

AUTONOMY AND PRACTICAL LAW

SAMUEL J. KERSTEIN

The University of Maryland, College Park

It's a pleasure to comment on *Agency and Autonomy in Kant's Moral Theory*. I have learned a great deal from Andy Reath's work over the years. For example, his essay 'Hedonism, Heteronomy, and Kant's Principle of Happiness', which is contained in the book, opened my eyes to new possibilities for interpreting Kant's views on non-moral action.¹ Reath's work on Kantian autonomy is striking for its rigour and depth. I know of no more helpful and engaging approach to the topic in the literature. So I'm going to focus my comments on an essay that appears for the first time in the book: 'Autonomy of the Will as the Foundation of Morality'. It's an intricate enterprise to reconstruct Kant's thinking about autonomy. So it's no surprise that the arguments in the essay are very complex. Some of my comments might simply reflect misunderstandings of them. But let me plunge ahead nonetheless.

Reath sees two ideas as central to Kant's view that autonomy is the basis of morality. The first is what he calls the 'Sovereignty Thesis', according to which "an agent who is subject to an unconditionally valid principle (i.e., a practical law) must be (regarded as) the legislator from whom it receives its authority" (p. 122).² Second is the idea that the Formula of Universal Law (FUL) "is the basic principle of a will with autonomy" (p. 124).

Let me explore briefly how, according to Reath, these ideas come together in Kant's thought on autonomy. The Sovereignty Thesis suggests that moral agents are in some sense legislators. Reath specifies two senses in which they are. Moral agents have "legislative capacity", that is, "the capacity to carry out the reasoning that makes a principle a law" (p. 123). Moreover, an agent who acts from duty, that is, out of respect for the moral law, "actually carries out the reasoning that makes a principle a law" and gives law through her will (p. 123). It is, says Reath, "by following [the FUL] that one gives law through one's will". The FUL is the principle that is constitutive of an agent's legislative power.

It might be helpful to illustrate this view with an example. Suppose that an agent is considering acting on a maxim of making a false promise solely for financial gain. She asks whether she can act on the maxim and, at the same time, will that it become a universal law. She finds that she cannot do this. Acting on her maxim would violate the FUL. On the grounds that it would, she refrains from doing so. According to Reath, she has established a law against the maxim's adoption "by conferring the status of 'impermissible'" on the maxim (p. 144). She has given herself and others "authoritative reason"

1. I disagree with some aspects of his interpretation. See Samuel J. Kerstein, *Kant's Search for the Supreme Principle of Morality* (Cambridge University Press, 2002), pp. 22–29.
2. All references in the text are to Andrews Reath, *Agency and Autonomy in Kant's Moral Theory*.

to accept that it is wrong to act on the maxim in question (p. 144). Implicit in the Sovereignty Thesis, according to Reath, is that all agents have the capacity and thus the authority to make the law that our agent has made here.

The FUL, says Reath, sets out a procedure that enables one to give law and that one must follow in order to give law. The FUL “confers lawgiving power” on the agent (p. 148); it makes it possible for her to establish the law that it is wrong to adopt the maxim of false promising. But an agent could not give law by following some other, incompatible procedure, such as that of *rejecting* any maxim that could be a universal law. If an agent could give law by following this procedure, then she would presumably be able to establish the law that is wrong to act on a maxim of sincere promising. In sum, the FUL both enables an agent’s giving of law to take place and provides a significant constraint on it.

A final preliminary note: Reath refers to particular principles such as the law that it is wrong to act on maxims of false promising as well as to the FUL as practical laws (e.g., p. 165, n. 21). He suggests that the will is subject to two kinds of practical laws, ones that are “constitutive of its legislative powers” and ones that “it legislates for itself through their exercise” (p. 131). It will be convenient to refer to the former as constitutive practical laws and the latter as particular practical laws. Practical laws of both kinds are universally valid (p. 164, n. 16) and unconditionally binding; all of us are rationally compelled to act in accordance with them regardless of our desires or contingent interests (p. 140).

Reath reconstructs Kant’s arguments for both the Sovereignty Thesis and the idea that the FUL is the basic principle of a will with autonomy. In what follows I probe these arguments. For the most part, I do not ask to what extent they square with Kant’s text (namely the *Groundwork of the Metaphysics of Morals*, in which Reath claims to find them), but rather raise some issues regarding their philosophical plausibility. I am unsure of the extent to which Reath aims to defend the arguments as opposed to offer them as textually plausible interpretations. He frequently uses language that suggests a defence (see pp. 135, 148, 152). For example, he claims that the Sovereignty Thesis “follows analytically from Kant’s conception of a practical law” (p. 5). Perhaps what he means here is really just that, *according to Kant*, this thesis follows analytically from the concept of such a law. If his project is simply interpretative—and in my view there would be nothing at all wrong with that—then my comments will not contain much criticism of his essay. In any case, I hope that they serve as a springboard for discussion.

I take the claim that the Sovereignty Thesis follows from the concept of a practical law to mean that, upon reflection, it is self-contradictory or incoherent to affirm that there is a practical law and yet to deny that “agents subject to moral requirements must be regarded as their legislators and the source of their authority” (p. 123). In order to get from the concept of a practical law all the way to the Sovereignty Thesis, Reath implies (e.g., p. 138), Kant appeals to (and needs to appeal to) the idea that “a practical law must contain the reasons for its validity and the source of its own authority in itself” (p. 132). Let’s call this the Validity Condition on a practical law. Reath emphasizes that

the Validity Condition does not imply that a practical law needs no justification. “Rather,” he says, “one must look for its justification (both of a particular law and of the higher order law stated by the Categorical Imperative) in a certain place, namely in some internal feature of the principle” (p. 165, n. 21).

I suspect that there are two problems with the Validity Condition. First, it seems too strong. If we focus on the grounds Kant offers for the validity of the FUL in the *Groundwork*, it seems that in his own view the FUL itself fails to fulfil the condition. Second, the Validity Condition seems to be without sufficient warrant. I don’t believe it follows from the concept of a practical law since one can, I think, coherently count as such a law a principle that fails to fulfil the condition.

Regarding the first possible problem, it is unclear to me precisely when the justification of a principle can or cannot legitimately be said to lie in some internal feature of it. But I gather that if a principle is self-evident, then its justification lies in such a feature. Moreover, Reath suggests that the suitability of a maxim to be acted on and at the same time willed to be a universal law constitutes an internal feature of the maxim. If a maxim is justified just in case it has this feature, then its justification lies in some internal feature of it. In contrast, if a principle owes its justification solely to God’s having proclaimed it, then its justification fails to lie in an internal feature.

In his *Groundwork* III attempt to prove the validity of the FUL (or an equivalent principle), Kant seems to appeal to the premise that a rational being must assume that its theoretical judgements are spontaneous.³ But this premise does not constitute an internal feature of the FUL which is, after all, a practical not a theoretical principle. So if Kant does appeal to this premise, FUL fails to fulfil the Validity Condition. Of course, it’s beyond the scope of these remarks to enter into controversies regarding the interpretation of the *Groundwork* III deduction. But one of the most careful, textually sensitive contemporary scholars working on it, Dieter Schönecker, makes the case that Kant appeals in his argument to an “ontoethical principle”, namely to the principle that “[T]he world of understanding and thus the pure will as a member of this world of understanding are ontically superior to the world of sense . . .”.⁴ The FUL does not contain this principle. If Schönecker is correct and Kant indeed relies on the ontoethical principle for his justification of the FUL, then in Kant’s own (*Groundwork*) view, the FUL fails to fulfil the Validity Condition.

But if in the *Groundwork* Kant implies that the FUL fails to fulfil this condition, then we have reason to doubt that Kant there embraces the condition in the first place. We have grounds for wondering whether in the *Groundwork* Kant really holds that a practical law must contain *the* reasons for its validity and the source of its own authority in itself.

3. Immanuel Kant, *Groundwork of the Metaphysics of Morals*, tr. Mary Gregor, in *Immanuel Kant: Practical Philosophy* (Cambridge University Press, 1996), p. 448. I am referring to *Preussische Akademie* (vol. IV) pagination, which is included in the margins of the Gregor translation.

4. ‘How is a categorical imperative possible?’, in Christoph Horn and Dieter Schönecker (eds.), *Groundwork for the Metaphysics of Morals* (de Gruyter, 2006), p. 318.

In reply to the objection that the Validity Condition is too robust for Kant to accept, Reath might appeal to Kant's doctrine of the 'fact of reason'. Recall that in the *Critique of Practical Reason*, Kant states:

the moral law is given, as it were, as a fact of pure reason of which we are a priori conscious and which is apodictically certain, though it be granted that no example of exact observance of it can be found in experience. Hence the objective reality of the moral law cannot be proved by any deduction, by any efforts of theoretical reason, speculative or empirically supported, so that, even if one were willing to renounce its apodictic certainty, it could not be confirmed by experience and thus proved a posteriori; and it is nevertheless firmly established of itself.⁵

Reath might claim that Kant here abandons his *Groundwork* position that the validity of the FUL can be established through philosophical argument. According to Kant's considered view, he might say, the moral law (i.e., the FUL) is a self-evident practical law and thereby contains the reasons for its validity and the source of its own authority in itself.

Now let me move to my second concern, namely that the Validity Condition has insufficient warrant. Reath seems to defend the condition in the following way. He notes that, according to Kant, practical laws must be such that the moral duties that stem from them are necessary, that is, "apply independently of and limit the reasons given by an agent's desires", and universal, that is, capable of being regarded as valid by anyone (p. 132). In order to fulfil these criteria, says Reath, a practical law must fulfil two conditions. First, it must provide sufficient justification for an action. "It must provide the final reason why the action is required or right in that it, or the pattern of reasoning which it initiates, brings the search for reasons to a close" (p. 132). Second, a practical law "must have some kind of justification or rational support", support that "explains why it is valid and which is available to move an agent to adhere" to it (p. 132).

Let us grant for the sake of argument that a practical law's having to fulfil these two conditions does, as Reath suggests, follow from the very concept of such a law. Reath then moves immediately to embracing the Validity Condition. He implies that since the two conditions are contained in the concept of a practical law, the Validity Condition is as well (pp. 132–133).

Now at a couple of points in the chapter, Reath discusses rational intuitionist principles. But it doesn't seem to me that he provides a promising path toward eliminating the possibility that such principles fulfil the two conditions in question, yet fail to satisfy the Validity Condition.

Let me explain. Take the principle 'Act so as to perfect your rational and physical capacities'. Its defender believes it to be a categorical imperative:

5. Immanuel Kant, *Critique of Practical Reason*, tr. Mary Gregor, in *Immanuel Kant: Practical Philosophy* (Cambridge University Press, 1996), p. 47; see also pp. 48 and 93. I am referring to *Preussische Akademie* (vol. V) pagination, which is included in the margins of the Gregor translation.

a principle that we, dependent rational beings, always have overriding reason to comply with. She holds that this principle states a truth about right action that obtains in virtue of the intrinsic goodness of rational and physical capacities' perfection. As a result of having reason, which is among other things a *recognitionnal capacity*, any agent is able to discern this intrinsic goodness, the defender insists. The validity of the principle, and thus its authority, is established through reason's recognitionnal capacity. Finally, an agent's grasping the principle's validity has motivating force. (Much of this description of a rational intuitionist position is mentioned by Reath himself (p. 137).) Notice that the justification of this principle would involve appeal to something outside the principle itself, for example the intrinsic goodness of well-developed theoretical and physical capacities. So the principle would not fulfil the Validity Condition. Yet, as far as I can tell, it *could* fulfil the two conditions in question: it *could* count as both providing sufficient justification for an action and as having some sort of rational support.

If, as Reath seems to think, the concept of a practical law really led via just the set of the two conditions in question to the Validity Condition, then any principle that failed to fulfil the latter would fail to fulfil the former. But that doesn't seem to be the case. (Of course, my aim here isn't to defend a perfectionist version of rational intuitionism. I'm just trying to make a conceptual point.)

Kant might respond to a rational intuitionist principle, suggests Reath, by claiming that its normative force remains unexplained and that it appears to be rationally rejectable. One "can coherently ask why one should follow it" (p. 166, n. 22). But the same points can be made regarding the FUL, at least if adherence to the Validity Condition forces Kantians to "justify" it by stating that, as a fact of reason, it is firmly established of itself. Stating this neither explains the FUL's normative force nor renders incoherent the question of why one should follow it.

Reath says he aims to show that "reflection on the concept of a practical law leads to the conclusion that such a principle must contain the grounds of its validity and authority in itself" (p. 135). He is, I gather, trying to demonstrate that the Validity Condition follows from the concept of a practical law, not merely that, rightly or wrongly, Kant thinks it does. I don't believe that Reath has shown this. Moreover, if, as Reath seems to hold, in order to get from the concept of a practical law to the Sovereignty Thesis we must proceed via the Validity Condition, then in my view he has also fallen short of establishing that the Sovereignty Thesis follows from the concept of a practical law.

Now let me turn to the second aspect of Reath's chapter that I would like to probe. Another significant Kantian claim he appears to defend is that if we assume that an agent has autonomy, then we are rationally compelled to hold that the FUL is binding on her. From materials provided by Kant, Reath constructs the following argument:

- (c1) A will with autonomy is not bound to any external authority and has the power to create law through its will.
- (c2) The principle of an autonomous will is the FA: the principle of acting only from maxims through whose adoption one can regard oneself as giving law through one's will.

- (c3) If an autonomous will is to create law through its willing, it cannot take its reasons for enacting law from any contingent interests, or enact laws whose authority comes from contingent interests in those whom they address. It must guide its legislative activity by reasoning that anyone can regard as valid, and thus will enact principles that can serve as practical laws (have the form of law).
- (c4) Thus the fundamental legislative principle of a will with autonomy, through which it exercises its sovereignty, is to will only principles with the form of law (principles that can be willed to serve as practical laws in accordance with the FUL) (c1–3).
- (c5) Since a legislator is bound to its own laws, a will with autonomy has the principle of acting only from maxims that can be willed to serve as practical laws (c4). (pp. 151–152)

In effect, this argument purports to show that the constitutive practical law for a will with autonomy, that is, the principle that both enables it to give particular laws and restricts the particular laws it can give, is the FUL. But I don't believe the argument eliminates the possibility that other, different, principles serve as the constitutive practical law for a will with autonomy. (The reasons for my doubts here are similar to the reasons I give elsewhere for concluding that neither Korsgaard nor Allison offers a philosophically successful reconstruction of Kant's derivation of the categorical imperative.)⁶

I'll consider two principles, but I don't think it would be hard to generate more. Take first the weak principle of universalization WU: 'Act only on that maxim which, when generalized, could be a universal law.' Notice that WU is not equivalent to the FUL, in Kant's view. As he suggests in the *Groundwork*, a maxim of non-beneficence could, when generalized, constitute a universal law.⁷ Since a world where no one acted beneficently is indeed a coherent possibility, acting on a maxim of non-beneficence does not violate WU. Kant does, of course, think that acting on such a maxim runs afoul of the FUL. It does so, he believes, because as a rational agent it is not possible to act on it and, at the same time, will that its generalization be a universal law. Let's suppose that Kant is right about all of this.

As far as I can tell, Reath's argument fails to demonstrate that an autonomous agent could not legislate a particular (permissive) law of non-beneficence on the basis of the constitutive law WU. Premises c1–c2 obviously don't rule out WU as a constitutive law. Regarding (c1), WU does not invoke any external authority. And c2 just reminds us what FA is. So if WU did get eliminated, it would have to be on the basis of c3, which, again, states the following: "If an autonomous will is to create law through its willing, it cannot take its reasons for enacting law from any contingent interests, or enact laws whose authority comes from contingent interests in those whom they address. It must guide its legislative activity by reasoning that anyone can regard as valid, and thus will enact principles that can serve as practical laws (have the form of law)." But why

6. See Kerstein, *Kant's Search*, pp. 33–45 and 74–76.

7. Kant, *Groundwork*, 423.

would an agent's authority to 'legislate' a (permissive) law of non-beneficence through WU need to be based on any contingent interest? And why would it not be possible, rationally speaking, for each and every agent to regard this law as valid? Of course, in answering the latter question, it would be circular to claim that the principle of non-beneficence fails to accord with the FUL and thus fails to have the 'form of law'. For this claim just assumes the truth of the proposition we are questioning, namely the proposition that it is accordance with FUL, but not with WU, that can give a particular principle the form of law. The principle of non-beneficence *would* have the form of law if the constitutive law of a will with autonomy was WU.

Next consider the bizarre principle BP: 'Act only on that maxim such that you *cannot*, at the same time, will that it become a universal law.' It is not hard to find maxims that would be in accordance with BP, but not FUL. Once again, the maxim of non-beneficence Kant discusses in *Groundwork* I fits the bill. If the maxim of non-beneficence fails the FUL test (as Kant holds it does), then it passes the BP test. Indeed, it seems that if BP were a constitutive practical law for autonomous agents, then one of the particular laws such agents would legislate would be one according to which it is wrong to act on a maxim of beneficence. Where does Reath's argument rule out the possibility that BP is a constitutive practical law? Granted, such a practical law must guide the giving of particular laws by reasoning that anyone *can regard* as valid. But why would it, rationally speaking, be impossible for each and every agent to view as valid the procedure for legislating particular laws specified in BP? So far as I can tell, Reath's argument does not give us an answer to this question.

I don't believe Reath's argument shows the FUL to be the only principle suited to be a constitutive practical law for an agent with autonomy. I should mention that in a different essay, namely 'Agency and Universal Law', Reath seems to offer a slightly different argument for this sort of conclusion (p. 206). But this argument seems to me to suffer from the same shortcoming, namely that of failing to eliminate rival 'formal' candidates for status as constitutive practical laws.

Needless to say, I'm not an advocate of BP. And I have no wish to defend WU. But I reject these principles not because a priori analysis reveals incompatibility between them and the very concepts of a practical law or of an agent with autonomy, but rather because they generate counterintuitive practical prescriptions. I am much more doubtful than Reath seems to be that exploring these concepts can yield striking philosophical results.