

Autonomy and the Idea of Freedom: Some Reflections on *Groundwork* III

ANDREWS REATH
University of California, Riverside
Email: reath@ucr.edu

Abstract

This article explores a set of questions about the ‘idea of freedom’ that Kant introduces in the fourth paragraph of *Groundwork* III. I develop a reading that supports treating it as a normative notion and brings out its normative content in some detail. I argue that we should understand the idea as follows: that it is a general feature of reasoning and judgement that it understands itself to be a correct or sound application of the normative standards of the relevant domain of cognition, not influenced by irrelevant or external factors. Reasoning and judgement are thus normatively committed to these standards of correctness. A second and related concern is to explore connections between the idea of freedom and Kant’s conception of autonomy and to identify different points at which autonomy plays a role in the argument of *Groundwork* III. In the final section, I mine the idea of freedom for a set of normative commitments specific to rational agency that play a foundational role in Kant’s moral conception.

Keywords: Kant, autonomy, freedom, authority of morality

This article explores a set of questions about the well-known ‘idea of freedom’ (*Idee*, but in what follows simply ‘idea’) that Kant introduces in the fourth paragraph of *Groundwork* III. Various commentators have suggested that Kant’s idea of freedom is fundamentally a normative rather than a metaphysical notion. I develop a close reading of relevant texts that supports this general approach and that brings out the normative content of the idea of freedom in some detail. I argue that we should understand it as follows: that it is a general feature of reasoning and judgement that it understands itself to be a correct or sound application of the normative standards of the relevant domain of cognition, not influenced by irrelevant or external factors. Reasoning and judgement thus proceed

under the idea of freedom in that they are normatively committed to these standards of correctness. A second and related concern is to explore connections between the idea of freedom and Kant's conception of autonomy. I argue that a notion of autonomy is a key component of the idea of freedom, and recognition of this fact supports the normative reading of the latter. I also identify different points at which autonomy plays a role in the argument of *Groundwork* III, beyond its appearance in the opening argument in §2, that have not been generally recognized.

There is general consensus that Kant intends to ground morality in the normative commitments that we have as free rational agents, or as agents with autonomy. Once one understands Kant's idea of freedom as fundamentally a normative notion, it is a natural place to look for a set of normative commitments that play a foundational role in Kant's moral conception. In the final section of the article I suggest that the idea of freedom contains a set of individual normative commitments that provide input to the categorical imperative and that point to a core set of substantive principles in Kant's moral conception. This section of the article is less textually based, but still I hope on target.

The article is structured around a series of questions that are given in the section headings. Section 1 explores the connection between Kant's conception of freedom of will and his conception of autonomy. Sections 2 and 3 develop a reading of the idea of freedom that shows that it contains an 'Autonomy Condition' on reasoning and that explains its normative content. Section 4 is a digression on what metaphysical implications, if any, attach to the idea of freedom. Section 5 explores the role played by the idea in the balance of the preliminary arguments that set up the 'deduction' of the categorical imperative in *Groundwork* III; here I suggest some reasons for finding the overall argument of *Groundwork* III to be inconclusive. Finally, in section 6 I mine the idea of freedom for a set of normative commitments specific to rational agency that can be assigned a foundational role in Kant's moral conception.

1. What is the Relation between Freedom of the Will and Autonomy of the Will?

Kant tends to identify the freedom of the will with its autonomy (*G*, 4: 446.24ff., 'what else, then, can freedom of the will be, but autonomy . . .?', 450.23, 461.15; *KpV*, 5: 33).¹ However as I understand autonomy of the will, that is not quite right. The will's autonomy, rather,

is a component and a necessary condition of its freedom, but they are not exactly the same property.

Although Kant does not yet clearly distinguish *Wille* and *Willkür* in the *Groundwork* and the second *Critique*, presumably freedom of the will in these works is the freedom of *Willkür* (the power of choice). It is thus a feature of the *causality* or *executive aspect* of the will in the broad sense as the capacity to realize an action that has been represented as rationally supported or good. The metaphysical (or ‘cosmological’) idea of free causality is that of a first causality – an ‘absolute causal spontaneity beginning from itself a series of appearances’ that is not determined by any temporally prior cause (*KrV*, A446/B474; cf. A533/B561). When this idea is applied to the power of choice, it is specified as ‘the power (*Vermögen*) of pure reason itself to be practical’ independently of sensible incentives (*MdS*, 6: 214). Thus Kant ultimately understands freedom of the power of choice as the power to act from principles of pure practical reason: the motivational capacity to act from moral principles.

However, the autonomy of the will, as I understand it, is not (in the first instance) a feature of the *causality* of the power of choice. Rather it is a feature either of *Wille* (of the ‘legislative’ aspect of volition) or of the will in the broad sense (comprising both a cognitive or legislative aspect and a causal or executive aspect). Kant defines autonomy of the will as ‘the property of the will by which it is a law to itself (independently of any property of the objects of volition)’ and argues that the categorical imperative is that law (4: 440.16–20, 447.1–6). As I interpret him, the thesis that the rational will is a law to itself is the claim that the nature of rational volition, or of practical reason, is the source of its fundamental norm.² That principle is the formal principle (Kant’s terminology) or the internal constitutive principle (contemporary terminology) that specifies what it is to exercise the will, and therefore authoritatively governs all volition. So understood, Kant’s thesis of the autonomy of the will is introduced as part of his foundational argument for the authority of morality: a practical principle applies unconditionally only if it ‘arises from’ the nature of the will *a priori* (433.2).

Although Kant’s thesis of the autonomy of the will appears in the second half of *Groundwork* II (after 4: 431), earlier arguments are intended to show that the will, or practical reason, is a law to itself. I have in mind the arguments that introduce first the Formula of Universal Law (FUL) (412–21), then the Formula of Humanity (FH) (427–9), where Kant moves from a conception of practical reason to a formula of the

categorical imperative (CI). In these arguments Kant articulates a conception of practical reason as a capacity – a conception of what practical reason is, how it operates, or its formal aim, etc. – and moves from there to a statement of its formal principle. If these arguments succeed, they show that the will, or practical reason, is a law to itself.

To clarify the essential points of my reading of autonomy: first, to say that the will is a law to itself (independently of any property of the objects of volition) is to say that the nature of rational volition or practical reason is the source of its fundamental normative principle *a priori*. This principle is the formal principle that specifies what it is to exercise the faculty (its internal principle of operation) that as such unconditionally regulates its exercise. Second, Kant's claim that the will has autonomy is supported by, and can be understood in terms of, the arguments that move from an analysis of rational volition or practical reason as a faculty (from its nature, its formal aim) to the first two formulations of the CI. Finally, the thesis of autonomy of the will is introduced as a key component of Kant's explanation of the unconditional authority of morality.³

Why then are the freedom and the autonomy of the will not quite the same property? Freedom is a feature of the power of choice – of the causal or executive aspect of volition. It is the capacity to act from moral principle. But the autonomy of the will is a feature of the 'legislative' or 'cognitive' aspect of the will – the fact that the nature of rational volition, or of practical reason, is the source of its own formal norm. So autonomy and freedom attach to different aspects of the will.

However, this is a fairly small point. We may certainly assume that a will with autonomy in this sense has the capacity to act from its own principle. If so, a will with autonomy is a free will. The reason to make this rather subtle distinction is that isolating the property of autonomy in this way makes it clear why the capacity to act from moral principle is a practical specification of the concept of freedom. Free causality is a 'causal spontaneity beginning from itself' – a self-determining power. If the rational will is to be genuinely self-determining, one might think that it must in some way be the source of the basic normative principles that guide its exercise – it must be 'the author of its own principles' or have autonomy. Thus autonomy of the will is a feature that has to be in place for the will, in its power of choice, to be genuinely self-determining; it is a necessary condition of the will (power of choice) being free. The pure will is free because it has the capacity to act from 'self-given' principles.

I will suggest presently that this conception of autonomy can be generalized to any rational or cognitive power governed by its own formal principles. In this case the rational power will be ‘the author of its own principles’. The obvious thought here is that for a normatively governed capacity to be self-determining or free (a capacity for spontaneity), it must have autonomy.

2. What is the Idea of Freedom Introduced in *Groundwork III*, Paragraph 4?

The first sub-section of *Groundwork III* is titled ‘The concept of freedom is the key to the explanation of the autonomy of the will’. I have just suggested that Kant reverses the order of dependence here: that the autonomy of the will is, if not the key, at least a key component of the explanation of its freedom. I pursue that suggestion further in this section by taking up the idea of freedom as it appears in *Groundwork III*.

In §§1–2, Kant uses an ‘explication’ (*Erklärung*) of the concept of free will to argue that such a will is ‘under moral laws’. He has applied the cosmological idea of freedom (as spontaneous first causality) to the rational will, and his analysis notes that the resulting concept (of free will) has both a negative and a positive aspect. Negatively, a free will operates ‘independently of alien causes determining it’ (4: 446.9). Call this the ‘Causal Independence Condition’. The context indicates that Kant takes it to mean independence of determination by prior natural causes (independence of *Naturnothwendigkeit*): the activity of a free will is not determined by any prior natural cause or governed by any natural causal law. I take it that this Causal Independence Condition is intended to secure a robust notion of causally open alternatives at the time of action and alternative possibilities of choice that Kant thinks is needed for moral responsibility (see *KrV*, B582ff.; *KpV*, 5: 94–9).

Kant’s ‘explication’ of the concept of free will leads to the positive concept of freedom – the law according to which a free will operates. He continues by introducing a ‘Normative Law’ condition: since a free will is a form of *causality* it must operate according to some law, and since it is a *rational* causal power, the principle governing its operation must be a normative principle. Kant thinks that the Causal Independence Condition and the Normative Law Condition taken together imply the Autonomy Condition – that a free will has autonomy, ‘the property of being a law to itself’. The will must in some sense be the source of its basic normative principle *a priori*. And (as I have noted) the arguments that introduce the first two formulations of the CI show that it is the law that the will, or

practical reason, is to itself. Thus the law according to which a free will operates – its formal principle – is the moral law/categorical imperative. This argument shows that a free will is ‘under moral laws’ and that the power of freedom of choice can be specified in practical terms as the power to act from moral principle.

As many commentators point out, Kant’s analytical argument shows nothing about human beings – in particular it does not show that we stand under moral laws. There should be no doubt that we human beings have a faculty of desire, so the obvious next question is whether our faculty of desire is a free will: is it a rational faculty of desire that includes a capacity for pure volition?⁴ In §4 of *Groundwork* III Kant takes only a first step towards resolving this question. His foundational argument requires that the grounds for ascribing freedom to our will (or faculty of desire) must be grounds for ascribing freedom to the will of any rational being. Such grounds are provided by his famous assertion ‘that we must necessarily lend (*leihen*) to every rational being that has a will also the idea of freedom, under which alone it acts’ (4: 448.9–11).

How does Kant support the claim that any rational being with a will necessarily acts under the idea of freedom (that we must ‘lend’ this idea to any such agent)? Here is the relevant passage in full (key sentences numbered for further discussion):

Now I assert: that we must necessarily lend to every rational being that has a will also the idea of freedom, under which alone it acts. [S1] For in such a being we conceive a reason that is practical, i.e., has causality with regard to its objects. [S2] Now one cannot possibly think of a reason that would self-consciously (*mit ihrem eigenen Bewußtsein*) receive guidance from any other quarter with respect to its judgments, since the subject would then attribute the determination of judgment not to his reason, but to an impulse. [S3] Reason must view herself (*sich . . . ansehen*) as the author (*als Urheberin*) of her principles, independently of alien influences, and [S4] must consequently, as practical reason, or as the will of a rational being, by herself be viewed as free; i.e., its will can be a will of its own only under the idea of freedom . . . (4: 448)

Kant claims first (S1) that ‘in such a being we conceive a *reason* that is practical, i.e., has causality with regard to its objects’ (my italics). That is, rational volition is *reason* applied to *action*. He then asserts that reason

proceeds under the idea of freedom (S2–3). Thus practical reason, or the will of a rational being, must regard itself as free because it is a general feature of all reasoning and judgement that it proceeds under the idea of freedom.

Let me make three extended comments.

First, the grounds for ascribing freedom to the will of any rational being come from perfectly general features of reasoning and judgement that are not specific to practical reason. Practical reason or rational volition inherits whatever freedom it has from general features of reasoning and judgement. Dieter Henrich calls the latter ‘logical freedom’ or ‘freedom of judgement’ (Henrich 1998: 311–16). This idea of freedom of judgement applies to any form of rational judgement – logical inference, judgement about spatiotemporal objects and the application of empirical concepts, inference from empirical evidence to explanatory principles, the systematization of a body of empirical laws, as well as to the familiar range of practical judgements (instrumental, prudential, moral, etc.). If so, the ‘spontaneity’ of rational volition is just a special case of the spontaneity of judgement in general.

Second, the idea of freedom of judgement articulated here has both a negative and a positive component, just like the concept of freedom of the will found in §§1–2. S2 expresses an Independence Condition on rational judgement: it is part of the self-consciousness of rational judgement that it does not receive direction ‘from elsewhere’ (*anderwärts her*). I take Kant to be saying that it is part of the self-understanding of rational judgement that it is not determined by anything external to its application of the relevant normative standards, whatever they may be. Rational judgement is of course responsive to sensible or externally given material (to sensible intuition, to empirical fact, to inclination or empirically given ends, etc.), but the normative force to be given to such input is a function of the relevant rational standards. Further, rational judgements can be in error. Judgements that are mistaken or based on faulty inference, judgements that improperly weigh relevant factors, judgements based on bias or irrelevant (even non-rational) factors, and so on, are still judgements if the error falls within a certain range. So it is part of the Independence Condition on judgement that, in judging, one takes oneself *not* to be mistaken or influenced by irrelevant factors. But that is to say that one takes oneself to be correctly following the applicable rational standards.

One further point about this Independence Condition: it is not clear that this Independence Condition on judgement is identical with the Causal Independence Condition on free will seen in §2. It is often assumed that Kant's idea of freedom includes causally open alternatives at the time of choice or judgement. For example, philosophers sometimes expect Kant to hold that it is a presupposition of practical deliberation that when we are deciding how to act, we take it that *what* we will decide is undetermined until we have reached a conclusion. This presupposition is often extended to deliberation and judgement about anything, including what to think about some (theoretical) question. But Kant does not clearly make such a claim at this point in the *Groundwork*, and it is an open question whether the idea of freedom of judgement, as articulated here, does presume causally open alternatives.⁵ That is a point to which I will return (section 4). For now, observe that the immediate focus in the fourth paragraph is that judgement is not determined by anything other than the subject's application of the relevant normative standards. Now Kant's conception of the independence of judgement does exclude a certain kind of empirical causation: a representational state that was directly caused by an impulse (an inclination, sensible intuition, or indeed any prior representational state) in accordance with some natural law would not be a judgement properly speaking. Some independence from natural causal determination is part of the idea of freedom, and I would characterize it, minimally, as follows: a judgement is determined other than causally (other than by some natural causal or psychological law). Judgement understands itself to be normatively rather than causally determined in that the connection of a judgement to its supporting grounds is normative rather than causal. A judgement cannot be understood in its character as rational to be the result of a natural causal process. Further, the idea of judgement requires an active subject who moves from some ground to the judgement through his or her grasp of a relation of rational support that holds between them. So we should include in the Independence Condition that judgement is determined other than causally – that it is a spontaneous act that is normatively rather than causally determined. But it is not clear to me that this addendum introduces causally open alternatives at the time of choice, the claim that judgement is causally undetermined.

Turning to the third sentence (S₃), I suggest that the occurrence of *Urheberin* – that 'reason is the author of her own principles, independently of alien influences' – points to Kant's conception of autonomy. As I read this sentence, it states the positive component of the idea of freedom of judgement, analogous to the positive conception of freedom

of will. Moreover, it states an Autonomy Condition on reasoning and judgement that parallels the autonomy of the rational will (at 4: 440, 447). I interpret the notion of authorship here through the idea that the basic norms of a domain of reasoning or cognition are the formal principles by which it is constituted. These principles ‘arise from’ or are ‘given by’ the nature of that rational or cognitive power, and they normatively govern a subject who understands herself to be engaging in that form of cognition. ‘Independently of alien influence’ here means: based only on the nature of the (relevant) form of cognition *a priori*. Just as the will is a law to itself – the nature of the will is the source of its fundamental formal norms – so reason more generally (a domain of reasoning) is the ‘author’ of the basic normative principles by which it is guided, and the principles authored by a form of reasoning are the *a priori* formal (constitutive) principles that serve as its fundamental norms.

My suggestion, then, is that the positive component of the idea of freedom of judgement is that any domain of reasoning and judgement is constitutively guided by its own formal principles, and in that sense has a form of autonomy. If I am right about this, autonomy makes a second appearance at this point in *Groundwork* III as a component of the idea of freedom of judgement. Furthermore, part of the explanation of why the idea under which all reasoning proceeds is the idea of *freedom* (of judgement) is that reason in general has autonomy. (Again, autonomy is a necessary condition of a rational power being free or genuinely self-determining.) Everyone presumably agrees that the Independence Condition is a general feature of reasoning and judgement. Is it also a general feature of reasoning that any domain of cognition is guided by its own formal norms (and in that sense has a form of autonomy)? I am inclined to think that Kant accepts this idea, but since it is a large thesis that needs much more development than I can provide here, I offer it here somewhat speculatively.

A third comment: various commentators have suggested that this idea of freedom (of judgement) is not a metaphysical but a normative notion. Henry Allison says that this idea ‘regulates the conception of ourselves as rational agents, that is as centers of thought and action’, and as such has normative force. In acting or judging under the idea of freedom we take ourselves to be subject to rational norms (of belief and action).⁶ David Sussman says that it is not a metaphysical presupposition, but rather the regulative principle of making up one’s mind on grounds that would be accepted by a genuinely transcendently free agent, namely, in ways that are independent of purely causal and psychological processes,

or on grounds available to rational subjects generally. Simplifying, it amounts to a commitment to form one's judgement on normatively sound grounds (Sussman 2008: 55–6)

I agree that Kant's idea of freedom is normative rather than metaphysical. The idea of freedom of judgement contains claims about how a judging subject must regard herself and her cognitive activity. It is part of the self-consciousness of judgement that one takes oneself to be basing one's judgement on the sound application of the relevant normative standards, without being influenced by irrelevant or non-rational factors. The self-consciousness of judgement thus contains a *commitment* to sound or correct judgements, or as Sussman says, a commitment to the universal communicability of the reasons on which one bases one's judgement.⁷

3. What is the Normative Content of Kant's Idea of Freedom?

The above comments lead to the question: what is the normative content of the idea of freedom of judgement introduced in §4 of *Groundwork* III, and how does it capture a form of self-determination? Why, in other words, is it an idea of *freedom*?

We should note a detail of Kant's phrasing. He says that reason does not 'self-consciously' (*mit ihrem eigenen Bewußtsein*) receive external guidance and 'must view herself' (*sich ... ansehen*) as the author of her own principles. Both components of the idea of freedom include reference to the way in which a reasoner must understand her activity. So when we put the entire passage together, the general feature of reasoning and judgement that goes into the idea of freedom of judgement is this: it is part of the self-consciousness or self-understanding of rational judgement that (a) it is determined by the correct application of the formal principles of the relevant domain of reasoning (the Autonomy Condition); and (b) judgement is determined *only* by the correct application of these principles and not by anything external to their application. External stimuli and inputs to judgement are to determine judgement only in ways specified by the relevant formal norms (the Independence Condition). That means that reasoning and judgement are driven solely by the application of their own formal norms without the need for 'external influence' to move them forward. Before I can make a judgement about belief or action, I may need some sensible input or some sensibly based material to work with. But when e.g. I move from some sensible input to belief – when the input provides grounds for accepting some belief – in forming my judgement I am guided by the relation of rational support that holds between the input and the belief, and that is a function of the applicable

normative standards. Judging is thus guided by its own normative standards. A judgement can, of course, be incorrect or in error – so the important take-away here is that it is part of the idea of freedom (of judgement) that a judgement understands itself to be correct (to correctly apply the relevant normative standards), and thus is normatively committed to this standard of correctness.

Why is acting under this idea of freedom a form of self-determination? I have suggested that reason is a self-determining faculty in the sense that it proceeds according to its own formal norms, independently of outside influence. But this very abstract characterization of the operation of reason as a faculty (*Vermögen*) needs to be brought down to the level of individual reasoners: in what sense are agents who reason under this idea of freedom (of judgement) as such self-determining? Bringing in the agent's self-conception or self-consciousness helps to complete the picture. In reasoning and judgement, one is self-determining because one is normatively guided by one's own self-consciousness in the following sense: one's conception of what one is doing provides a normative standard by which one guides one's rational or cognitive activity. Rational activity is self-conscious in that a subject who is in the business of making a certain kind of judgement (a logical inference, determining what to believe about some matter of fact, deciding how to act) is self-aware that one is so judging.⁸ Call this 'the self-consciousness of rational judgement'. This self-consciousness includes some (perhaps tacit) understanding of the formal norms that define the relevant form of judgement, as well as an awareness that one is judging in this way, which one does by applying these norms. This self-consciousness thus gives the relevant formal norms a normative hold on the subject that guides his or her cognitive activity. In judging, one is self-aware that one is judging and is normatively committed to judging correctly. (In this respect the formal norms both describe and normatively regulate one's cognitive activity.) Perhaps we can partially unpack the 'spontaneity' of judgement through this normative guidance by one's self-consciousness. And such normative guidance is recognizably a form of self-determination.

How does this idea of freedom of judgement apply to rational volition? As I have said, Kant claims in §4 of *Groundwork* III that the grounds for ascribing freedom to the rational will come from the perfectly general features of reasoning and judgement discussed so far. At 4: 412, he describes volition as the capacity to 'derive actions from laws' or principles. This passage and others suggest that rational volition begins from an agent's ends or principles, and moves from there to a representation of an

action that, in the agent's circumstances, will achieve the end or is called for by the principle. Kantian maxims, as I understand them, are intended to capture the practical reasoning and judgement that guide volition, and thus represent an action as rationally supported by facts about an agent's ends, principles and circumstances. Volition is the capacity to move from end or rational principle to action through an agent's representational capacities by judging the action to be rationally supported or good (where that judgement is captured in the agent's maxim).

If rational volition inherits its freedom from the freedom of reasoning and judgement, then the 'freedom' of rational volition resides in its self-understanding that its activity (of deriving actions from ends and principles) is determined by one's sound application of the formal norms of volition and *only* by their sound application, whatever these norms may be; further, it is a function of the way in which volition is normatively guided by this self-understanding. The idea of freedom of judgement is not sufficient by itself to specify the formal principles of volition, but Kant has argued in §2 of *Groundwork* III that the Moral Law (ML) is the formal principle of free agency. That suggests that for a rational agent to act or will under the idea of freedom is to understand one's activity to be guided by the ML and to be normatively committed to this principle in some form. I think that this leads to what I have discussed elsewhere as Kant's 'rationalist conception of the will'.⁹ According to this thesis, rational volition begins from ends and principles taken to be good and it understands itself to move *correctly* or *in a warranted way* from that starting point to action. That is to say that it acts on maxims that are taken to satisfy a condition of universal validity.¹⁰ (That is what it is for the ML to be the formal principle of free agency.)

If these arguments hang together – and there are a frightening number of 'ifs' here – they suggest a connection between the necessity of acting under the idea of freedom and a certain 'guise of the good' thesis. The idea of freedom is a necessary structural feature of rational self-consciousness. The suggestion is that to act or will under that idea is to understand oneself to be acting on maxims that satisfy a condition of universal validity, and to be normatively guided by that self-understanding.

4. Does the Idea of Freedom have any Metaphysical Implications?

This section digresses briefly on whether the idea of freedom of judgement has any 'metaphysical implications'. For instance, does it imply that in deliberation what one will judge is causally open in advance – undetermined

by any temporally prior cause? It may, but it is not clear to me that it does; I am unsure what to think.

I have argued that the idea of freedom (of judgement) should be understood as follows: that it is part of the self-consciousness of rational judgement that it is determined by the correct application of its own formal principles and not by anything external to these normative standards. It is part of this idea, and arguably a presupposition of deliberation, that we can, in some sense of ability, form a judgement that is valid relative to our existing principles, ends, beliefs, information, etc. (Note that these prior principles and beliefs can be in error or defective, etc.) However, must we assume that our judgements are undetermined in advance of our judging? In a case where you judge incorrectly (relative to your prior ends and beliefs, etc.), or where your judgement was affected by bias, perhaps you presuppose that in those same circumstances you could have assessed the available evidence correctly, or that you could have become aware of and set the bias aside. This assumption could be understood as an analogue to alternate possibilities of choice in the practical sphere. But it is not clear that we must make that presupposition, or if we do, that it need be understood in terms of open possibilities at the time of judgement.¹¹ We are normatively committed to avoiding such errors, and to revising our judgement when we see that it was in error, but that may be all.

My earlier suggestion (section 2) was that the Independence Condition on judgement requires that judgement be determined other than causally – that it be normatively rather than causally governed. This assumption may allow that there are (empirical) causal processes underlying judgement at some level, or that given sufficient information about a person's mental states and dispositions, empirical generalizations of some kind may predict how a person will judge in some situation. Indeed, Kant thinks that human activity viewed as phenomenon is causally determined in this way, and the same could apply to a person's judgements. (Cf. *KpV*, 5: 99: given enough information about a 'human being's cast of mind . . . we could calculate a human being's conduct for the future with as much certainty as a lunar or solar eclipse'.)

What is clear, however, is that we cannot understand our judgements in their character as rational if we view them as the results of a causal process governed simply by empirical laws or psychological principles. A naturalistic explanation of a judgement leaves out the spontaneity that makes it a rational judgement – the act of the subject moving from supporting ground to judgement through a grasp of the relation of rational

support – and cannot capture its character as normatively governed. The phrase ‘governed other than causally’ is an attempt to capture this idea, without going so far as to assert that our judgements are causally undetermined in advance.

These reflections raise the question whether the Causal Independence Condition in the concept of free will (G III, §§ 1–2) and the Independence Condition in the idea of freedom of judgement (§ 4) are equivalent, as Kant appears to assume? To explain this issue: §§ 1–4 aim to establish in two basic steps that any rational being with a will stands under moral laws. As we saw in section 2, § 4 uses analysis of the self-consciousness of rational judgement to argue that a rational being with a will necessarily acts under the idea of freedom (and that a being that necessarily acts under the idea of freedom is ‘free in a practical respect’ (4: 448.6) and subject to any principles that hold for an agent whose freedom is established on theoretical grounds).¹² §§ 1–2 use an ‘explication’ of the concept of free will to argue that a free will stands under moral laws. Thus:

- (1) A rational being with a will necessarily acts under the idea of freedom (and is ‘free in a practical respect’, etc.) (§ 4)
- (2) A free will stands under moral laws. (§§ 1–2)
- (3) So a rational being with a will stands under moral laws (etc.).

For the opening four paragraphs to achieve their aim, the conception of freedom in (1) and (2) must be the same. Furthermore, the support for ascribing the concept of freedom (in §§ 1–2) to the will must come from features of the idea of freedom of judgement (since Kant argues that rational volition inherits its freedom from the freedom of judgement). That is, applying the idea of freedom of judgement to the rational will should lead to the idea of a free will whose causality is independent of ‘alien determining causes’.

I have assumed that Kant thinks that the Causal Independence Condition (‘negative concept’) of free will contains a robust notion of causally open alternatives. But is that part of the idea of freedom of judgement? I find this unclear, and the possibility that it is not reveals another potential fissure in the argument of *Groundwork* III that would complicate the transition from (1) to (2). Here there seem to be two routes that Kant, or a defender of his argument, could take. One is to accept that the idea of causally open alternatives is built into the concept of free will and to argue that it is secured for the will through the idea of freedom of

judgement. For example, either the idea of freedom of judgement includes metaphysically open alternatives, or when this idea is applied to the will it introduces alternative possibilities of action at the time of choice. The other route is to hold that the idea of causally open alternatives at the time of action is not part of the idea of freedom of judgement, but that it is not necessary for the concept of free will. Rather, what is needed for freedom of will (and the argument of §2) is a robust notion of a self-determining will, and the requisite notion of self-determination is secured by the idea of freedom of judgement. (Here see section 3 above: reason is a self-determining faculty because its exercise is guided by the self-conscious application of its own formal norms, independently of outside influence.)

This is a vexing issue that goes beyond the scope of this article.¹³

5. How does the Idea of Freedom Figure in the Balance of the Preliminary Arguments of *Groundwork III*?

Let me now return to the preliminary arguments of *Groundwork III* that set up the ‘deduction’ of the CI (at 4: 453.17ff.), and to some questions that they raise.

As we have seen, in §4 Kant asserts that a rational being with a will necessarily acts under the idea of freedom because of general features of reasoning and judgement. Reasoning and judgement in general necessarily proceed under the idea of freedom, and since the will is practical reason, volition inherits the freedom of judgement. This seems like a good argument for asserting that an idea of freedom is a necessary feature of the practical self-consciousness of a rational being with a will. Kant does not refer back to this specific move later in *Groundwork III*, but there is no indication that he has any reservations about it. I have also tried to unpack the sense in which activity that proceeds under this idea of freedom is a genuine form of self-determination.

Now it is also generally agreed that the material in §4 does not warrant the ascription of freedom to human beings. Kant does not assert that *human beings* necessarily act under the idea of freedom until §15 (4: 452.31–3), after he has shown that human beings must view themselves ‘as intelligences’, as members of an intelligible world (450.30–453.2).

Why does §4 not show that human beings necessarily act under the idea of freedom? One standard answer is that nothing so far establishes that human beings have wills in the required sense: ‘a capacity distinct from a

mere desiderative faculty (namely to determine itself to action as an intelligence, hence according to laws of reason, independently of natural instincts)' (4: 459.11–13; cf. 457.25–9). As previously noted, Dieter Henrich and others point out that Kant cannot simply infer the existence of 'transcendental freedom' from 'logical freedom'. It is clear that we have a faculty of desire, and whether we have a capacity for theoretical judgement is not in doubt. So the point here is that from the fact that we have understanding and the capacity for theoretical judgement, we cannot assume that our faculty of desire is a rational will with a capacity for pure volition. Allison suggests that Kant takes seriously the possibility that we are creatures governed by instinct and that our agential capacities are illusory (Allison 2011: 239). The fact that we have understanding does not rule out the possibility that our sense of determining our actions through rational choice is epiphenomenal. According to Allison, Kant sees the need to dispel this possibility.

Kant takes on this unresolved question about the ascription of freedom to human beings in the paragraphs in which he escapes from the 'circle'. Depending on how one understands it, escape from the circle requires either some *grounds* for ascribing freedom to our faculty of desire (other than the fact that we ordinarily assume that we have free agency – 4: 455.11), or some *grounds* for ascribing freedom to ourselves that are *independent of moral consciousness* (our tendency to recognize the authority of morality).¹⁴ Since Kant provides such grounds at 450.30–453.2 (§§10–15), the next question is what these paragraphs add to §4 that warrants the ascription of free agency to human beings.

There are two important moves in the intervening paragraphs. First, drawing the distinction between the sensible and the intelligible worlds gives us two standpoints from which to consider human action, and thus creates space for thinking of human activity as free and normatively governed. Arguably we need the standpoint of an intelligible world in order to capture the norm-guided spontaneity of rational activity, including the judgemental activity of understanding.¹⁵ Second, Kant appeals to the 'pure self-activity' displayed by theoretical reason that 'is elevated even above the understanding' to give us human beings membership in the intelligible world. That he is referring to theoretical reason is shown by the fact that he distinguishes reason from understanding and by the pressure coming from the threat of the circle to find grounds for ascribing freedom to our faculty of desire that do not depend on moral consciousness. The pure spontaneity of theoretical reason is displayed in our ability as theoretical reasoners to draw the distinction between the sensible and

intelligible worlds that sets the limits for the use of the understanding. Further it includes the capacity of reason to generate the regulative ideas that guide the employment of the understanding. If so, this is a third point in *Groundwork* III at which one can discern autonomy playing a role. Reason has the formal aim of introducing systematic unity and completeness into a body of given material (here the laws of nature discovered by the understanding), and this ideal of systematicity generates various regulative principles that guide the extension of the empirical use of the understanding. Reason, again, in its theoretical employment is a law to itself (in the sense that its nature, or a conception of what theoretical reason does, is the source of the fundamental normative principles that guide its exercise).

The upshot is that the ‘pure self-activity’ or ‘pure spontaneity’ of theoretical reason gives us, as reasoners, membership in the intelligible world – it shows that we possess certain a kind of normative capacity – and thus as ‘intelligences’ we necessarily act under the idea of freedom:¹⁶

As a rational being, hence as one that belongs to the intelligible world, a human being can never think of the causality of his own will otherwise than under the idea of freedom. (4: 452.3 I–2)¹⁷

How does the spontaneity of theoretical reason resolve the question of whether our faculty of desire is a free rational will? I suggest that citing the spontaneity of theoretical reason brings out structural parallels with pure practical reason – that theoretical reason is a normative capacity with the same features as those that would have to be in place in a pure practical reasoner. The ‘pure spontaneity’ of theoretical reason makes us members of the intelligible world by showing that we have normative capacities with a certain self-determining structure. Since we have this capacity for pure theoretical reasoning, we may conclude that our faculty of desire is likewise a rational capacity with comparable self-determining features, including a capacity for pure volition. Let me develop this possible reconstruction.

Step 1

It is part of the ‘pure spontaneity’ of theoretical reason that it generates the ideal of a systematic whole of empirical inquiry on its own, *a priori* and independently of any sensible input, and this ideal is the source of normative principles that are sufficient to guide its employment to empirical inquiry. Theoretical reason thus satisfies the Independence Condition seen in §4, but also an Autonomy Condition – it is ‘the author of its own

principles, independently of alien influence'. If I am right, understanding also satisfies an Autonomy Condition, since it is a cognitive power constituted by its own formal principles, and Kant says that (unlike sensibility) it 'is also self-activity'. However, 'it can produce from its activity no other concepts than those which serve merely to bring sensuous representations under rules and thereby unite them in one consciousness' (4: 452.10–17). Understanding generates empirical concepts that combine given intuitions in one consciousness, but it can think nothing substantive on its own without input from sensible intuition. The autonomy of reason (versus that of understanding) is distinguished by the fact that it generates regulative principles and ends with enough substance to guide its own exercise. It thus shows a degree of self-determination that goes beyond that of the understanding.¹⁸

Our capacity for theoretical reason gives us membership in the intelligible world. It is a normative capacity that is the source of ideas and ends of inquiry that go beyond anything given either by sensibility or understanding, and these regulative ideas and ends are sufficient to guide the exercise of reason (in application to material provided by the empirical use of understanding). That is to say that it is a normative capacity with the features comparable to pure practical reason.

Step 2

That we have a theoretical power with these features supports the supposition that our faculty of desire is a rational will with corresponding features – it is a law to itself, its law is sufficient to guide its own exercise and to determine conduct, and so on, independently of input from sensibility, and is thus a capacity for genuine self-determination. Otherwise put, our faculty of desire includes pure practical reason.

From here I see two ways to continue the argument. One is that having now established that we may regard our faculty of desire as a rational will, the argument of §4 applies directly, and we must indeed *act* under the idea of freedom (of judgement). (This way of continuing uses what §4 establishes about rational agents in general, but makes no use of the difference between understanding and reasoning.) A second possibility would apply the earlier argument from §4 at a somewhat higher level, appealing to the fact that theoretical reason displays a kind of spontaneity that goes beyond anything seen in the understanding. The specific form of spontaneity seen in theoretical reason transfers to our practical faculties – our will – and we necessarily act under that idea of freedom. Either way, we would have some independent warrant for supposing a

capacity for free agency that does not depend directly on ordinary moral consciousness.

Is this a good argument for the claim that we human beings necessarily act under the idea of freedom? It is not a bad argument, but it strikes me as inconclusive. Beyond that I am not sure what to say. So let me close with some questions that I will not attempt to answer, but which point to reasons for being unsatisfied with the argument.

First recall Henrich's point (Henrich 1998: 311–16) that we cannot infer transcendental from logical freedom (of judgement) – that from the fact that we have the capacity for theoretical judgement, we cannot assume that our faculty of desire is a rational will with a capacity for pure volition. Does the argument that I have sketched address this concern? Could one raise a similar worry: that from the fact that we have theoretical reason (which shows 'pure self-activity') we cannot infer that we have a will with a pure component? Perhaps the force of this worry is blunted once we unpack the spontaneity of theoretical reason – but perhaps it is not.¹⁹

Second, why is not the capacity for understanding sufficient to get us into the intelligible world and to support the ascription of freedom to our faculty of desire? (That is, if our theoretical capacities support membership in the intelligible world, why is it not enough to appeal to understanding, rather than going all the way to reason?) The understanding is a normative capacity that satisfies the Independence Condition and an Autonomy Condition, and it displays spontaneity and a robust notion of self-determination. It has all the features that go into the idea of freedom. But Kant appeals to our capacity for theoretical reason to get us into the intelligible world. One possible response to this question is that we need reason to get into the intelligible world because we need reason to conceive of the intelligible world. Since understanding cannot form this conception on its own, it cannot by itself warrant the assumption that we have normative capacities that make us members of an intelligible world.

Third, reason appears to display a form of self-determination that goes beyond that of understanding. Among other things, it sets its own ends in ways that understanding does not. Is this an addition to the idea of freedom as laid out in §4 – and one that is needed to support the ascription of freedom to our wills? Or is the idea of freedom basically the same throughout *Groundwork* III? Perhaps the 'pure spontaneity' of reason

adds features to the earlier idea of freedom that are needed to support the idea that we have the kind of practical capacities that are presupposed by morality.

6. What Further Normative Commitments are Built into the Idea of Freedom of Action?

My focus so far has been on a very basic normative commitment that is part of all reasoning and judgement. I have argued that to proceed under the idea of freedom is to take one's judgement to be determined by one's correct application of the formal norms of the relevant domain of cognition and not by anything external. I have also suggested (though not fully argued) that extending this idea of freedom to rational volition leads to the idea that rational volition understands itself to be guided by maxims that satisfy a condition of universal validity. But the idea of freedom so understood also contains a richer set of normative commitments specific to the self-consciousness of rational volition, with the potential to play a role in Kant's normative conception. I consider this thought briefly in closing.

To *act* under the idea of freedom is to understand oneself to be moving correctly from ends or principles to action through a sound judgement of what is rationally supported in one's circumstances. I have been stressing the condition of sound judgement, but this judgement is necessarily the *agent's own* and an act of spontaneity. Thus it is part of the idea of freedom *of action* that in rational volition one understands one's judgement of what is rationally supported to be determined by *one's own* (sound) application of the relevant normative standards and by nothing external to one's own application of these standards. Further, it is part of the idea of freedom of action that one is to *guide one's use of one's causal powers* by one's own judgement of what is rationally supported. These features of the practical self-consciousness of rational volition – of our self-conception as agents – lead to a set of individual normative commitments specific to rational agency: for example, that in rational volition one is committed to forming one's own judgement of what is rationally supported without being determined by anything external to one's judgement and to deploying one's own causality to realize the action and any purposes that one has judged to have rational support. In addition, we might think that the idea of freedom (one's self-conception as an agent) leads to a commitment to the general material conditions needed to support the successful exercise of one's (finite) agency.

Any such necessary commitments would provide input to the categorical imperative. First, they set constraints on universalizability: universalization of any maxim must be consistent with any such individual commitments that are part of the idea of freedom. If the universalization of some maxim conflicts with these commitments, the maxim cannot be willed as universal law without contradiction. Second, these individual commitments can figure in the ‘direct route’ to substantive moral principles suggested by Stephen Engstrom.²⁰ Applying the condition of universality to what one necessarily wills for oneself leads straightforwardly to some general moral principles. Either way, necessary commitments that are part of the idea of freedom of action are available as an *a priori* source of (substantive) input to moral reasoning.

Let me develop this suggestion very briefly. It is part of one’s self-conception in acting that one forms a judgement about what is rationally supported in one’s circumstances and that the resulting judgement of an action or end as rationally supported is to guide one’s causality. This feature of practical self-consciousness leads to two very general commitments. The first is a commitment to forming your own judgement of what is rationally supported in your circumstances, independently of influence external to your judgement. That does not mean that nothing external to your judgement bears on what you have reason to do, since many independent facts are obviously relevant (including facts about your abilities, alternatives and circumstances, facts about the intentions and beliefs of others, etc.).²¹ Rather, it means that the rational weight that any such factors have is a function of the relevant normative standards, and that they are to influence your judgement through your application of these standards. Call this a commitment to the *independence of judgement*. Second, since agency involves *acting* on your judgement of what is rationally supported, it involves a commitment to realizing the action or purpose that you represent as rationally supported through your own causality. Call this a commitment to the *independence of action*. These two general commitments appear to be formal components of what one wills whenever one acts on a maxim, which is to say formal components of willed action.

You can probably see where I am going. Maxims that involve interfering with or trying to control another person’s judgement – maxims of deception, manipulation, threat, coercion and so on, for reasons of self-interest – cannot be willed as universal law consistently with the commitment to the independence of judgement. Likewise maxims that involve interference with the freedom of action of others (for reasons of self-interest) cannot be willed as universal law consistently with the

commitment to the independence of action. Grant that I am committed to the independence of my own judgement and action as part of what I will any time I act on a maxim. Maxims of deception or manipulation, and maxims that involve interfering with the freedom of others – all for reasons of self-interest – conflict with these commitments when willed universally. Therefore one cannot consistently will such maxims and their universalization at the same time.²² Alternatively, to take the direct route to moral principles: applying the condition of universality to the commitments to the independence of judgement and of action (what I necessarily will for myself I am to will for everyone) leads to standard Kantian prohibitions against attempts to control the judgement of others or to interfere with their freedom of action (for reasons of self-interest).

This line of reasoning is a route to duties of justice that classify as impermissible maxims that involve direct interference with the agency of another. To extend this proposal to the duties of virtue, one needs to argue that it is part of our self-conception as finite agents that we are committed to wanting the general conditions of human life needed to support the continuing and successful exercise of one's agency. I assume that this can be done, but will not pursue it here.²³

Since there is no detail here, this is just a proposal: that we look to Kant's idea of freedom of judgement and action as an *a priori* source of input to his normative conception. The target here is a set of *individual* commitments that are plausibly seen as *necessary features* of our self-conception as free agents. Such commitments point to a core set of moral principles to be further specified by bringing in general facts about human beings and human social life, and which require judgement and understanding for their application.²⁴

Notes

- 1 Citations of Kant are, with the exception of *KrV*, to the Akademie edition, using the following abbreviations and translations (with occasional modifications): *G* = *Groundwork of the Metaphysics of Morals* (Kant 2011, citations sometimes including line numbers); *KpV* = *Critique of Practical Reason* (Kant 2015); *KrV* = *Critique of Pure Reason* (Kant 1998, using standard A/B pagination); *MdS* = *The Metaphysics of Morals* (Kant 1996a). Page references are included in parentheses in the body of the article where practical.
- 2 Kant identifies the will with practical reason (among other places) at *G*, 4: 412.29–30, 448:18–19; and *MdS*, 6: 212.
- 3 I develop the reading of autonomy of the will in these paragraphs in Reath 2013 and especially Reath 2019.
- 4 A point often stressed by Allison. See e.g. Allison 2011: 309, 324–30.
- 5 Here cf. Allison 2012: 88ff. Remarks in Kant's 'Review of Schulz' bear on this point (8: 13–14; included in Kant 1996b). Here Kant argues that 'the most confirmed fatalist'

must 'always act as if he were free', where 'free' appears to mean that his choice is causally undetermined. But Kant also claims that the fatalist necessarily assumes a 'freedom to think' characterized as follows: 'he has assumed in the depths of his soul that understanding is able to determine his judgment in accordance with objective grounds that are always valid and is not subject to the mechanism of merely subjectively determining causes, which could subsequently change; hence he always admits freedom to think, without which there is no reason' (8: 14). That Kant here ties 'freedom to think' to the ability to 'determine one's judgment in accordance with objective grounds that are always valid' supports that claim that it is part of the idea of freedom of judgement that the subject takes herself to be correctly following applicable rational norms.

- 6 Allison 2011: 306, where he writes: 'To act in conformity [with the idea of freedom] is to place oneself in the logical space of (practical) reasons, thereby subjecting oneself to norms of both a moral and a prudential sort.' The idea of deliberation supposes that one is operating in the space of reasons – that one has the capacity to determine one's judgement through one's assessment of the relevant reasons and is subject to various norms of thought and action. Cf. also Allison 2012: 90–2, 113.
- 7 Clearly there is more to say about the specific normative implications of this idea. But to give one simple kind of example, if you realize that a judgement that you have made is based on bias or some irrelevant factor, you are committed to countering that source of bias and revising the judgement.
- 8 Here see Pippin 1987: section IV.
- 9 See Reath 2015. Similar conceptions of rational volition are developed by Herman 2007 (chs 10–11) and by Engstrom 2009 (ch. 4) and 2010.
- 10 I leave this 'condition of universal validity' undefined to allow for the possibility that bad willing can take itself to be 'universally valid', or to allow that an agent could operate with a defective conception of what it is for a maxim to be universally valid. For further discussion, see Reath 2015, especially 238–40, 247–8, 251–2.
- 11 The 'could' in 'could have assessed the evidence correctly' might refer to one's having certain general capacities, not to how one applies them in this case. If so, it does not require causal indeterminism. Thanks to an anonymous referee for this journal, whose suggestion I adopt.
- 12 Since it does not affect the immediate point at hand, I set aside this detail, explained further in *G*, 4: 445n: that 'the same laws that would bind a being that was actually free yet hold for a being that cannot act otherwise than under the idea of freedom'. For discussion, See Hill 1992: 116–20.
- 13 But for the record, the second route appears sufficient for Kant's purposes. First, as explained in section 3 above, the idea of freedom of judgement contains a notion of self-determination that is a general feature of all reasoning – and that does not clearly involve metaphysical commitments. Reasoning and judgement in some domain of cognition understand themselves (a) to be guided by the correct application of the formal principles of that domain of reasoning that satisfy an Autonomy Condition; and (b) to be guided only by the application of these principles. (b) is the Independence Condition on judgment. But second, (b) can capture the Independence Condition on free will, that it operates 'independently of alien causes determining it'. For it implies that the relevant rational power is normatively rather than causally governed (it is governed other than causally), and if so it is part of its self-consciousness that it is not determined by anything external to the application of these principles. Finally, this rendition of the Independence Condition is enough to launch a reconstruction of the argument of *Groundwork* III, §2: (1) As a rational causal power, a free will is governed other than causally and is not determined in its operation by external causes. (Independence

- Condition). (2) So it is governed by normative principles. (3) These normative principles cannot be external but must be ‘internal’ to the power – that is, they are formal principles based in the nature of the power that satisfy an Autonomy Condition. (4) Therefore a free will has autonomy, the ‘property of being a law to itself’. (5) Earlier arguments of *Groundwork* II have shown that the CI (or ML) is that law that the will, or practical reason, is to itself. (6) Therefore a free will is under moral laws.
- 14 For discussion, see Quarfoot 2006, Schönecker 2006, and Allison 2011: 309–16.
 - 15 I am inclined to think that at root membership in the intelligible world amounts to the possession of normative capacities, including a capacity for spontaneity or normative self-guidance, that cannot be understood naturalistically – e.g. capacities whose operation cannot be reductively explained in terms of any natural causal process.
 - 16 It is clear to me that in these paragraphs, Kant claims that general features of reason make us ‘intelligences’ and that we necessarily view our wills as free (that we think of our faculty of desire as a rational will with a capacity for pure volition, hence as a free will) because of our membership in an intelligible world. That is, we must first get into the intelligible world before we can view our wills as free. (See also *G*, 4: 453.16–18; 457.9–13, 22–3; and 457.29–30.) But there are also passages where Kant says that it is by thinking of ourselves as free that we transfer ourselves into the intelligible world – e.g. 453.12, 454.7, 454.25ff. This seems to me a *prima facie* different argument. Perhaps these two strains can be combined; but perhaps not.
 - 17 Cf. also *G*, 4: 453.17: ‘A rational being counts itself as an intelligence among the world of understanding and, merely as an efficient cause that belongs to it, it calls its causality a will.’
 - 18 Eric Watkins has objected that understanding does not have autonomy, only reason. But I am not claiming that understanding has autonomy in the full sense in which reason does. It is a cognitive power whose nature is determined by its own formal principles, and in that sense it is ‘the author of its own principles, independently of alien influence’. So it shares some features with the autonomy of reason.
 - 19 A referee asks whether it strengthens Kant’s argument as described here by taking into account his distinction between theoretical arguments and arguments made from the practical standpoint; though unconvincing as a theoretical argument, the argument might be sufficient if made from the practical standpoint. To respond, in §4 of *Groundwork* III, Kant asserts that a being ‘that cannot act otherwise than under the Idea of freedom is actually free, *in a practical respect*’, where the italicized phrase means that such a being should acknowledge any normative principles that apply to a being proven free by a valid theoretical argument. Since any rational being with a will necessarily acts under the idea of freedom, such agents should accept that they are free for practical purposes, and acknowledge whatever normative consequences follow. Kant then needs to establish whether we human beings have rational wills, in which case we may ascribe freedom to ourselves ‘in a practical respect’. This raises the question: is the ascription of rational wills made from the theoretical or the practical standpoint? I am inclined to think that it is from the practical standpoint. First, it is not clear that ‘rational will’, as a normative capacity, can be defined or used theoretically. (‘First cause’ can be theoretically defined, but the will is a complex normative capacity.) Second, the ascription of a rational will is supported by reflection on the capacity for theoretical reason, which is likewise a normative capacity that gives us membership in an ‘intelligible world’ – viz. a normative capacity that cannot be explained in terms of any natural causal process. This leads me to think that most of *Groundwork* III operates within the practical standpoint – ‘the standpoint that reason sees itself necessitated to take outside of appearances, in order to think of itself as practical’ (*G*, 4: 458.19–21).

- Back now to the question: it is not clear to me that understanding the argument (from G, 4: 450–3) as made from the practical standpoint addresses e.g. Henrich's worry. It may – however my point is not that the argument fails, but that it is inconclusive.
- 20 Engstrom 2009: 138–40, 216–20. Similar ideas are found in Cureton 2013, who argues that other-regarding moral requirements are grounded in the universalization of rational self-regarding dispositions (such as 'dispositions to be free, to perfect ourselves and even to promote our own happiness', 169).
 - 21 Likewise it does not mean that one may never take advice, rely on information from others or consult their opinions, and so on.
 - 22 The material in this section needs a paper of its own, but let me add a brief comment. Universalization of maxims that interfere with the independence of judgement and action arguably leads to a contradiction in conception, since it is inconsistent with commitments that are formal features of willed action. To exercise your rational agency is to form and act on your own judgement of what you have reason to do, and your self-conception as an agent commits you to these forms of independence. But universalizing a maxim of e.g. deception or interference with another's freedom is to will that agents be free to act in ways that undermine the conditions of willed action. In adopting any maxim, one wills one's own independence of judgement and action, but in universalizing these particular maxims, one wills that agents are generally free to undermine the independence of judgement and action in others, including in oneself. Universalizing this maxim is thus deeply incoherent.
 - 23 For other developments of this approach, see Herman 1993: ch. 6, especially pp. 119–25, and Reath 2006: 211–20.
 - 24 Versions of this article were presented to a Workshop on Kant on Freedom at the University of Groningen, the Netherlands (June 2016), at Williams College (October 2017), at the Central Division of the APA in Kansas City (March 2017), and at a Symposium on 'The Many Faces of Kantian Freedom', University of Tennessee-Knoxville (April 2017). I am grateful to the organizers of and audiences of these events for their comments and questions, and also for several comments by an anonymous referee for *Kantian Review*.

References

- Allison, H. E. (2011) *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Oxford/New York: Oxford University Press.
- (2012). *Essays on Kant*. Oxford/New York: Oxford University Press.
- Cureton, A. (2013) 'From Self-Respect to Respect for Others'. *Pacific Philosophical Quarterly*, 94, 166–87.
- Engstrom, S. (2009) *The Form of Practical Knowledge*. Cambridge, MA: Harvard University Press.
- (2010) 'Reason, Desire and the Will'. In L. Denis (ed.), *Kant's Metaphysics of Morals: A Critical Guide* (New York/Cambridge: Cambridge University Press), pp. 28–50.
- Henrich, D. (1998) 'The Deduction of the Moral Law: The Reasons for the Obscurity of the Final Section of Kant's *Groundwork of the Metaphysics of Morals*'. Trans. P. Guyer. In P. Guyer (ed.), *Groundwork of the Metaphysics of Morals: Critical Essays* (Lanham, MD: Rowman & Littlefield), pp. 303–41. (First publ. in Alexander Schwann (ed.), *Denken im Schatten des Nihilismus*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1975.)
- Herman, B. (1993) *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- (2007) *Moral Literacy*. Cambridge, MA: Harvard University Press.

- Hill, T. E., Jr. (1992) *Dignity and Practical Reason*. Ithaca: Cornell University Press.
- Kant, I. (1996a) *The Metaphysics of Morals*. Trans. and ed. M. Gregor. Cambridge/New York: Cambridge University Press.
- (1996b) *Practical Philosophy*. Trans. and ed. M. Gregor. Cambridge/New York: Cambridge University Press.
- (1998) *Critique of Pure Reason*. Trans. and ed. P. Guyer and A. Wood. Cambridge/New York: Cambridge University Press.
- (2011) *Groundwork of the Metaphysics of Morals*. Trans. and ed. M. Gregor and J. Timmermann. Cambridge/New York: Cambridge University Press.
- (2015) *Critique of Practical Reason* (revised edition). Trans. M. Gregor, ed. Andrews Reath. Cambridge/New York: Cambridge University Press.
- Pippin, R. (1987) 'Kant on the Spontaneity of Mind'. *Canadian Journal of Philosophy*, 17.2, 449–76.
- Quarfoot, M. (2006) 'The Circle and the Two Standpoints (GMS III, 3)'. In C. Horn and D. Schönecker (eds), *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter), pp. 285–300.
- Reath, A. (2006) *Agency and Autonomy in Kant's Moral Theory*. Oxford/New York: Oxford University Press.
- (2013) 'Kant's Conception of Autonomy of the Will.' In O. Sensen (ed.), *Kant on Moral Autonomy* (New York/Cambridge: Cambridge University Press), pp. 32–52.
- (2015) 'Did Kant Hold that Rational Volition is *sub Ratione Boni*?'. In M. Timmons and R. Johnson (eds), *Reason, Value and Respect: Kantian Themes from the Philosophy of Thomas E. Hill, Jr.* (Oxford/New York: Oxford University Press), pp. 232–55.
- (2019) 'What Emerged: Autonomy and Heteronomy in the *Groundwork* and Second *Critique*'. In S. Bacin and O. Sensen (eds), *The Emergence of Autonomy in Kant's Moral Philosophy* (New York/Cambridge: Cambridge University Press), pp. 176–95.
- Schönecker, D. (2006) 'How is a Categorical Imperative Possible? Kant's Deduction of the Categorical Imperative'. In C. Horn and D. Schönecker (eds), *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter), pp. 301–24.
- Sussman, D. (2008) 'From Deduction to Deed: Kant's Grounding of the Moral Law'. *Kantian Review*, 13(1), 52–81.