Philosophical Books Vol. 49 No. 2 April 2008 pp. 125–137

AUTONOMY, TAKING ONE'S CHOICES TO BE GOOD, AND PRACTICAL LAW: REPLIES TO CRITICS

ANDREWS REATH
The University of California, Riverside

I

Several essays in my book try to understand what Kant means by autonomy and how that concept figures in his foundational project. Since political and juridical terminology runs through much of Kant's moral theory, I suggest that he models autonomy on a notion of sovereignty. I interpret autonomy not as a motivational capacity, but as the sovereignty of the will over itself: the rational will is not bound to any outside authority and it has the capacity to give universal law. I shall begin with an overview of some of these themes.

Since Samuel Kerstein asks whether I am interpreting or defending Kant, let me first say something about my approach. My guiding aim is to reconstruct the underlying ideas and arguments so as to make them philosophically successful, where 'reconstruction' permits pushing an idea farther in a certain direction than Kant did. My approach is interpretive, since I accept the constraint that the ideas that I develop should have a firm basis in the texts and the underlying arguments. Without claiming that there is a uniquely right reading, I want to get at the deep structure of concepts and arguments that animate the movements of thought in Kant's texts. But I am also engaged in defending Kant, in that I am trying, as far as I can, to make the arguments work out. Thus a second constraint is that the ideas and arguments attributed to Kant turn out to be philosophically successful. These two constraints are not always compatible, and that is one reason why Kant is hard work (and sometimes frustrating).

II

To provide background for my responses, I begin with some remarks about the role of autonomy in the overall trajectory of the *Groundwork* that may clarify the senses in which Kant thinks that moral requirements are laws that the will gives to itself. (Here I am more concerned to lay out an overall picture of what goes on in the *Groundwork* than to defend it.) Kant thinks that it is part of common moral thought that moral requirements apply with absolute necessity and that we take them to be categorical imperatives. Once this feature of common moral thought is brought to our attention, the question that sets the problematic of the *Groundwork* naturally arises: how could there be normative requirements with this kind of necessity? *Groundwork* III attempts to ground the authority of moral requirements in a dual necessity. Greatly simplified, it argues that the Categorical Imperative [CI] is the constitutive or internal principle of a necessary self-conception. First, the CI is the constitutive principle of free

Philosophical Books Vol. 49 No. 2

volition. That is to say, it is the principle that captures the general form of free volition and is the guiding principle that one must follow to exercise one's free agency.[1] It is in that sense the 'formal principle' of free volition.[2] Second, we necessarily think of ourselves as free and identify with our capacities for free volition; and (Kant argues) independent support for ascribing ourselves free agency comes from features of theoretical reason. If the CI is the constitutive principle, or general form, of free volition, then we cannot coherently reject the authority of morality and continue to think of ourselves as free agents. (This analytic connection between the CI and free volition is the first element of normative necessity.) And we necessarily think of ourselves as free when we adopt the practical perspective. (This synthetic a priori feature of the practical perspective is the second element of normative necessity.)

In order to make these very large claims, Kant works through a series of reformulations of the CI in *Groundwork* II, one aim of which is to connect the CI to the nature of free volition. Let me quickly note some of the high points.

The Formula of Universal Law [FUL] states the basic principle of morality. It may be understood as a general principle that can be used to derive some substantive requirements on conduct (at a high level of generality) by specifying the form that a principle must have to serve as a practical law. With the FUL and the Formula of Humanity in hand, and still operating in the analytical mode, Kant turns to the authority of the CI: what grounds the authority of moral requirements for agents subject to them? At *Groundwork* 431, he makes the striking claim that "the will is not merely subject to the [moral] law but subject to it in such a way that it must be regarded as also giving the law to itself . . . [*so unterworfen das er auch als selbstgesetzgebend . . . angeshen werden muß*]". Kant's claim, which I call the 'Sovereignty Thesis', is that agents subject to moral principles must be regarded as the legislators from whom they receive their authority. It is a conceptual claim that follows from the concept of an unconditional practical requirement or practical law, and it leads to a new formula of the CI, the Formula of Autonomy [FA], which is equivalent to the FUL. That is, the Sovereignty Thesis implies that in acting from maxims that one can will as universal law, one acts in such a way that one can regard oneself as giving universal law through one's will. The general idea behind the Sovereignty Thesis is that since a categorical imperative applies unconditionally, its authority cannot depend on any contingent interest in those subject to it; instead it comes from the fact that agents subject to the principle are (in some sense) its legislator—or to make the same point in a different way, its authority comes from the fact that the principle arises from the agent's will. The idea

1. "An absolutely good will, whose principle must be a categorical imperative, will therefore . . . contain merely the *form of volition* as such and indeed as autonomy . . ." (*Groundwork of the Metaphysics of Morals*, in Mary J. Gregor (tr.) and Allen Wood (ed.), *Kant: Practical Philosophy*, (Cambridge University Press, 1996), 4: 444.) Citations to the *Groundwork*, abbreviated as *G*, give the Berlin Academy pagination.
2. For Kant a 'formal principle' is a principle that is constitutive of some rational activity. It gives the 'form of the activity' in the sense of providing the guiding norm that one must follow if one is to engage in that activity. I discuss this idea further in 'Formal Principles and the Form of a Law', to appear in A. Reath and J. Timmermann (eds.), *A Critical Guide to Kant's 'Critique of Practical Reason'* (Cambridge University Press, forthcoming).

that moral subjects are law-giving introduces the concept of autonomy, because autonomy is the capacity of the will to give itself law.

The Sovereignty Thesis leads to a further striking idea that sets out the target for Kant's foundational exercise: that a satisfactory moral theory must represent moral requirements as in some sense self-legislated. If moral requirements apply unconditionally, and it is a conceptual truth that agents subject to them must be viewed as the legislators from whom they receive their authority, then a moral theory that does not represent moral requirements as self-legislated cannot explain their unconditional authority. Their normative necessity can only be grounded in the autonomy or law-giving activity of the will.

Why, we might ask, does Kant present us with these two alternatives: either one is bound to a practical principle by some contingent interest (in which case it applies conditionally), or one is its 'legislator' (in which case it applies unconditionally)? Why does he think that the unconditional authority of a principle can be explained only if the agent is its legislator, and what is the force of saying that moral agents must be regarded as 'legislating' moral principles? The answer is that if a practical principle applies unconditionally, it applies to one simply in virtue of being a rational agent. That is to say that one cannot coherently reject the authority of the principle while continuing to think of oneself as a rational agent, or without abandoning one's status as a rational agent. A practical principle can only have that status if it is constitutive of rational agency—that is, it is the formal principle describing what it is to exercise one's agency, and the capacity to follow the principle is what makes you an agent. A principle that is constitutive of rational agency is internal to the nature of rational agency, and that permits one to say that it 'arises from the nature of rational agency' or 'from the nature of the will'. I believe that the idea that a principle has authority because it arises from one's will (or from the nature of the will) is the central notion that leads Kant to think of moral agents as 'legislating' and underlies his frequent references to the will 'giving universal law'. A legislator is an agent with the power to give law through his will. Civil legislation gives us a transparent model of how agents can give law through their willing. In the political sphere, we understand how a duly constituted agent can enact law by following an authoritative legislative procedure, thereby creating reasons of authority for citizens to adhere to the law. Whenever a law gets its authority from the fact that it originates in an agent's will, it is appropriate to talk about an agent 'legislating' or 'giving' the law—or so we should understand Kant's use of this term. Kant thinks that moral subjects are legislators because a practical principle applies unconditionally only if it originates in one's will, and agents that can originate law through their will are legislators.

This line of thought offers a good argument for the conceptual claim that an agent subject to a practical law must be regarded as the legislator from whom it receives its authority, beginning from the unconditional authority of a practical law:

1. A principle has unconditional authority only if you cannot reject it and continue to think of yourself as a rational agent.

2. A principle is rejected at the cost abandoning one's rational agency only if it is constitutive of rational volition (i.e., only if it is a formal principle of rational volition).
3. A principle that is constitutive of rational volition arises from or originates in the nature of the will; in other terms, it's a principle that the will gives to itself.
4. Therefore, a principle has unconditional authority only if it originates in the nature of the will.
5. Since an agent with the power to originate law through his will is a legislator, an agent subject to a practical law must be regarded as its legislator.

This account needs some refinements because different kinds of practical principles 'arise from one's will' (are 'legislated') in different ways. The CI and particular categorical imperatives require different stories, and the above argument applies most naturally to the former. To connect these stories, it is useful to think of the FUL as a kind of legislative procedure that (following Kant's official view) can be used to determine whether a principle has the form of law. Given the procedural elements in Kant's understanding of moral assessment, we may think of the FUL as the 'constitution' of the rational will—a principle that sets out a legislative procedure whose application gives substantive principles of action (at a certain level of generality) their normative status. We use the FUL (the constitutive practical law) to determine whether a substantive principle has the form of law, and in this way arrive at particular categorical imperatives and substantive principles that can be used to justify actions (particular practical laws).

With this picture in mind, in what sense do moral agents 'legislate' particular categorical imperatives and substantive principles of justification? I believe that Kant's idea is that since a practical law applies unconditionally, its authority cannot be based on appeal to contingent interests, but instead must come from the reasoning (deliberative procedure) that makes it a law. In other words, sufficient reasons to accept the principle are given by whether it can be willed as universal law through the FUL. (It is important here that universalizability *confers* normative status on proposed principles of action.) That means that an agent is bound to the law by the reasoning that makes it a law. In holding that an agent is bound in this way, we presuppose that the agent has the capacity to carry out the deliberative procedure that makes the principle a law. But an agent who can employ that deliberative procedure has the same capacities as would be required of its legislator. I want to say that agents with this rational capacity can be regarded as a kind of legislator and that Kant's references to giving law through one's will are best understood in this light. Furthermore, an agent who acts out of respect for the moral law does carry this reasoning out and thus displays the volitional state of a legislator. I want to say that, in an extended sense, this agent may be regarded as giving law through his will (is in the volitional state that makes a principle a law).[3]

3. See *Agency and Autonomy*, pp. 99–104, 137–145. (All unattributed page references are to this work.)

These arguments show that in thinking of an agent as subject to practical laws, we presuppose that the agent has a legislative capacity. As one might say, 'subject' and 'legislator' are different aspects of the relation that we bear to moral principles. And this legislative standing is the basis of the dignity of the moral agent (*G* 4: 435–6, 440). The argument also goes some way toward establishing a restatement of the FUL as the FA. The principle that specifies the form of law at the same time specifies the form of law-giving—the principle that one must follow in order to give one's maxims the form of law, thus to give law through your will. Taking the autonomy of the will to be its law-giving capacity, the FUL is the constitutive principle of autonomy.

Much of the burden of these arguments is carried by the idea that the FUL is a legislative procedure that gives practical principles their normative status. But only an *authoritative* legislative procedure can be used to create law. Where does the FUL get its authority? That question is answered by connecting the FUL to the form of free volition (as Kant attempts in *Groundwork* III). Granting for a moment that Kant can link these ideas, in what sense would the FUL/FA be a law that 'arises from the will'? Here we need to turn to Kant's claim that the will "is a law to itself (independently of any property of the objects of volition)" (*G* 4: 440). I understand this as the claim that the FUL/FA is the constitutive principle of free volition: it is the principle that one must follow in order to exercise the power of free volition, and as such is not coherently rejected by any agent who regards himself as free and rational. It is thus the principle that arises, not from any particular act of volition, but from the nature of free volition. That makes the FUL/FA the principle that the will gives to itself.[4]

The idea that moral agents must be regarded as legislating universal law is not forced on the texts. This is Kant's language.[5] I have tried to explain why this language is appropriate by pushing certain parallels between moral law and civil legislation as far as one can—for instance, by thinking of the FUL as the 'constitution' of the rational will that sets out a 'legislative procedure' in the ways just sketched. However, these parallels give out at certain points, two of which I'll mention. First, a civil legislative body is, within certain limits, free to decide what principles to enact as law. There is an element of discretion in positive civil law (at both constitutional and ordinary level) that does not apply to moral principles. Second, civil law is not made until a legislative process is actually carried out. Moral principles, by contrast, are grounded in an idealized deliberative (constructive) procedure that can be carried out by

4. Chapter 4 of my book (pp. 104–9) offers a different take on the idea that the will is a law to itself that is closely related to the idea sketched in this paragraph. Beginning at *G* 412, Kant appears to extract the concept of an unconditional requirement (categorical imperative) from the idea of practical reason as a faculty of principles. He then derives a statement of the CI from the concept of a categorical imperative. The fact that the CI can in this way be derived from the nature of practical reason is one way to unpack the idea that the rational will is a law to itself.

5. At the key passage at *G* 431 Kant says, not that moral agents *are* legislators, but that moral agents *must be regarded* as legislators, or as author of the law. The more measured claim can be sustained if there are significant parallels between the bases of the authority of civil and moral laws, even if the parallels are imperfect.

any rational agent, and repeatedly.[6] The limits to these parallels would be more troubling to my approach if the main idea were that moral agents legislate the *content* of morality, but that is not the focus. Kant introduces the idea that moral agents legislate moral law to account for the *normative authority* of moral requirements. The different senses in which moral agents legislate or give moral principles for themselves attempt to spell out different aspects of Kant's principal idea—that moral requirements get their authority from the fact that they originate in our will.

To return to the overall trajectory of the *Groundwork*, the contribution of autonomy is to provide a bridge from the FUL to the nature of free volition. I have suggested that the *Groundwork* strategy for establishing the authority of the moral law is to establish that the FUL is the constitutive principle of a necessary self-conception. Our present concern is the claim that the FUL is the general form of free volition. The move to autonomy shows that the form of law is also the form of law-giving, and that the FUL is the constitutive principle of the autonomy, or law-giving capacity, of the will. To complete the overall argument, Kant needs to show that the form of law-giving is the general form of free volition; in other words, that autonomous law-giving and free volition have the same constitutive principle. Kant argues for this identity by claiming that autonomy—the capacity of the will to give itself laws independently of any external authority—specifies the positive concept of freedom. (This is a step in the argument that "a free will and a will under moral laws are one and the same" (*G* 4: 446–7).) The general idea is that free will is the capacity to initiate action without being determined by external influence—it is a self-originating cause. Since volition is the capacity to determine action through a principle, a will that can initiate action independently of external influence must be the source of the principles that guide its judgements about reasons. It must have the power to give itself laws out of its own resources, independently of external influence and authority—in other words, without accepting principles taken from another source. That power is autonomy and the FA its constitutive principle.

This argument is intriguing, but not wholly satisfactory. One problem concerns the conception of free agency. I think that one can show that the FUL is the constitutive principle of a sufficiently rich notion of free agency. The problem then is to justify ascribing that form of agency to ourselves on grounds other than its connection to morality. Kant apparently recognized this problem, since he abandons the strategy of *Groundwork*, III in favour of the Fact of Reason in the second *Critique*.

A more interesting question is whether one can connect the FA to a form of free volition that we may reasonably ascribe to ourselves. A conception of free agency should include the capacity to initiate action in some sense 'independently of external influence'. But 'initiate action independently of external influence' can be understood simply as 'initiate action normatively', where that involves adopting a principle of action through one's judgement that the

6. I address the limits of this parallel in *Agency and Autonomy* at pp. 103–4, 111–2, 122, 124, 142–9.

principle is worthwhile. The standard of choice here is one's own judgement, and it imposes almost no limits on what principles can be adopted. Here I am prepared to say (with other philosophers) that rational volition involves 'giving oneself a law': volition involves commitment to a practical principle that you find reason to adopt, where your commitment creates an additional content-independent reason, or normative requirement, to follow through and act on the principle. You are giving yourself a *law* in the sense that your choice or commitment creates an additional *requirement* to follow through that excludes the force of competing reasons for action, unless the commitment is reconsidered. Some principle of acting in such a way as to give yourself a law would be constitutive of rational volition, so understood. The problem is that when formulated, this principle is not Kant's FUL/FA. The form of agency just sketched involves making some general principle—*any* principle—into a law for *yourself*. But the force of Kant's imperative is to act from principles that can serve as practical laws for anyone.

Can one argue that CI captures the form of free volition that we reasonably may ascribe to ourselves? I'll take this question up in my response to Hill's remarks, to which I now turn.

III

Hill raises questions about the sense in which rational wills 'legislate' the basic moral law (the CI), and about the version of the view that rational choice is undertaken *sub ratione boni* that I ascribe to Kant. I'll address both points.

Regarding the first, Hill says that the language of 'legislating' or 'giving law' for oneself aptly captures the idea that the CI is not imposed by any external authority. But as he notes, "legislating" in this context would involve "turning a neutral principle into an authoritative 'law' that all rational human beings should respect" through the execution of an actual legislative process (p. 103 above). Since there is an a priori justification of the authority of the moral law, this model does not appear to apply.

In reply, two points. First, the term 'legislate' is indeed stretched in this context. But note that, although commentators, myself included, routinely talk about the rational will (or rational individuals) legislating the moral law, what Kant actually says in one canonical text is that the rational will "is a law to itself . . ." (*G* 4: 440). I interpret this claim through the idea that the FUL is the constitutive principle of free agency, which in turn gives sense to the idea that the FUL 'arises from the nature of rational volition or free agency'—though not of course from any particular act of volition. In this picture, no individual gives law, nor is there room for an actual deliberative process to create reasons of authority (in the way that civil legislation does). Thus we need not take Kant to claim either that we, or that the rational will, 'legislates the moral law' in the ordinary sense.

However, second, since this 'is' (in 'is a law to itself') is naturally read as 'gives', Hill can still ask: What is the point of saying that the rational will gives itself—i.e., legislates—the moral law even in an extended or metaphorical sense?

The answer is that the only possible ground of the unconditional authority of the moral law (Kant thinks) is that it 'arises from the will' or 'originates' in the nature of free, rational volition in the ways we have seen. 'Giving law' or 'legislating' may be at best a metaphor here, but since a legislator has the power to create law through his will, it is the appropriate metaphor.

Turning now to the thesis that rational choice is undertaken *sub ratione boni*, in several essays I ascribe to Kant the view that free and rational choice is guided by considerations that an agent takes to be reasons that are sufficient in the strong sense of carrying normative force for others. As I put the idea in one essay, all free and rational choice carries an implicit claim to justification in that it is guided by maxims that the agent regards as having some form of universal validity (p. 19). Hill doubts that Kant held this thesis (though we agree that the texts are inconclusive), and he has philosophical doubts about the thesis. However, he offers a useful critical discussion, and his suggestions about how I intend the interpretation, status, and role of the thesis are on target. First, the thesis, as he nicely puts it, is that rational choice is guided by standards that "prefigure and model the full Kantian justification" (p. 105 above). Second, regarding the status of the thesis, I agree that it is implausible as an empirical hypothesis and of limited interest as a stipulated feature of an ideal of rational conduct. I am inclined to think that a defence of the thesis needs to treat it, not as a conceptual thesis, but as a synthetic a priori feature of the practical perspective that informs our experience of ourselves as agents. This is obviously a difficult task, and I won't take it on here. Finally, Hill argues that the thesis cannot be used to launch a "substantial argument for the rationality of moral commitment" (p. 106 above)—to the effect that, since rational choice is implicitly guided by a standard of universal validity, any agent who acts for reasons is committed to the Categorical Imperative. I was reaching for the idea that the thesis could play this role in certain essays (especially Chapter 3), but I now share his misgivings: since a theorist is unlikely to find this conception of agency acceptable without already buying into large portions of Kant's moral conception, it is doubtful that the thesis can mount an independent argument for the conclusion that any rational agent is committed to the Categorical Imperative.

By way of response to Hill's challenge, let me consider what it might do for Kant to accept the idea that rational choice is undertaken *sub ratione boni* in the strong sense (without claiming to settle either the interpretive or the philosophical issues). I suggest that the thesis has some promise not primarily as a claim about acting for reasons, but as a component of a Kantian conception of free agency that can be used to make out the claim that the Categorical Imperative is the formal principle of free agency.

It is standard to characterize the negative dimensions of free agency—its freedom from certain kinds of influence—in terms of the causal and motivational independence of the will.[7] The positive dimension of free choice

7. 'Causal independence' is the fact that desires do not exert causal influence on choice. 'Motivational independence' is the fact that the presence of a desire is not per se a reason for action. See *Agency and Autonomy*, p. 154 (and references to Hill, Allison and Korsgaard in note 48).

involves, minimally, a spontaneous act of endorsement, such as the incorporation of an incentive into a maxim. To make sense of spontaneity, we need to think of it as normatively guided in some fashion—as a doing that is structured by some norm rather than as an arbitrary act. A natural move is to think of the spontaneity of choice as a 'taking to be good'. Factoring in the causal and motivational independence of free agency will place some distance between the act of spontaneity and the agent's desires and private interests, and that suggests that the taking to be good should be understood as an endorsement whose normative force is regarded as extending beyond the agent's private standpoint. Alternatively we might say that the requisite critical distance between the agent and his private interests (causal and motivational independence) is secured by taking these interests to be the basis of reasons that have some kind of normative force for others. Along these lines one might argue that free agency involves a taking to be good that makes an implicit claim to universal validity. That is, maxims are *freely* adopted on the supposition that they can serve as universal law in Kant's sense. The FUL would then provide the general form, or constitutive principle of free volition so understood: it is the norm that sets out how to exercise one's agency, and the capacity to follow this norm would be what makes you a free agent.

If the FUL is the formal principle of free volition, then it is implicated in or tacitly guides all exercises of free agency, even if imperfectly—so that actions in no way guided by the principle are not true instances of agency. One might think that this introduces a problem for this view: How can bad action (action on non-universalizable maxims) be guided by the FUL? In response, first, it is not hard to think of examples of actions that are guided by a constitutive rule that at the same time fail to conform to the rule. Someone who makes a mistake in addition (or a syntactical error, a mistaken inference, etc.) is in some sense following the rule of addition (or a rule of syntax, a principle of inference), although she violates the rule. Here one has to hold that bad action is guided by the FUL in the way that someone who makes a mistake in addition is guided by or follows the rule of addition. There are genuine puzzles here, but the fact that we have examples of actions that are guided by, while at the same time failing to conform to, a rule suggests that the puzzles are soluble.

Second, this problem appears to be the familiar problem of how, given Kant's identification of free agency with the capacity to act from the FUL, bad action can be free action. A standard resolution of the latter version of the problem is to hold that free agency is a capacity that is sometimes exercised, sometimes not. It is enough for an action to be free (therefore imputable) if chosen by an agent with the capacity to follow the Categorical Imperative, whether or not the capacity was exercised in that instance. The thesis that free action is undertaken *sub ratione boni* in the strong sense adds a further wrinkle. Since it is the thesis that all free action is guided by maxims regarded as having some form of universal validity, it allows one to say that the FUL guides or is prefigured in all free choice. Acting *sub ratione boni* is the form of free agency common to both good and bad choice. Bad choice is no less free, but is distinguished by the defective manner in which it instantiates the fundamental norm.

In sum, the thesis that free action is undertaken *sub ratione boni* offers a way to characterize the normative dimension of the spontaneity needed for free choice, in a way that indicates how the Categorical Imperative is prefigured in all forms of free choice, including choice that fails to conform to moral standards.

IV

Samuel Kerstein picks up on a discussion in 'Autonomy of the Will as the Foundation of Morality' in which I claim that it is part of the concept of a practical law that a practical law contains the ground of its authority in itself. He is quite right to raise questions about this idea; it may well be abstract nonsense. But let me try to reconstruct what I thought I was getting at. To summarize a few points: First, I understand a practical law as a practical principle that applies unconditionally and can authoritatively settle questions of justification. Reflection on this concept suggests two conditions: (1) Since a practical law authoritatively settles questions of justification (by bringing the search for justifications to an end), it cannot get its support from any further principle.[8] (2) Since a practical law has authority for rational agents, there must be some explanation of its authority. (We cannot expect rational agents to accept a principle unless we can explain why they should.) I summarize these two conditions by saying that a practical law contains the ground of its authority in itself (in some internal structural feature). (Kerstein calls this the 'Validity Condition', but I intend it as a way to combine these two conditions, and not a separate condition.)

Second, I speculate that the idea that a practical law contains the ground of its authority in itself can explain some puzzling formulations of the CI, which direct us to act from maxims that can "contain in themselves their own universal validity".[9] This alleged feature of practical laws seems to be reflected directly in these formulations of the CI. Third, I unpack the idea that a practical law contains the ground of its own authority through Kant's idea that the authority of a practical law comes from its having the form of law. If the normative authority of a principle comes from its form, then it does not get its support from any further or external principle; but there is something to say about the basis of its authority—namely, it has the form of a law. Why does that give it normative authority? Certainly the rejoinder, 'Because that is what it is to be a law', is not responsive. The full *Groundwork* response is that the form of law is the form of free volition, and one cannot reject the regulative norm of acting only from maxims that have the form of law and continue to think of oneself as a free agent.

I sketch a defence of the Sovereignty Thesis that does not require the claim that a practical law contains the ground of its own authority in the argument given in section II above. The main idea, again, is that in asking how a principle can apply unconditionally, we are led to the idea that its authority

8. Cf. Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge University Press, 1996), pp, 55–67.
9. These formulations of the CI are listed in *Agency and Autonomy* at p. 134, and p. 166 n. 29.

is based in the agent's will, in which case, the agent is the 'legislator' from whom it receives its authority. But while the Sovereignty Thesis does not depend on this idea, I do not want to abandon it too quickly. Kerstein doubts that the FUL satisfies this joint condition and argues that principles that Kant would not consider practical laws might well satisfy it. But these critical points can be answered if one assumes some machinery from the *Groundwork* and unpacks 'contains the ground of its authority in itself' in terms of 'has the form of law'.

Given the fact that Kant's attempted deduction relies on claims that have nothing to do with 'internal features' of this principle, does the FUL contain the ground of its authority in itself? According to the *Groundwork*, Yes. The authority of the FUL comes from the fact that it expresses the form of free volition (coupled with the claims that we necessarily act under the idea of and identify with our freedom). This story takes some work, and it is not evident in the initial statements of the principle. But one aim of the argument launched at *G* 420 is to show that the principle that expresses the form of a practical law also specifies the form of free volition.

Do particular practical laws contain the ground of their authority in themselves? Substantive principles of this sort get their normative force from the fact that they result from an authoritative deliberative procedure. Agents are bound to these principles by the reasoning that makes them laws, and that reasoning applies a procedure based in the nature of free volition. Doesn't that mean that they are grounded in some higher principle (thus do not contain the ground of their own authority)? No, because this procedure shows that these principles have the form of law—that the components of these principles are related in such a way that they can serve the justificatory role of a practical law. And that means that an agent who acts from such a principle exemplifies the form of free volition.

Couldn't practical principles that rational intuitionism regards as self-evident truths about right action contain the ground of their own authority? No, not as they are understood by intuitionism. In treating fundamental principles as self-evident, intuitionism proceeds as though no explanation of their normative authority is necessary, or is possible. So understood, such principles don't contain the ground of their own authority, because there is no such ground—i.e., there is no explanation of their normative grip. (The intuitionist may *take* these principles to be authoritative, but what explains their normativity?) If the intuitionist, after claiming the intrinsic goodness of developed human capacities, tells us that it is necessary to perfect one's rational and physical capacities, we can coherently ask for an account of their intrinsic goodness and for an explanation of why self-perfection is a necessary end. Now self-perfection may be a necessary end. Likewise, if the intuitionist tells us that the relation in which promisor stands to promisee makes it fit and incumbent upon the former to fulfil his promise, that principle may have unconditional authority. But the explanation of the authority of these principles will be Kant's (that self-perfection is a necessary end because we are committed to willing the necessary conditions for exercising our agency, that deception for self-interest cannot coherently be willed as law for agents with autonomy—in short, because the relevant principles have the form of a law, which in turn is the form of free volition).

Toward the end of his comments, Kerstein asks whether I have shown that the FUL is the only principle that can serve as the constitutive practical law of a will with autonomy. Why couldn't his WU ('weak universalization') or BP (the bizarre principle of acting only from non-universalizable principles) be accepted as constitutive principles of autonomy? If a community of agents regarded, say, WU as authoritative, they would be committed to a principle of non-beneficence. But how could agents with autonomy come to regard WU (or BP) as an authoritative deliberative principle?

Since a constitutive principle establishes the possibility of a certain activity, it cannot be selected arbitrarily. Here we are talking about the enabling conditions of autonomy, understood as a capacity to give authoritative law for agents with autonomy. Some of Kerstein's worries can be handled by noting that law-giving is addressed to some community of agents. Since autonomy is the capacity to give *unconditional* law, this community consists of agents who share the same legislative capacities. To succeed in 'giving universal law', one must guide one's will by reasoning that can move other agents with the same legislative capacities to accept the resulting principles. Kerstein and I both agree that WU (and BP) lead to counter-intuitive conclusions. I would put this by saying that they lead to substantive principles that agents with autonomy have no reason to accept. I take that to show that WU does not establish a procedure that enables one to give universal law. A principle that is constitutive of autonomy must enable one to give universal law that can be accepted as law by agents with the same legislative capacities, and, moreover, through the reasoning that goes into the principle. WU does not do that, and as far as I can see, only Kant's (or an equivalent) principle does.

V

Space constraints permit only a brief reply to Jens Timmermann's thorough analysis of Kant's principles of imputation. In my essay, I suggested that these principles can be extended to strict moral duties, so that, for example, any bad consequences of a violation of a perfect (though not legally enforceable) duty are imputable to the agent. Timmermann is certainly right that Kant focuses on violations of juridical duties, where he holds that the agent is legally liable for all subsequent bad consequences. I am happy to concede that these principles should not extend to duties of beneficence, even where the balance of reasons requires an agent to provide aid. Timmermann makes the useful point here that assistance can be both necessary (required on balance) and meritorious.[10] However, Kant's example of the servant who lies on the householder's order (*MdS* 6: 431) indicates that Kant applies a form of these

10. However, I don't think that the account in my paper would obliterate the distinction between killing and failing to save. They remain different kinds of deeds, since deeds are partly individuated by moral principles. If a death were (morally) imputed to an agent who failed to give life-saving aid in a situation where aid was required, it would be death due to a failure to save. That is not the same result as death from killing.

principles to perfect moral duties. Moreover, I see no argument for the claim that the principles of imputation apply only to legal duties and not perfect moral duties. In the latter case, of course, we do not have legal liability, but moral imputability; as Kant says, the servant's 'conscience' (not civil law) imputes the results of the lie to himself.

Imputing unforeseeable consequences of actions that violate strict (juridical or perfect moral) duties is indeed problematic. I doubt that the issue can be resolved conclusively, since we lack fixed shared intuitions about such cases. However, Timmermann overlooks the idea of 'authorship' in my account that may provide some rationale for imputing unforeseeable consequences to an agent up to a point. The idea is that when you comply with a strict requirement, you act under the authority of the law (civil or moral), not your own. For that reason, you are not the author of any resulting bad consequences. But when you violate a strict requirement, you 'act on your own authority'. In that case, we might think that you assume certain moral risks, and should be regarded as the author even of unforeseeable bad consequences, which may therefore be imputed 'to your account'.