http://www.kantstudiesonline.net)  
© 2012 Adrian M. S. Piper & Kant Studies Online Ltd.

10	[...] die Idee <i>des Willens jedes vernünftigen Wesens als eines allgemein gesetzgebenden Willens</i> .	... the idea of <i>the will of every rational being as a universal lawmaking will</i> .	431.20-22
11	[...] die Idee des Willens eines jeden vernünftigen Wesens <i>als allgemeingesetzgebenden Willens</i> .	... the idea of the will of each and every rational being <i>as a universal lawmaking will</i> .	432.03-04
12	[...] das <i>Prinzip</i> eines jeden menschlichen Willens <i>als eines durch alle seine Maximen allgemein gesetzgebenden Willens</i> [...].	... the principle of each and every human will <i>as a will that makes universal law through all of its maxims</i> . . . .	432.14-16
13	[...] daß es, eben um der Idee der allgemeinen Gesetzgebung willen, <i>sich auf kein Interesse gründet</i> und als unter allen möglichen Imperativen allein <i>unbedingt</i> sein kann [...]	... precisely for the sake of universal legislation, <i>it is not grounded on any interest</i> and among all possible imperatives alone can be <i>unconditioned</i> . . . .	432.17-20
14	[...] wenn es einen kategorischen Imperativ gibt (d.i. ein Gesetz für jeden Willen eines vernünftigen Wesens) [...]	... [i]f there is a categorical imperative (i.e. a law for every will of a rational being) . . . .	432.21-23
15	[...] so kann er nur gebieten, alles aus der Maxime seines Willens als eines solchen zu tun, der zugleich sich selbst als allgemein gesetzgebend zum Gegenstand haben könnte [...]	... it can only command to do everything from the maxims of its will as one that at the same time could have as its object itself as universally lawmaking . . . .	432.23-25
16	[...] daß er <i>nur seiner eigenen</i> und dennoch allgemeinen Gesetzgebung unterworfen sei [...]	... that [the human being] is subject <i>only to her own</i> and yet universal legislation . . . .	432.33-34
17	[...] daß er nur verbunden sei, seinem eigenen, dem Naturzwecke nach aber allgemein gesetzgebenden Willen gemäß zu handeln.	... that he is only bound to act according to natural ends, but in conformity with a universally lawmaking will.	432.34-36
18	Der Begriff eines jeden vernünftigen Wesens, das sich durch alle Maximen seines Willens als allgemein gesetzgebend betrachten muß, [...].	The concept of each and every rational being that must consider itself as universally lawmaking through all of the maxims of its will . . . .	433.15-17
19	[V]ernünftige Wesen stehen alle unter dem Gesetz, daß jedes derselben sich selbst und alles anderen <i>niemals bloß als Mittel</i> , sondern jederzeit <i>zugleich als Zweck an sich selbst</i> behandeln solle.	All rational beings stand under the law that each of them should treat itself and every other <i>never simply as means</i> , but rather always <i>at the same time as an end in itself</i> .	433.31-34
21	[...] keine Handlung nach einer anderen Maxime zu tun als so, daß es auch mit ihr bestehen könne, daß sie ein allgemeines Gesetz sei, [...]	... to perform no action according to any other maxim than such that it could be consistent with its being a universal law, . . . .	434.12-14

22	[...] und also nur so, daß der Wille durch seine Maxime sich selbst zugleich als allgemein gesetzgebend betrachten könne.	... and thus only such that the will through its maxim could consider itself at the same time as universally lawmaking.	434.14-16
23	[...] daß die Maximen so müssen gewählt werden, als ob sie wie allgemeine Naturgesetze gelten sollten [...].	... that the maxims must be so chosen as though they should be valid as universal laws of nature . . . .	436.19-21
26	[H]andle nach der Maxime, die sich selbst zugleich zum allgemeinen Gesetze machen kann.	Act according to the maxim that can make itself at the same time into a universal law.	437.01-02
27	[H]andle jederzeit nach derjenigen Maxime, deren Allgemeinheit als Gesetzes du zugleich wollen kannst [...].	Always act according to that maxim whose universality as law you can at the same time will . . . .	437.12-14
28	Handle nach Maximen, die sich selbst zugleich als allgemeine Naturgesetze zum Gegenstand haben können.	Act according to maxims that can at the same time have as their object themselves as universal laws of nature	437.21-23
30	[H]andle nach einer Maxime, die ihre eigene allgemeine Gültigkeit für jedes vernünftige Wesen zugleich in sich enthält [...].	Act according to a maxim that at the same time contains within itself its own universal validity for every rational being . . . .	438.02-04
31	[...] daß ich meine Maxime im Gebrauche der Mittel zu jedem Zwecke auf die Bedingung ihrer Allgemeingültigkeit als eines Gesetzes für jedes Subjekt einschränken soll [...]	... that in their use as means to any ends, I am to limit my maxims to the condition of their universal validity as a law for every subject . . . .	438.04-07
33	[...] daß jedes vernünftige Wesen als Zweck an sich selbst sich in Ansehung aller Gesetze, denen es nur immer unterworfen sein mag, zugleich als allgemein gesetzgebend müsse ansehen können [...]	... that every rational being, as an end in itself, must only be able to see itself as at the same time universally lawmaking, with regard to any laws whatever to which it may be subject . . . .	438.12-15
34	Demnach muß ein jedes vernünftige Wesen so handeln, als ob es durch seine Maximen jederzeit ein gesetzgebendes Glied im allgemeinen Reiche der Zwecke wäre.	Accordingly, each and every rational being must thus act as though it were, through its maxims, always a lawmaking member in a universal realm of ends.	438.25-27
35	Das Prinzip der Autonomie ist also: nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Wollen zugleich als allgemeines Gesetz mit begriffen seien.	The principle of autonomy is thus: not to choose otherwise than so that the maxims of its choice are at the same time comprehended as universal law in the same willing.	440.20-23
37	[D]ie Tauglichkeit der Maxime eines jeden guten Willens, sich selbst zum allgemeinen Gesetze zu machen, ist selbst das alleinige Gesetz, das sich der Wille eines jeden vernünftigen Wesens selbst auferlegt [...].	The suitability of the maxims of every good will to make itself into universal law is itself the only law that the will of every rational being imposes on itself . . . .	444.35-38

38	[D]er Wille ist in allen Handlungen sich selbst ein Gesetz [...]	[T]he will itself is a law in all actions . . . .	447.04-05
39	[...] das Prinzip, nach keiner anderen Maxime zu handeln als die sich selbst auch als ein allgemeines Gesetz zum Gegenstande haben kann.	. . . the principle to act according to no other maxim, than that which can also have for an object itself as universal law.	447.05-08
40	[E]in schlechterdings guter Wille ist derjenige, dessen Maxime jederzeit sich selbst, als allgemeines Gesetz betrachtet, in sich enthalten kann [...].	[A]n absolutely good will is one whose maxims always can always contain itself considered as universal law . . . .	447.14-16
41	[...] daß die subjektiven Grundsätze der Handlungen, d.i. Maximen jederzeit so genommen werden müssen, daß sie auch objektiv, d.i. allgemein als Grundsätze gelten, mithin zu unserer eigenen allgemeinen Gesetzgebung dienen können.	. . . that the subjective principles of actions, i.e. maxims, must always be taken in such a way that they also count as objective, i.e. universal principles, and consequently can serve for our own universal legislation.	449.10-14
43	[...] so zu handeln, daß das Prinzip der Handlungen der wesentlich Beschaffenheit einer Vernunftursache, d.i. der Bedingung der Allgemeingültigkeit der Maxime als eines Gesetzes gemäß sei.	. . . to act in such a way that the principle of actions is in accordance with the essential quality of a rational cause, i.e. with the condition of universal validity of the maxim as a law.	458.15-18
44	[...] macht den Begriff einer intelligibelen Welt [...] notwendig, [...] als bloß ihrer <i>formalen</i> Bedingung nach, d.i. der Allgemeinheit der Maxim des Willens als Gesetzes, mithin der Autonomie des letzteren [...].	. . . makes necessary the concept of an intelligible world, . . . as simply according to its <i>formal</i> condition, i.e. the universality of the maxim of the will as law, consequently the autonomy of the latter . . . .	458.32-33, 34, 35-36
45	[S]o ist die Erklärung, wie und warum uns die <i>Allgemeinheit der Maxime als Gesetzes</i> , mithin die Sittlichkeit interessiere, uns Menschen gänzlich unmöglich.	[T]hus the explanation of how and why the <i>universality of the maxim as law</i> , and consequently morality, would interest us, is entirely impossible for us human beings.	460.26-29
46	<i>[W]ie das bloße Prinzip der Allgemeingültigkeit aller ihrer Maximen als Gesetze</i> [...] für sich selbst eine Triebfeder abgeben [...].	<i>[H]ow the simple principle of the universal validity of all of [human reason's] maxims as law</i> . . . by itself yields an incentive . . . .	461.31-32, 35-36
47	Von der reinen Vernunft, [...] bleibt nach Absonderung aller Materie, d.i. Erkenntniß der Objecte mir nichts als die Form übrig, nämlich das praktische Gesetz der Allgemeingültigkeit der Maximen [...].	After culling all matter from pure reason, i.e. cognition of objects, nothing remains to me other than its form, namely the practical law of the universal validity of maxims . . . .	462.15-16, 16-17, 17-19

Table 4:

Seven Intentional-Object Universal Law Formulations of  
the Moral Law in Kant's *Groundwork of the  
Metaphysic of Morals*

This table contains all universal law formulations that conceive rational deliberation as stipulating a particular intentional object of the will, rather than as a manner of willing.

No.	German	English	Citation GMS Ak. 04:
1	[I]ch soll niemals anders verfahren als so, <i>daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden.</i>	I am never to proceed otherwise than such <i>that I could also will that my maxim should become a universal law.</i>	402.09-11
2	Kannst du auch wollen, daß deine Maxime ein allgemeines Gesetz werde?	Can you also will that your maxim become a universal law?	403.26-27
4	<i>[H]andle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde.</i>	<i>Act only in accordance with that maxim through which you can at the same time will that it become a universal law.</i>	421.08-09
6	Man muß <i>wollen können</i> , daß eine Maxime unserer Handlung ein allgemeines Gesetz werde [...].	One must <i>be able to will</i> that a maxim of our action become a universal law . . . .	424.02-03
15	[...] so kann er nur gebieten, alles aus der Maxime seines Willens als eines solchen zu tun, der zugleich sich selbst als allgemein gesetzgebend zum Gegenstand haben könnte [...]	. . . it can only command to do everything from the maxims of its will as one that at the same time could have as its object itself as universally lawmaking . . . .	432.23-25
28	<i>Handle nach Maximen, die sich selbst zugleich als allgemeine Naturgesetze zum Gegenstand haben können.</i>	Act according to maxims that can at the same time have as their object themselves as universal laws of nature	437.21-23
39	[...] das Prinzip, nach keiner anderen Maxime zu handeln als die sich selbst auch als ein allgemeines Gesetz zum Gegenstande haben kann.	. . . the principle to act according to no other maxim, than that which can also have for an object itself as universal law.	447.05-08

Table 5:

24 Manner-of-Willing Universal Law Formulations of the  
Moral Law in Kant's *Groundwork of the  
Metaphysic of Morals*

This table contains all universal law formulations that conceive rational  
deliberation as a manner of willing rather than as stipulating a  
particular intentional object of the will.

No.	German	English	Citation GMS Ak. 04:
3	[S]o bleibt nichts als die Allgemeinheit eines Gesetzes überhaupt übrig, welchem die Maxime der Handlung gemäß sein soll, und welche Gemäßheit allein der Imperative eigentlich als notwendig vorstellt.	[N]othing remains to which the maxim of the action shall conform except the universality of a law in general; and which conformity alone the imperative actually represents as necessary.	421.03-06
5	<i>[H]andle so, als ob die Maxime deiner Handlung durch deinen Willen zu m allgemeinen Naturgesetze werden sollte.</i>	<i>So act as though the maxims of your action should become UNIVERSAL LAWS OF NATURE through your will.</i>	421.21-23
7	[...] daß wir wirklich nicht wollen, es solle unsere Maxime ein allgemeines Gesetz werden, [...]	... that we do not really will such that our maxim should become a universal law, ...	424.19-20
10	[...] die Idee des Willens jedes vernünftigen Wesens als eines allgemein gesetzgebenden Willens.	... the idea of the will of every rational being as a universal lawmaking will.	431.20-22
11	[...] die Idee des Willens eines jeden vernünftigen Wesens als allgemeingesetzgebenden Willens.	... the idea of the will of each and every rational being as a universal lawmaking will.	432.03-05
12	[...] das Prinzip eines jeden menschlichen Willens als eines durch alle seine Maximen allgemein gesetzgebenden Willens [...].	... the principle of each and every human will as a will that makes universal law through all of its maxims ...	432.14-16
18	Der Begriff eines jeden vernünftigen Wesens, das sich durch alle Maximen seines Willens als allgemein gesetzgebend betrachten muß, [...].	The concept of each and every rational being that must consider itself as universally lawmaking through all of the maxims of its will ...	433.15-17

19	[V]ernünftige Wesen stehen alle unter dem Gesetz, daß jedes derselben sich selbst und alles anderen <i>niemals bloß als Mittel</i> , sondern jederzeit <i>zugleich als Zweck an sich selbst</i> behandeln solle.	All rational beings stand under the law that each of them should treat itself and every other <i>never simply as means</i> , but rather always <i>at the same time as an end in itself</i> .	433.31-34
21	[...] keine Handlung nach einer anderen Maxime zu tun als so, daß es auch mit ihr bestehen könne, daß sie ein allgemeines Gesetz sei, [...]	... to perform no action according to any other maxim than such that it could be consistent with its being a universal law, ...	434.12-14
22	[...] und also nur so, daß der Wille durch seine Maxime sich selbst <i>zugleich als allgemein gesetzgebend betrachten könne</i> .	... and thus only such <i>that the will through its maxim could consider itself at the same time as universally lawmaking</i> .	434.14-16
23	[...] daß die Maximen so müssen gewählt werden, als ob sie wie allgemeine Naturgesetze gelten sollten [...].	... that the maxims must be so chosen as though they should be valid as universal laws of nature ...	436.19-21
26	[H]andle nach der Maxime, die sich selbst <i>zugleich zum allgemeinen Gesetze machen kann</i> .	<i>Act according to the maxim that can make itself at the same time into a universal law</i> .	437.01-02
27	[H]andle jederzeit nach derjenigen Maxime, deren Allgemeinheit als Gesetzes du zugleich wollen kannst [...].	Always act according to that maxim whose universality as law you can at the same time will ...	437.12-14
31	[...] daß ich meine Maxime im Gebrauche der Mittel zu jedem Zwecke auf die Bedingung ihrer Allgemeingültigkeit als eines Gesetzes für jedes Subjekt einschränken soll [...]	... that in their use as means to any ends, I am to limit my maxims to the condition of their universal validity as a law for every subject ...	438.04-07
33	[...] daß jedes vernünftige Wesen als Zweck an sich selbst sich in Ansehung aller Gesetze, denen es nur immer unterworfen sein mag, zugleich als allgemein gesetzgebend müsse ansehen können [...]	... that every rational being, as an end in itself, must only be able to see itself as at the same time universally lawmaking, with regard to any laws whatever to which it may be subject ...	438.12-15
34	Demnach muß ein jedes vernünftige Wesen so handeln, als ob es durch seine Maximen jederzeit ein gesetzgebendes Glied im allgemeinen Reiche der Zwecke wäre.	Accordingly, each and every rational being must thus act as though it were, through its maxims, always a lawmaking member in a universal realm of ends.	438.25-27
35	Das Prinzip der Autonomie ist also: nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Willen zugleich als allgemeines Gesetz mit begriffen seien.	The principle of autonomy is thus: not to choose otherwise than so that the maxims of its choice are at the same time comprehended as universal law in the same willing.	440.20-23

37	[D]ie Tauglichkeit der Maxime eines jeden guten Willens, sich selbst zum allgemeinen Gesetze zu machen, ist selbst das alleinige Gesetz, das sich der Wille eines jeden vernünftigen Wesens selbst auferlegt [...].	The suitability of the maxims of every good will to make itself into universal law is itself the only law that the will of every rational being imposes on itself . . . .	444.35-37
40	[E]in schlechterdings guter Wille ist derjenige, dessen Maxime jederzeit sich selbst, als allgemeines Gesetz betrachtet, in sich enthalten kann [...].	[A]n absolutely good will is one whose maxims always can always contain itself considered as universal law . . . .	447.14-16
41	[...] daß die subjektiven Grundsätze der Handlungen, d.i. Maximen jederzeit so genommen werden müssen, daß sie auch objektiv, d.i. allgemein als Grundsätze gelten, mithin zu unserer eigenen allgemeinen Gesetzgebung dienen können.	. . . that the subjective principles of actions, i.e. maxims, must always be taken in such a way that they also count as objective, i.e. universal principles, and consequently can serve for our own universal legislation.	449.10-14
43	[...] so zu handeln, daß das Prinzip der Handlungen der wesentlich Beschaffenheit einer Vernunftursache, d.i. der Bedingung der Allgemeingültigkeit der Maxime als eines Gesetzes gemäß sei.	. . . to act in such a way that the principle of actions is in accordance with the essential quality of a rational cause, i.e. with the condition of universal validity of the maxim as a law.	458.15-18
44	[...] macht den Begriff einer intelligibelen Welt [...] notwendig, [...] als bloß ihrer <i>formalen</i> Bedingung nach, d.i. der Allgemeinheit der Maxim des Willens als Gesetzes, mithin der Autonomie des letzteren [...].	. . . makes necessary the concept of an intelligible world, . . . as simply according to its <i>formal</i> condition, i.e. the universality of the maxim of the will as law, consequently the autonomy of the latter . . . .	458.32-33, 34, 35-36
46	<i>[W]ie das bloße Prinzip der Allgemeingültigkeit aller ihrer Maximen als Gesetze</i> [...] für sich selbst eine Triebfeder abgeben [...].	<i>[H]ow the simple principle of the universal validity of all of [human reason's] maxims as law . . . by itself yields an incentive . . . .</i>	461.31-32, 35-36
47	Von der reinen Vernunft, [...] bleibt nach Absonderung aller Materie, d.i. Erkenntniß der Objecte mir nichts als die Form übrig, nämlich das praktische Gesetz der Allgemeingültigkeit der Maximen [...].	After culling all matter from pure reason, i.e. cognition of objects, nothing remains to me other than its form, namely the practical law of the universal validity of maxims . . . .	462.15-16, 16-17, 17-19



Table 6:

Seven Other Universal Law Formulations of the Moral Law  
in Kant's *Groundwork of the Metaphysic of Morals*

This table contains all universal law formulations not otherwise sorted. All are in the indicative mood.

No.	German	English	Citation GMS Ak. 04:
13	[...] daß es, eben um der Idee der allgemeinen Gesetzgebung willen, <i>sich auf kein Interesse gründet</i> und als unter allen möglichen Imperativen allein <i>unbedingt</i> sein kann [...]	... precisely for the sake of universal legislation, <i>it is not grounded on any interest</i> and among all possible imperatives alone can be <i>unconditioned</i> . . . .	432.17-20
14	[...] wenn es einen kategorischen Imperativ gibt (d.i. ein Gesetz für jeden Willen eines vernünftigen Wesens) [...]	... [i]f there is a categorical imperative (i.e. a law for every will of a rational being) . . . .	432.21-23
16	[...] daß er <i>nur seiner eigenen</i> und dennoch allgemeinen Gesetzgebung unterworfen sei [...]	... that [the human being] is subject <i>only to her own</i> and yet universal legislation . . . .	432.33-34
17	[...] daß er nur verbunden sei, seinem eigenen, dem Naturzwecke nach aber allgemein gesetzgebenden Willen gemäß zu handeln.	... that he is only bound to act according to natural ends, but in conformity with a universally lawmaking will.	432.34-36
30	[H]andle nach einer Maxime, die ihre eigene allgemeine Gültigkeit für jedes vernünftige Wesen zugleich in sich enthält [...].	Act according to a maxim that at the same time contains within itself its own universal validity for every rational being . . . .	438.02-04
38	[D]er Wille ist in allen Handlungen sich selbst ein Gesetz [...]	[T]he will itself is a law in all actions . . . .	447.04-05
45	[S]o ist die Erklärung, wie und warum uns die <i>Allgemeinheit der Maxime als Gesetzes</i> , mithin die Sittlichkeit interessiere, uns Menschen gänzlich unmöglich.	[T]hus the explanation of how and why the <i>universality of the maxim as law</i> , and consequently morality, would interest us, is entirely impossible for us human beings.	460.26-29

## References

- Förster, Eckart, Ed., *Kant's Transcendental Deductions: The Three 'Critiques' and the 'Opus Postumum'* (Stanford: Stanford University Press, 1989)
- Gibbard, Allan, "Utilitarianisms and Coordinations" (Ph.D. diss., Harvard University, 1971)
- Hintikka, Jaakko, *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic* (New York: Oxford University Press, 1973 and 2002)
- Hobbes, Thomas, *Leviathan*, Ed. Michael Oakeshott (New York: Macmillan/Collier Books, 1977)
- Kant, Immanuel, *Kritik der reinen Vernunft*, Herausg. Raymund Schmidt (Hamburg: Felix Meiner Verlag, 1976)
- \_\_\_\_\_, *Grundlegung zur Metaphysik der Sitten*, Herausg. Karl Vorländer (Hamburg: Felix Meiner Verlag, 1965)
- \_\_\_\_\_, *Groundwork of the Metaphysic of Morals: A German-English Edition*, trans. and ed. Mary Gregor and Jens Timmermann (Cambridge, UK: Cambridge University Press, 2011)
- Kitcher, Patricia, "What is a Maxim?" *Philosophical Topics* 31, 1 & 2 (April 2005)
- Lewis, David, *Convention: A Philosophical Study* (Cambridge, Mass.: Harvard University Press, 1969)
- Mill, John Stuart, *A System of Logic, Ratiocinative and Inductive, Volume I* (London: John W. Parker, 1843)
- Millgram, Elijah, "Does the Categorical Imperative Give Rise to a Contradiction in the Will?" *The Philosophical Review* 112, 4 (October 2003), 525-560.
- Neiman, Susan, *The Unity of Reason: Rereading Kant* (New York: Oxford University Press, 1994)
- Nell (née O'Neill), Onora, *Acting on Principle: An Essay on Kantian Ethics* (New York: Columbia University Press, 1975)
- O'Neill, Onora, *Constructions of Reason: Explorations of Kant's Practical Philosophy* (Cambridge: Cambridge University Press, 1989)
- Piper, Adrian M. S., "Moral Theory and Moral Alienation," *The Journal of Philosophy* LXXXIV, 2 (February 1987), 102-118
- \_\_\_\_\_, "Kant on the Objectivity of the Moral Law," in Andrews Reath, Christine M. Korsgaard and Barbara Herman, Eds. *Reclaiming the History of Ethics: Essays for John Rawls*, (New York: Cambridge University Press, 1997), 240-269; also at [http://adrianpiper.com/docs/WebsiteKantOnMorObjRwlsFstschrft\(1994\).pdf](http://adrianpiper.com/docs/WebsiteKantOnMorObjRwlsFstschrft(1994).pdf).

- \_\_\_\_ *Rationality and the Structure of the Self, Volume I: The Humean Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20and%20the%20Structure%20of%20the%20Self%20Volume%20I-%20The%20Humean%20Conception.pdf>, APRA Foundation Berlin, 2008)
- \_\_\_\_ *Rationality and the Structure of the Self, Volume II: A Kantian Conception* (Berlin: <http://adrianpiper.com/rss/docs/Rationality%20and%20the%20Structure%20of%20the%20Self%20Volume%20II-%20A%20Kantian%20Conception.pdf>, APRA Foundation Berlin, 2008).
- \_\_\_\_ “Kant’s Two Solutions to the Free Rider Problem,” *The Kant Yearbook 4/2012: Kant and Contemporary Moral Philosophy*, 113-142.
- \_\_\_\_ “Practical Action: First Critique Foundations,” in *Kant und die Philosophie in Weltbürgerlicher Absicht: Akten des XI. Internationalen Kant-Kongresses*, Herausg. Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca, und Margit Ruffing (Berlin: Walter de Gruyter, forthcoming)
- Rawls, John, *A Theory of Justice* (Cambridge, Mass.: Harvard University, 1971)
- \_\_\_\_ “Themes in Kant’s Moral Philosophy,” in Eckart Förster, Ed., *Kant’s Transcendental Deductions: The Three ‘Critiques’ and the ‘Opus Postumum’* (Stanford: Stanford University Press, 1989)
- Ullman-Margalit, Edna, *The Emergence of Norms* (Oxford: Clarendon Press, 1977).