

## CHAPTER 2

*Formal principles and the form of a law**Andrews Reath*

## I. INTRODUCTION

One aim of the *Critique of Practical Reason* is to establish that reason alone can determine the will. To show that it can, it suffices to show that there are practical principles given by reason alone – what Kant terms ‘practical laws’, or (roughly) requirements of reason on action. Chapter I of the *Analytic* accomplishes this aim by arguing that the moral law is an authoritative practical principle given as a ‘fact of reason’. The chapter begins in section 1 with a ‘Definition’ (*Erklärung*) of a practical law as a practical principle that holds necessarily for every rational being (as a principle that ‘determines the will simply as will’ without presupposing any subjective or rationally contingent interests). From this definition Kant uses conceptual arguments to derive three ‘Theorems’ and various ‘Corollaries’ that state conditions that any practical law must satisfy. The principal condition, stated in Theorem III, is that a practical law determines choice through its form – through ‘the mere form of giving universal law’ – rather than through its matter (*CpV* 5:27). The arguments culminate in section 7 with the statement of the ‘Fundamental Law of Pure Practical Reason’, which Kant identifies with the moral law: ‘So act that the maxim of your will could at the same time always hold as a principle in a giving of universal law’ (*CpV* 5:30). Section 7 includes a formulation of the fundamental law that follows from the Definition and Theorems – just as in the *Groundwork* a formula of the Categorical Imperative is derived first from the analysis of the concept of duty (*G* 4:402), then from the concept of a categorical imperative (*G* 4:421). But since the moral law is put forward here, in imperative form, as an authoritative practical law, section 7 is the synthetic claim that *Groundwork* III attempts to establish. Thus in short order, these pages of the second *Critique* cover the terrain of the extended argument of the *Groundwork*.

The analytic portions of this chapter contain several striking claims about features of practical laws and fundamental practical principles that centre on

the notion of 'form'. Because Kant's foundational works aim to establish that there are moral requirements with genuine rational necessity, the second *Critique* begins with a definition of a practical law. Kant then claims that a fundamental principle of morality is in some sense a formal principle (Theorem II, Corollary, *CpV* 5:22) and that a practical law provides a ground of choice through its form rather than its matter (Theorem III). Moreover, he claims that only a formal principle can be the basis of categorical imperatives that apply with necessity.<sup>1</sup> He also argues for analytical connections between freedom of the will and the form of a law: that a will 'for which the mere lawgiving form of a maxim can alone serve as a law is a free will' and that if a will is free, the lawgiving form of a maxim is 'the only thing that can constitute a determining ground of the will' (Problems I and II, *CpV* 5:28–9). To assess these claims, we need a handle on certain concepts: What does Kant mean by the form of universal lawgiving, or the form of a practical law? What does it mean to think of the lawgiving form of a practical principle as the determining ground of choice? What does he mean by a 'formal principle' and why does he think that fundamental normative principles must be 'formal'? In particular, why does he think that a practical principle is normatively necessary only if it is a formal principle in his sense (to be discussed in section III below) and provides a ground of choice through its form?

The difficulty in unpacking these claims is to some extent a matter of their familiarity. The arguments found in this chapter of the second *Critique* are compact versions of arguments given in the *Groundwork* that Kant, perhaps, took greater pains to develop in the earlier work. These claims, as they appear in the second *Critique*, rely on various distinctions that Kant now takes for granted. For example, Theorem III simply assumes that there is an exhaustive distinction between the form and the matter of a principle, and that a practical principle provides a ground of choice either through its matter (through a contingent interest in its matter) or through its form. Given the claim that a principle that provides a ground of choice through its matter cannot serve as a practical law (convincingly argued in Theorem I), it readily follows that a practical law provides a ground of choice through its form. The arguments of Problems I and II identify the

<sup>1</sup> '[S]ince material principles are quite unfit to be the supreme moral law (as has been proved), the *formal practical principle* of pure reason (in accordance with which the mere form of a possible giving of universal law through our maxims must constitute the supreme and immediate determining ground of the will) is the *sole* principle that can *possibly* be fit for categorical imperatives, that is, practical laws' (*CpV* 5:41). See also *CpV* 5:64: 'Only a formal law, that is, one that prescribes to reason nothing more than the form of its universal lawgiving as the supreme condition of maxims, can be a priori a determining ground of practical reason.'

matter of a principle with items given in appearance and assume that lawgiving form is not found among appearances. Kant accordingly concludes that the form of a law is a ground of choice that is independent of the causal laws that govern appearances, thus that a will that can be determined by the form of law is transcendently free. Because arguments like these are so easily made within Kant's philosophical framework, it is hard to get an independent perspective on the basic ideas at issue.

The form of a practical law (the form of universal lawgiving) is commonly understood as the formal feature of necessary and universal applicability – the universal validity of some practical principle for all rational beings, or the fact that a principle makes a necessary demand on all rational agents that excludes the force of competing reasons for action. Stephen Engstrom has pointed to Kant's distinction between 'objective universal validity' and 'subjective universal validity' to suggest a gloss on this notion. Objective universal validity is the applicability of a principle to all objects that fall under the principle. The objective universal validity of a practical principle is the fact that it applies to or governs the actions of all agents in the circumstances covered by the principle. Subjective universal validity is the validity of a principle for all knowing subjects – the fact that it governs judgment and determines how all subjects should think about the matter, 'so that all such subjects could agree in the matter and share the same judgment'.<sup>2</sup> Engstrom observes that the two kinds of universal validity coincide in the case of practical principles:

The principle applies to the will of every rational being, and every such being can recognize this universal applicability. This is as much as to say that principles of practical cognition are necessarily such that every subject can agree to every subject's acting on them, as would actually happen if all subjects were jointly to legislate this principle for themselves.<sup>3</sup>

According to this suggestion, the form of a practical law is the fact that all subjects can agree to any subject's acting on the principle. To say that the form of a law is a ground of choice is to say that the necessity and universality of a principle (the fact that it is suited to hold as universal law, or makes a necessary demand) is a sufficient justifying reason to adopt or comply with the principle. The lawgiving form of the principle is the ground of a particular agent's choice (i.e. is that agent's reason for acting) when the agent complies simply because the principle meets these criteria and makes a

<sup>2</sup> Stephen Engstrom, 'Introduction', in Kant, *Critique of Practical Reason*, trans. Werner Pluhar (Indianapolis: Hackett Publishing Company, 2002), xxxiv–xxxv. See also his *The Form of Practical Knowledge* (Cambridge, MA: Harvard University Press, 2009), 115–17, 122–7.

<sup>3</sup> Engstrom, 'Introduction', xxxv.

necessary demand on any rational agent. A formal practical principle – or better, since Kant appears to think that there is only one, *the* formal principle of volition – is presumably the directive to act only from maxims that have the form of a law and to regard this demand as a sufficient reason for choice. This principle is seen in the universal law version of the Categorical Imperative, which Kant refers to, variously, as ‘the formal principle of volition’ and ‘the formal principle of pure practical reason.’<sup>4</sup> I accept these basic understandings, but by raising various questions, I will try to fill out what they amount to.

In the next section, I attempt to say something further about Kant’s notion of the form of a law. In section III I develop an understanding of what Kant means by a ‘formal principle’ that explains why formal principles are uniquely suited to apply with normative necessity. In the concluding section I shall use these points to give readings of the arguments for Kant’s Theorems I and III and Problem I.

## II. THE FORM OF A LAW

Let me rehearse several familiar points that provide some parameters for understanding the idea of the form of a practical law. First, Kant thinks that the form of a law is common to both natural laws and practical laws. Second, some practical principles have the form of a law and some do not. Third, the form of a practical principle inheres in some substance or content that is the ‘matter’ of the principle. The last two points raise the questions how a principle can provide a ground of choice in virtue of having the form of a law, and whether it is only principles with the form of a law that provide a ground of choice through their form.

On the first point, textual support that Kant thought that the form of a law is common to both natural causal laws and practical laws, should one need it, is found in the ‘Typic of Pure Practical Reason’. This brief but obscure section of the second *Critique* gives the theoretical underpinning for the law of nature variant of the Categorical Imperative found in the *Groundwork* (*G* 4:421, ¶33). Kant wishes to explain how we can judge whether a maxim specifying a proposed choice that can be carried out in the sensible world (e.g. an action to be done for some end or reason, all specifiable in experiential terms) satisfies the criterion of morally good choice – that it have the form of universal law. Practical (moral) judgment is ‘subject to special difficulties’: since the criterion of morally good choice is a purely rational

<sup>4</sup> *G* 4:402 and *CpV* 5:41.

requirement, it cannot be represented in experiential terms in a way that can be applied directly to a maxim – that is, to subjective principles that can be implemented in the sensible world – to determine whether the maxim ‘is a case of’ or satisfies the criterion.<sup>5</sup> It is not completely clear why (or whether) this is a problem. But perhaps the idea is, first, that whether a maxim has the form of a law cannot be read directly off the matter of the maxim; and second, that if that we could determine whether a criterion is satisfied simply by considering the matter of the maxim, that criterion would employ a ‘material principle of morality’.<sup>6</sup> However, in lieu of applying the fundamental criterion of good choice directly to maxims of action, Kant writes that we can use a law of nature as ‘a type for the appraisal of maxims in accordance with moral principles. If the maxim is not so constituted so that it can stand the test as to the form of a law of nature in general, then it is morally impermissible’ (CpV 5:70–1). In other words, we determine whether a maxim has or is consistent with the form of a practical law (satisfies the fundamental criterion of pure practical reason) by determining whether it could hold as a universal law of nature. The fact that we can assess the legislative form of a maxim in this way indicates that ‘the form of lawfulness in general’ (CpV 5:70) is shared between natural laws and practical laws.<sup>7</sup>

<sup>5</sup> As Kant describes the problem: ‘it seems absurd to want to find in the sensible world a case which, though as such it stands only under a law of nature, yet admits of the application to it of a law of freedom and to which there could be applied the supersensible idea of the morally good, which is to be exhibited *in concreto*’. The ‘special difficulty’ is that ‘a law of freedom is to be applied to actions as events that take place in the sensible world and so, to that extent, belong to nature’ (CpV 5:68). The issue is how laws of freedom and the rational idea of the morally good can be applied to actions that are events in the natural world, given the fact that moral principles and ideas of the good cannot be instantiated in experience.

<sup>6</sup> To elaborate the second part of this suggestion, the idea is that any criterion of moral goodness that could apply directly to maxims of action would assess maxims through their matter, and that any such method of assessment would employ a ‘material principle of morality’. To see why, assume that the criterion of moral goodness were such that we could determine whether it is satisfied by examining the content of the maxim, or some properties of the maxim accessible through observation and sensible experience (as Kant understands it). That is, assume that we could read off from the content of the maxim whether or not it is morally good. For that to be possible, the criterion of moral goodness would have to be specifiable in experiential terms. One possible criterion of this sort might assess maxims in terms of their consequences – e.g. whether adoption of the maxim promotes some value in which we as a matter of psychological fact take an interest. Another might assess maxims in terms of our disinterested responses to them – e.g. whether disinterested reflection on the maxim would lead to feelings of pleasure or satisfaction (of approval). There may be other possibilities as well, but the thought is that any criterion of goodness specifiable in experiential terms would end up employing what Kant calls a ‘material principle of morality’. But no such principle, Kant argues, can have the requisite normative necessity. I’ll say more about this point in section IV (a).

<sup>7</sup> Here Kant writes: ‘It is permitted to use the nature of the sensible world as the type of an intelligible world, provided that I do not carry over into the latter intuitions and what depends upon them but refer only to the form of lawfulness in general [*die Form der Gesetzmäßigkeit überhaupt*] ... For to this extent laws are the same, no matter from what they derive their determining ground’ (CpV 5:70).

Regarding the second point, it seems equally clear that some maxims have (or are consistent with) the form of a law and that some do not; otherwise lawgiving form would not provide a criterion of morally good choice. Kant thinks that ‘the most common understanding can distinguish without instruction what form of a maxim makes it fit for a giving of universal law and what does not’ (*CpV* 5:27).<sup>8</sup> How then might we characterize the form of practical principles that are *not* suited for a giving of universal law? One suggestion is that what Kant calls the ‘principle of happiness or self-love’ captures the form of a practical principle that is to be contrasted with the form of a law.<sup>9</sup>

The principle of happiness is the principle of making ‘a rational being’s consciousness of the agreeableness of life uninterruptedly accompanying his whole existence ... the supreme determining ground of choice’, or the principle of choice that places ‘the determining ground of choice in the pleasure or displeasure to be felt in the reality of some object’ (*CpV* 5:22). In other words, it is the fundamental principle of finding reasons for action in the fact that an object of choice offers prospective satisfaction. An agent who makes happiness his principle will act so as to secure or increase overall satisfaction of his desires, or will do what he desires most strongly on balance. As I understand Kant’s principle of happiness, it sets no limits on the objects in which one takes satisfaction. For example, the objects of desire need not be self-regarding, as we learn from Kant’s friend of humanity in *Groundwork* I.<sup>10</sup> (In this respect it is somewhat misleading of Kant to refer to it as the principle of ‘self-love’.) The principle of happiness is thus a ‘formal’ principle in the contemporary sense – it directs action to ensure

<sup>8</sup> Some reason to think otherwise might come from remarks such as the following: ‘Now, all that remains of a law if one separates from it everything material, that is, every object of the will (as its determining ground) is the mere *form* of giving universal law’ (*CpV* 5: 27). One might take that to mean that once we abstract away the content of any maxim or practical principle (or set aside reasons for adopting the principle based in a contingent interest in the content), what is left is the form of a law – in which case the form of a law would be the general form of any practical principle. But that reading would imply that all practical principles have the form of a law, which is implausible. In this remark, however, Kant is talking about practical laws (*Gesetze*), not principles (*Prinzipien*). That said, it may still be Kant’s view that a *taking* an action to be objectively good, or *taking* a maxim to be universally valid, is a formal element of all free choice. For discussion of this point, see Reath, ‘Autonomy, Practical Law, and Taking One’s Choices to Be Good: Replies to Critics’, *Philosophical Books* 49:2 (April 2008), 132–4.

<sup>9</sup> I made this suggestion in an earlier essay, ‘Hedonism, Heteronomy, and Kant’s Principle of Happiness’, reprinted in *Agency and Autonomy in Kant’s Moral Theory* (Oxford and New York: Oxford University Press, 2006). See sections III–IV and the Appendix.

<sup>10</sup> See Reath, *Agency and Autonomy*, 39–40. I also argue in this essay that Kant does not regard pleasure or satisfaction as the object of all choice done from the principle of happiness, thus that he does not accept a simple form of psychological hedonism for action on this principle.

satisfaction of your desires whatever their objects may be – although abstraction from the objects of volition is not enough to make it a formal principle in Kant’s specialized sense. What comes close to making it a formal principle in Kant’s sense is that the principle of happiness captures the defining feature or form of one model of human choice – that feature being that the reason for choice lies in expected satisfaction or the ways in which the object of choice answers to antecedently given object-dependent desires and dispositions that are part of an individual’s subjective motivational set.<sup>11</sup> This is the structural feature shared by what Kant calls ‘material practical principles’ – practical principles that have normative force only on the condition that one has an object-dependent desire for, or is antecedently disposed to take satisfaction, in their object.<sup>12</sup> The significance of the principle of happiness for Kant is that it is the basic principle underlying a certain form of practical principle or choice that may be contrasted with choice guided by moral principle.

The third point is that (with the exceptions noted below) every practical principle has a form and a matter, and the form inheres in some matter or content and is the form of some substantive principle. So a practical principle with the form of a law will be a substantive maxim or principle that has the form of a practical law, for example, a substantive principle that makes a necessary demand on anyone.<sup>13</sup> Likewise what has the form captured by the principle of happiness are specific material practical principles. The exceptions are the most general principles, such as the formal principle of morality or the principle of happiness, that abstract from the content that differentiates one specific principle (practical law or material practical principle) from another. I presume that the formal principle of

<sup>11</sup> See Williams, ‘Internal and External Reasons’, in *Moral Luck* (Cambridge and New York: Cambridge University Press, 1981), 101–13. For the idea of ‘object-dependent desires’, see Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge, MA: Harvard University Press, 2000), 46, 151–2. Rawls characterizes them as desires for objects or states of affairs ‘that can be described without the use of any moral conceptions, or reasonable and rational principles’. That makes them desires that arise independently of practical reasoning and judgments about value – desires that do not presuppose or are not generated by the acceptance of moral conceptions or the application of principles of reason.

<sup>12</sup> Such principles presuppose ‘an object (matter) of the faculty of desire as the determining ground of the will’ where ‘desire for this object precedes the practical rule and is the condition of its becoming a principle’ (*CpV* 5:21). For further discussion of the concept of a material practical principle, see Reath, ‘Introduction’, in Kant, *Critique of Practical Reason*, trans. and ed. Mary J. Gregor (Cambridge and New York: Cambridge University Press, 1997), xx.

<sup>13</sup> Substantive principles with the form of a law are ‘formal principles’ in Kant’s sense, though they need to be distinguished from *the* formal principle of morality, e.g. the Formula of Universal Law. In discussing formal principles in section III, I will have the latter in mind.

morality, in addition to expressing the requirement that one's maxims have the form of a law, itself has the form of a law.

When Kant holds that a practical law provides a ground of choice through its form, he is thinking in the first instance of substantive principles and maxims that involve either requirements or prohibitions (a principle of providing aid, of fulfilling a voluntarily undertaken obligation, of avoiding dishonesty and so on). In his discussion of the duty to promote the happiness of others, the determining ground of choice is the 'mere lawful form' of this particular principle.<sup>14</sup> So his view is best stated as follows: whether there is sufficient justification to act on some substantive principle turns on the form instantiated in the matter of that principle; the ground of choice is the fact that a certain substantive principle has (or is consistent with) the form of a practical law. The fact that a certain principle, such as a principle of mutual aid, has the form that makes it a requirement is sufficient (decisive) and overriding reason to adopt that principle, whether or not one is inclined or would profit from doing so. The fact that a certain principle does not have or is not consistent with the form of a practical law (cannot be made a universal law) is a sufficient reason to abandon a maxim that otherwise holds considerable interest. In the case of both requirements and prohibitions, the form of a practical law is a ground of choice in a meaningful sense, first, because the fact that a principle has a certain form is a reason for choice (either to adopt or to abandon the principle), and second, because that reason is sufficient to override or exclude competing material reasons based in desire or interest – e.g. that helping would be burdensome or inconvenient, that the deception would further an important personal aim, and so forth.

Permissible maxims introduce complications, but I assume that one can say that they have the form of practical law in an extended sense marked by the phrase 'consistent with the form of practical law'. A maxim contains a proposed action and some set of considerations regarded as good reasons for performing it (such as an end that the action will further, one's circumstances of action, and so forth). When a maxim can be made universal law without inconsistency or irrationality, all can agree to anyone's adopting the maxim (to regarding those considerations as good reasons) and no one may

<sup>14</sup> Kant writes: 'Thus the matter of the maxim ... must not be the condition of the maxim since the maxim would then not be fit for a law. Hence the mere form of a law, which limits the matter, must at the same time be a ground for adding this matter to the will ...' (*CpV* 5:34). Note also that Theorem III is concerned with how an agent can 'think of his maxims as practical universal laws', and that the following Remark illustrates with the example of a deposit whose owner has died; the issue is whether this maxim has the form of lawgiving.

complain of an agent who acts on those reasons. Here, the reasons stated in the agent's maxim do justify the action: they can be regarded by anyone as good reasons for that agent to perform the action. The fact that a maxim is in this way consistent with the form of a practical law can be an agent's reasons for action, again in an extended sense, when the agent recognizes the permissibility of the maxim as a limiting condition on its adoption – that is, when the agent would have rejected the maxim had the proposed reasons not met this standard of permissibility.<sup>15</sup>

Let me now raise two questions: First, can the form of a practical law provide a ground of choice by itself? Second, can Kant claim (as he wishes to) that only a practical law can provide a ground of choice through its form? The first question arises because the form of a law by itself cannot point the will in any specific direction. The higher-order commitment to act only from principles that have the form of a law is a fundamental practical orientation that determines which substantive considerations an agent treats as good reasons. But the substantive considerations that provide the matter of choice specify what is to be done. The higher order commitment cannot lead to action unless it is applied to some concrete circumstances. Nonetheless, Kant can hold that the lawgiving form of a maxim can provide a ground of choice by itself in the case of requirements and prohibitions (though not permissions): the fact that a maxim or substantive practical principle has the form of a law (the form of a requirement) is a sufficient reason to adopt it and the fact that it does not is a sufficient reason to abandon it. The matter provides direction for the will, but the ground of choice is the fact that the matter has (or lacks) the form of a practical law.

Regarding the second question, I have suggested that the principle of happiness is a kind of formal principle because it captures the form of choice on a material practical principle. The fact that an object of choice offers satisfaction or answers to one's desires and dispositions sounds like a reason of a very general sort – i.e. something that can be said on behalf of that choice; or at least it seems to be the form of a (kind of) reason. Someone who adopts and acts on a material practical principle is taking the fact that the object of choice answers to his independently given desires and dispositions as his reason for action at the most general level. This may seem a case where the form of a principle – here the form of a material practical principle – provides a ground of choice (a motivating reason, even if not a sufficient justifying

<sup>15</sup> The idea that the motive of duty, thus the permissibility of a maxim, can function as a 'limiting condition' on choice was introduced by Barbara Herman; see *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press, 1993), chapter 1.

reason).<sup>16</sup> Hence my question: does a practical law determine choice in virtue of its form in a way that a material practical principle does not? For the following reason I would say, 'yes'. A material practical principle (such as the principle of seeking all the comforts of life, or the principle of working and saving in one's youth to provide for one's old age) gets its normative force for an agent through its relation to the agent's interests – i.e. from the fact that its object answers to an agent's antecedent desires and dispositions. Since the reasons given by material practical principles have subjective conditions, any specification of those reasons must be completed by a reference to those interests. Thus one cannot say that the form of a principle by itself provides a reason for choice. By contrast, in the case of a practical law no reference to the interests of the agent is needed to substantiate the reason for adopting the principle. The ground of choice is the fact that the matter of the principle has the form of a law.<sup>17</sup>

Let me close this section with another way of characterizing the form of a practical law that fits well with the three familiar points just outlined. Laws of any kind state necessary connections that hold between some ground and some consequence. One might think of the form of a law as the structural relationship that holds between ground and consequence, the formal relationship characterized by the notion of sufficient reason.<sup>18</sup> Natural (causal)

<sup>16</sup> To illustrate, take the example of someone who seeks all the comforts of life and never willingly forgoes any comfort. The substantive considerations that he treats as good reasons include the appealing features of good meals, well-tailored clothes, well-designed furniture and cars, the avoidance of stress, and so on. He need not consciously have adopted the principle of seeking all comforts, but this principle unifies and explains his choices. His substantive reason for choice in a given instance will be an appealing feature of some object – of this meal, this article of clothing, this relaxing activity. But at the most abstract level, his reasons derive from the fact that this overall end (or this method of determining what count as reasons) appeals to him, and that these specific objects of choice fit together to further this end. The fact that he finds these objects of choice agreeable, given his desires and tastes, is the common structural feature (form) of his choices, and it seems to me that with a bit of abstraction, he can cite this fact as a reason for choice. Is it then fair to say that this formal feature is the ground of choice and that a material practical principle can determine choice through its form?

<sup>17</sup> I have been discussing substantive practical laws and have argued that a substantive practical law can determine choice through its form alone, while a material practical principle cannot. Can the Formula of Universal Law (FUL), in contrast to the principle of happiness (PH), determine choice through its form? Both FUL and PH can be the basis of higher-order commitments (basic practical orientations) to treat certain substantive considerations as reasons in certain ways, and neither can lead to action without application to some concrete circumstances. In the case of PH, choice requires further input from an agent's interests (so it cannot determine choice beyond being the basis of the aforementioned commitment). But as I argue in the next section, FUL can determine the will in virtue of its form: the formal principle or morality is a ground of choice in virtue of the fact that it is the formal principle of pure (i.e. free) volition, or is the form of autonomous willing.

<sup>18</sup> Here I draw on Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge and New York: Cambridge University Press, 1996), 102, and *The Sources of Normativity* (Cambridge and New York: Cambridge University Press, 1996), 107–8. I discuss this idea further in *Agency and Autonomy*, 133–4.

laws are concerned with necessary connections that obtain between antecedent conditions and some effect when those conditions are sufficient to produce the effect in all cases. A natural law states that whenever certain conditions obtain, a specified event follows as their effect. Likewise practical laws are concerned with the relationship between considerations that provide practical reasons and choices when those considerations constitute sufficient justification for the choice. A practical law states, in effect, that certain considerations or facts about the circumstances (that an individual needs life-saving aid that one can provide, that one has promised to repay) are in every case sufficient justification for some action (providing the aid, repaying).<sup>19</sup> This structural relationship – necessary connection between ground and consequence, or sufficient reason – is common to both natural laws and practical laws. It is also a relationship that in a relatively clear sense ‘inheres in the matter’ of a maxim or principle. A maxim has the form of a law when a certain relationship holds between the components of the maxim – when the considerations cited by the maxim are indeed a sufficient justification for the action. And this relation of sufficient reason will obtain in some practical principles but not others. Finally, practical laws will have both the objective universal validity and subjective universal validity to which Engstrom points: a practical law holds that a certain consideration is a sufficient justification for any agent in the situation covered by the principle; and it governs how all subjects should think about the issue – it is a reason that anyone can agree to or regard as sufficient.

### III. FORMAL PRINCIPLES (THE FORMAL PRINCIPLE OF VOLITION)

In this section I discuss one way of understanding Kant’s conception of a ‘formal principle’. As far as I know, Kant does not explicitly characterize formal principles in the way that I shall propose (though there is some textual evidence for attributing this view to him). But this account offers a natural explanation of why Kant thought that the fundamental principles of

<sup>19</sup> Again, the structural relationship between the considerations that provide reasons and the choice, and thus the notion of sufficient reason, are slightly different in requirements, prohibitions and permissions. In the case of a requirement, the relevant considerations – the fact that an agent needs life-saving aid that one is in a position to provide – are in every case decisive and overriding reason for the action (providing the aid). In the case of permissions, the relevant considerations provide sufficient justification. When a maxim of self-interest is universalizable, the agent’s reasons (e.g. the fact that an action will further my end) are sufficient justification for the action which anyone must recognize as such (permitting but not requiring adoption of the maxim). In the case of a prohibition, they never provide justification: the fact that dishonesty will further my self-interest never justifies dishonesty.

certain domains of rational activity are ‘formal’, and why formal principles, so understood, are uniquely suited to apply with normative necessity to some domain of rational activity. We may think of the *form* of some rational activity or object of cognition as the constitutive or defining features of an activity or entity of that kind – the features that an activity or entity must possess to count as an instance of that kind. The form of some rational activity or object of cognition will be associated with a *formal principle* that is constitutive of that rational activity or object of cognition. The formal principle of some rational activity would be the guiding internal or constitutive norm that a subject must follow in order to engage in that activity. By specifying the form of that activity, it provides a norm that anyone engaged in that activity must satisfy and that in some sense does guide any instance of the activity (even if defectively).<sup>20</sup> The formal principle of a kind of object of cognition would be a norm that sets necessary conditions on thought about such objects. The form of a law would be the defining features that a principle must have in order to qualify as a practical law. Since practical laws regulate volition, or more specifically pure volition, the associated formal principle will be a principle of pure volition – the internal guiding norm of pure willing, or the principle that one must follow in order to guide one’s will by reason alone (ironic aside: if that is important to you ...).

<sup>20</sup> Let me mention a problem that I am unable to deal with adequately here: if a formal principle is constitutive of an activity, it must be implicated in all instances of the activity. The problem is how to characterize mistakes or defective instances of an activity. As an example, if syntactical rules are constitutive linguistic rules, they are rules that one must follow in order to produce speech. This holds for grammatically well-formed statements as well as those containing grammatical errors. Syntactical rules in some sense guide even ungrammatical speech – if they did not, the utterance would not count as an instance of speech, in which case it could not be the kind of activity that contains a grammatical error or is subject to assessment by grammatical rules. Likewise, someone who makes a mistake in addition in some sense follows the rules of addition, while at the same time violating them. In such cases, we want to say that the subject is guided by or implicitly follows the constitutive rules, even in instances where she violates them. The question of how defective instances of an activity (those that fail to conform to the rules) are guided by constitutive rules is a problem that this account of formal rules needs to address.

Barbara Herman makes a promising suggestion about this issue. One might think that all rational or cognitive activity is guided by a *representation* of various formal or constitutive principles, and that in defective activity, the relevant principle is *misrepresented*. Herman specifically develops this idea to explain how Kant can hold both that the moral law is a law of freedom and that morally bad action is freely willed. Her idea is that all free action is derived from a representation of the formal principle of the will (the moral law), but that in bad or faulty action this principle is misrepresented. Bad action is free because derived from a representation of the principle of the will, but faulty because that principle is misrepresented. See Herman, *Moral Literacy* (Cambridge, MA: Harvard University Press, 2007), 171–2, 245–6.

When formal principles are understood as constitutive principles, certain features stand out. Formal principles are, first, *enabling principles* or conditions. By defining what counts as a certain kind of activity they make it possible to engage in that activity. Second, by constituting or defining an activity, they are *internal* to the activity. Since a formal principle specifies the form of an activity and serves as a norm for any subject engaged in the activity, it is a principle that arises from the nature of that activity – as one might say, a principle that the relevant rational activity ‘gives to itself’. Third, and this is the point to which I am leading, formal principles, understood as constitutive principles, *apply with necessity* to anyone engaged in the activity, and their constitutive role explains their necessity. Since they specify how one must proceed in order to engage in the activity (for what one does to count as an instance of the activity), they are not coherently rejected by anyone engaged in that activity. It is for this reason that formal principles are uniquely suited to govern an activity with unconditional necessity. It is somewhat beyond my competence to pursue this idea elsewhere in Kant’s work, but I presume that this understanding of a formal principle and of the connection between formal principles and necessity can be seen throughout his system. The principles of formal logic specify the form of thinking in general, and can be viewed as enabling conditions of thought. Their necessity lies in the fact that unless one is governed by these principles, one is not thinking (and, one should add, thinking is a necessary activity for rational creatures like us). The pure categories and principles of the understanding (principles of transcendental logic) specify the form of and enable thinking about objects of experience, thus necessarily govern all thought about objects of experience. And so on.<sup>21</sup>

A formal principle involves some abstraction from content: the form is what remains when one sets aside those features that differentiate one instance of an activity from any other (e.g. that distinguish one principle of volition from another). But focusing on abstraction from specific differences in content only characterizes formal principles negatively. It does not bring out the foundational role of such principles and the point of introducing them. For example, simply pointing to the fact that a formal

<sup>21</sup> For the most part, the formal or constitutive principles of various domains of cognitive activity that one finds in Kant’s thought are synthetic a priori principles, such as the Categorical Imperative or the a priori principles of the understanding. But the principles of formal logic are analytic. Can formal or constitutive principles include both analytic principles and synthetic a priori principles? (Can the analytic status of formal logic be due to the fact that it sets constraints on thinking about any kinds of object whatsoever?) This is one of several issues that the understanding of formal principles proposed here needs to address in greater depth. I am grateful to Pierre Keller for raising this question.

principle abstracts from ends (or does not get its authority from any interest in an end) provides almost no insight as to why Kant thinks that fundamental principles must be formal and where their normative authority comes from. This is one reason why the argument for Theorem III – that a practical law provides a ground of choice through its form – appears too quick. The arguments here and elsewhere explain why material practical principles, understood as principles whose normative force presupposes an empirically given interest, lack the normative necessity required of a law. But they do not give a clear positive argument as to why formal principles and only formal principles (or why principles that determine choice through the form of a law and only such principles) do have this necessity. Understanding formal principles as constitutive principles fills this lacuna by establishing a connection between form and necessity. For that reason, when I refer to formal principles, such as the formal principle of morality, I mean to draw attention to their role as constitutive, and therefore necessary, principles in some domain of rational activity, and not to the fact that they involve a certain abstraction from content.

To see what this conception of a formal principle might do for Kant, it may help to consider how the connection between form and necessity figures in the more ambitious argument of the *Groundwork*. The overall task of the *Groundwork* is to show (through a ‘deduction’) that the special authority that ordinary moral thought assigns to moral requirements is genuine – as Kant puts it, to show how the necessity thought in a categorical imperative is possible (*G* 4:417). Kant’s resolution of this problem, in a word, is to argue that the moral law is the constitutive principle of a necessary self-conception. First, as rational agents we necessarily think of ourselves as free, various features of our capacity for theoretical reason confirm ascribing freedom to ourselves, and we identify with our capacity for free agency.<sup>22</sup> And second, the moral law is the formal or constitutive principle of free volition. In other words, he identifies the formal principle of morality with the formal principle of our agency. Kant launches his argument by deriving a statement of the Categorical Imperative – the Formula of Universal Law (FUL) – from the concept of a categorical imperative. This derivation presents the FUL as the formal principle of conformity to moral requirement – the principle one must follow in order

<sup>22</sup> The claim that we necessarily act under the Idea of freedom comes, of course, at *G* 4:448. That our theoretical capacities confirm ascribing freedom to ourselves (and that we necessarily act under the Idea of Freedom) comes at *G* 4:451–3. The claim that we identify with our capacities for free agency is seen in the references to the ‘proper self’ at *G* 4:457 and 461. I believe that this idea is behind the claim at *G* 4:453, that the intelligible world is the ground of the sensible world.

to conform to moral law. While this principle tells us what morality demands of us, understanding the principle as constitutive of conformity to moral requirement does not speak to the question of its authority. That is the function of the sequence of formulas that sets up the arguments of Section III. The introduction of autonomy (at *G* 4:431) shows that the form of a law is also the form of lawgiving – that is to say, the formal principle of conformity to universal law is the principle that is constitutive of the autonomy, or lawgiving capacity, of the will; it is the principle that you must follow to give law through your will. And the argument in *Groundwork* III that autonomy specifies the positive concept of freedom leads to the idea that the form of lawgiving is the general form of free volition. (Free agency is the capacity to act from principles that one gives to oneself through the will's own principle, independently of determination by external conditions.)<sup>23</sup> The formal principle of morality, in other words, is at the same time the constitutive principle of free volition: it is the principle one must follow to exercise one's free agency and the capacity to follow this principle makes one a free agent. On the assumption that volition is free volition, the sequence of formulas permits Kant to claim that the FUL 'contains merely the form of volition as such' (*G* 4:444). If it is the capacity to follow this principle that makes us agents, we cannot coherently reject its authority and continue to think of ourselves as agents. Thus, given the claim that we necessarily think of ourselves as free and identify with our agency, the necessity of morality is vindicated by showing that the fundamental (formal) principle of morality is the formal principle of our agency.

Although the main reasons for thinking of formal principles as constitutive principles are philosophical, there is textual evidence in the second *Critique* that Kant does understand formal principles in this way. I'll cite two passages. The first is the Remark following the Definition of a practical law, in which Kant says that a practical law 'determines the will as will'. In this paragraph Kant draws a contrast between, on the one hand, principles that determine the 'conditions of causality of a rational being as an efficient cause merely with respect to the effect' and, on the other hand, principles 'that determine only the will, whether or not it is sufficient for the effect', 'principles that determine the will as will even before I ask whether I have the ability required for a desired effect or what I am to do in order to produce it', or principles that have to do 'only with [one's] will, regardless of whether the purposes the human being may have can thereby be attained'.

<sup>23</sup> For further discussion, see Reath, *Agency and Autonomy*, 154–5 and Reath, 'Autonomy, Practical Law, and Taking One's Choices to Be Good', 130–4.

In the latter case (the case of a law), the principle determines ‘the mere volition’ (*das bloÙe Wollen*) and ‘refer[s] only to the will, without regard to what is attained by its causality’ (5:20–1). I am not sure that I fully understand certain details of this contrast, but this much is clear.<sup>24</sup> Practical laws do not prescribe means to desired ends or give directives that are conditional on subjective interests, but rather determine what it is necessary for any rational agent to will. That they ‘determine the will as will’ suggests that practical laws are internal to (constitutive of) purely rational volition – the kind of principle that one must follow in order to will from reason alone. They apply in virtue of defining features of pure volition that hold for any agent with that capacity. The higher-order principle of conformity to practical law would be the principle that defines the capacity for purely rational volition, and therefore is not coherently rejected by any agent with this capacity. The formal principle of morality, in other words, is the formal principle of pure volition.

Second, in Remark I following Theorem IV, Kant writes: ‘the necessity that the law [of the pure will] expresses, since it is not to be a natural necessity, can therefore consist only in the formal conditions of the possibility of law as such’ (*CpV* 5:34). This Remark follows the claim that the moral law is the principle of the autonomy of the will, but it makes the point that the moral law states the constitutive or defining features that a principle must have to qualify as a practical law. That makes it the constitutive principle of conformity to moral requirement, which we know is the formal principle of pure volition.

On one influential reading of Kant’s Formula of Humanity, it can also be understood as a ‘formal principle’ in the sense explained here: it is the formal principle of finding sufficient reason to adopt some end and taking its achievement to be worth while or objectively good. According to this reading of the value of humanity or rational nature, the exercise of rational choice confers objective value on its objects. Rational choice is directed at ends taken to be objectively good, and since the value of these ends comes from the fact that they are rationally chosen by some person, rational nature is implicitly valued in all rational choice. One component of this view is that ends are adopted for the sake of some person (often myself, but possibly

<sup>24</sup> Kant says that laws ‘determine only the will, whether or not it is sufficient to the effect’ and ‘... even before I ask whether I have the ability required for a desired effect’. See also, *CpV* 5:45, 66. But why should practical laws set aside one’s ability to achieve some end? After all, pursuing an end that one knows to be beyond one’s capacity involves a kind of irrationality. Presumably his idea is that laws concern what we are to will for its own sake in some absolute sense, and what we are to will in that sense should no more depend on our practical abilities than on our subjective interests.

others as well), and the value of the person confers value on the achievement of the end. What makes the achievement of some (rationally chosen) end worthwhile is that it is adopted for the sake of an end in itself, and ends in themselves are persons. For example, in finding sufficient reason to pursue an end in the fact that it interests me and, having made it my end, thinking that its achievement matters, I value myself as an end in itself. So it is by adopting an end for the sake of some person and valuing that person (in this case, myself) as an end in itself that one finds the achievement of the end to be objectively worth while. And whenever one takes the achievement of an end to matter in this way, it is adopted for the sake of some person who is valued as an end in itself. In this way, one can argue that valuing humanity as an end in itself is the formal principle of finding objective value in the achievement of rationally chosen ends.<sup>25</sup>

In this section, I have claimed that since formal principles are constitutive of some domain of rational activity, they apply with normative necessity to anyone engaged in that activity. Whether such principles are unconditionally necessary, of course, depends on whether that form of rational activity is necessary, or one to which we are unavoidably committed. Here it is important to note that early in the first chapter of the second *Critique*, Kant asserts up front that the formal principle of morality is the formal principle of *pure* volition. The *Groundwork*, by contrast, aims at the more ambitious claim that the formal principle of morality is the formal principle of *volition as such* – that is, that it is constitutive of the brand of free volition that underlies and is exercised in *anything* recognizable as rational choice. If the *Groundwork* argument succeeds, it would show that the basic principle of morality is not coherently rejected by anyone in the business of choosing (which we are in most of the time). But it is hard to see how this argument could succeed. To make it work, one needs to assume that a strong and controversial conception of transcendental freedom underlies all forms of rational agency; but there are certainly coherent conceptions of agency that do not meet this standard (and one may indeed be true of us). In the second *Critique*, Kant pulls back from this claim, perhaps because he came to see the difficulties of showing that a rational agent necessarily acts under the idea of *transcendental* freedom, but certainly because of the obstacles to ascribing transcendental freedom to ourselves on grounds that are independent of moral consciousness. He takes the more modest line that the

<sup>25</sup> For examples of this reading, see Korsgaard, 'Kant's Formula of Humanity', in *Creating the Kingdom of Ends*, 114–24, and *The Sources of Normativity*, 120–6; and Wood, *Kant's Ethical Thought* (Cambridge and New York: Cambridge University Press, 1999), 125–32.

formal principle of morality is the formal principle of pure volition – an assertion with which one could hardly take issue.<sup>26</sup> The question to ask here is whether we must be in the business of pure willing. The fact of reason asserts, among other things, that we are in this business. The fact of reason, as I understand it, is our recognition of the authority of moral concerns in everyday moral thought, judgment and feeling, and the subsequent awareness of our capacity to act from moral concerns, even in the face of powerful countervailing reasons.<sup>27</sup> The fundamental (formal) principle of pure volition presents itself to us, synthetically, with an authority ‘not to be outweighed by any sensible conditions and indeed quite independent of them’ (*CpV* 5:29–30) that, on due reflection, we cannot reject without loss, and our recognition of its authority reveals the capacity for purely rational volition. We do, indeed, take the standards of the pure will to apply to us with normative necessity, though a deduction that establishes this necessity is neither possible nor required (*CpV* 5:47).<sup>28</sup>

#### IV. INTERPRETATIONS

In this section, I’ll use some of the points made above to reconstruct the arguments of Theorems I and III and Problem I.

##### (a) *Theorems I and III*

Theorems I and III claim, first, that no material practical principle can furnish a practical law, and second, that if a principle is to serve as a law, it must provide a ground of choice through the fact that it has the form of a

<sup>26</sup> The pullback in the second *Critique* is not complete. Throughout Chapter I, Kant is explicit that the moral law is the formal principle of pure volition, and I take that to differ from asserting that it is the formal principle of volition as such. But at *CpV* 5:32 he does appear to make the stronger claim:

Now this principle of morality, just on account of the universality of the lawgiving that makes it the formal supreme determining ground of the will regardless of all subjective differences, is declared by reason to be at the same time *a law for all rational beings insofar as they have a will*, that is, the ability to determine their causality by the representation of rules, hence insofar as they are capable of actions in accordance with principles and consequently also in accordance with a priori practical principles [italics added].

Here Kant claims that the moral law is ‘a law for all rational beings insofar as they have a will’ and goes on to define the will in the standard way. But an agent whose principles were based on desire and contingent interests would still have a will according to this conception. If so, he cannot claim that the moral law is the formal principle of volition as such.

<sup>27</sup> I take this understanding of the fact of reason from Rawls. See ‘Kant: Lecture X’, in *Lectures on the History of Moral Philosophy*, especially 255, 260.

<sup>28</sup> For further discussion of the fact of reason and the changes in argument between *Groundwork* III and the second *Critique*, see the essays by Pauline Kleingeld and Jens Timmermann in this volume.

law. These Theorems are the basis of Kant's claim that the fundamental principle of morality cannot be a material practical principle, and that only a formal practical principle can apply with the necessity implicit in the ordinary concept of duty (*CpV* 5:41, 64–5). To substantiate this claim, Kant must first show that there is an exhaustive distinction between practical principles that provide a ground of choice through their matter and those that do so through their having the form of a law. That distinction will be exhaustive only if Kant can make the case that his category of material practical principles is quite inclusive – that is, that all principles other than those that are formal in his sense 'presuppose an *object* (matter) of the faculty of desire as the determining ground of the will' (*CpV* 5:21). Second, he needs a positive account of why only formal principles (or principles that provide reasons through their having the form of a law) are normatively necessary.

A material practical principle gets its normative force from an antecedently given interest in some object, broadly construed to include a wide range of ends and values. Kant explains why all such principles are empirical and unable to ground laws in the following passage:

For, the determining ground of choice is then the representation of an object and that relation of the representation to the subject by which the faculty of desire is determined to realize the object. Such a relation to the subject, however, is called *pleasure* in the reality of an object. This would have to be presupposed as a condition of the possibility of the determination of choice ... [I]n such a case the determining ground of choice must always be empirical, and so too must be the practical material principle that presupposes it as a condition. (*CpV* 5:21)

The passage describes one way in which a practical principle can provide reasons – the form of choice, as noted in section II above, captured by the 'principle of happiness'. But it also describes the method of moral philosophy employed by theorists who 'begin from a concept of the good in order to derive from it laws of the will' (*CpV* 5:63). Kant rejects this method in Chapter II because a theory that proceeds in this way cannot produce genuine practical laws, for the reasons given here. Such theorists present an object of choice as intrinsically good or choiceworthy – e.g. Epicurean happiness, agreement with Hutchesonian moral feeling, Wolffian perfection, and so forth – and from this object derive rules of conduct (say, to act in ways conducive to or in agreement with this object). Here Kant claims that the reasons for accepting such a principle come from (first) the representation of the object and (second) the fact that the representation of the object has a certain relation to an agent through which the agent is

moved to take an interest in the object, a relation that Kant terms 'pleasure in the reality of an object'. I take that to mean that features of the object provide reasons because when represented to an agent, they produce interest – they elicit responses of interest or answer to antecedent dispositions that agents are assumed to have. These responses of interest are 'based on the receptivity of the subject' and belong to 'feeling' (*CpV* 5:22). They include any interests that an agent may have, and indeed could share with most other members of the species, as a matter of empirical fact, but that are not essential to rational capacity. In this way, any such principles are 'subject to an empirical condition' and thus not laws (*CpV* 5:27).

This argument shows that no principle of pursuing an obviously desire-based end can be a practical law. More importantly, it shows that empiricist moral theories, such as an Epicurean principle of happiness or Hutcheson's or Hume's moral sense theory, cannot ground the necessity that is part of the ordinary concept of duty (not that this fact would disturb them). But Kant needs to address this objection: Why should one think that all principles based on an end or substantive value are 'subject to an empirical condition' in this way? Some theorists will resist assimilating their substantive first principles to Kant's understanding of a 'material practical principle' (where the reason-giving force presupposes rationally contingent responses of interests). We can readily imagine principles based on ends or substantive values that are taken to be immediately or intrinsically good and to make claims on the will of any agent with adequate powers of appreciation—for example, Wolffian perfection, Cumberland's greatest happiness of all rationals or Clarke's eternal fitnesses. If such ends and intrinsic values make necessary claims on the will, as these theorists might hold, they would be the basis of substantive practical principles that hold for any rational agent, and thus hold as laws.

Kant's reply to this objection will have different facets. First, he can counter that many of the objects proposed as the basis of fundamental principles are not in fact necessary objects of volition; they can be rejected without obvious irrationality.<sup>29</sup> A second and more direct response is that any theory that 'begins with the concept of the good' does employ the method that he rejects, and that the representation of an end or value – as immediately or intrinsically good – is not enough to establish a necessary claim on the will.<sup>30</sup> How would such a theory establish that some end

<sup>29</sup> For discussion of this point, see Thomas E. Hill, Jr, *Dignity and Practical Reason* (Ithaca: Cornell University Press, 1991), 128–31.

<sup>30</sup> Here again I have learned from Engstrom, 'Introduction', xxxvi–xxxix.

(e.g. perfection) or a set of substantive values (e.g. relations between agents that ground the moral fitness of certain acts) is intrinsically choiceworthy and can ground practical principles? Presumably, all that theory can do is to present the end or value for our consideration – as good. If we respond to the representation of the end or value with interest, the associated principle will be reason-giving. But no necessary claim has been established unless it can be shown that one ought to respond with interest (regardless of antecedent dispositions), and a theory that begins by presenting what is good in this way lacks the resources to establish this claim. This is the methodological point of Chapter II: a theory that begins by proposing some end as good cannot establish true practical laws, because the representation of the end can provide reasons only in relation to antecedently given dispositions and interests that are not part of rational capacity, and are in that sense subjective.<sup>31</sup> To show that an end is necessary, one must show that it is required by a principle to which any rational agent is committed; that is, one must begin from a practical law.

Along these lines one can refine Kant's argument that any principle based on the prior presentation of an end or value as intrinsically good is 'subject to an empirical condition'. Kant's positive account of why the necessity of a practical law is conferred by its form, very briefly, has to do with constitutive role of the moral law in pure willing. When one takes the lawgiving form of one's maxims as a sufficient reason for action, one acts from the formal principle of morality, and that is what it is to exercise one's pure will. The formal principle of morality is the constitutive principle of pure volition, thus necessary for anyone engaged in that activity.

### *(b) Problem I*

Problems I and II argue that 'freedom and unconditional practical law reciprocally imply each other' (*CpV* 5:29). I understand Kant to be claiming that the formal principle of pure practical reason is the formal (constitutive) principle of transcendently free agency and that the capacity to act from this principle confers transcendental freedom. Let's see how far Kant can get in solving 'Problem I', which reads as follows:

<sup>31</sup> See *CpV* 5: 63: 'Suppose that we wanted to begin with the concept of the good in order to derive from it laws of the will ... Now, since this concept had no practical a priori law for its standard, the criterion of good and evil could be placed in nothing other than the agreement of the object with our feeling of pleasure and displeasure.'

Supposing that the mere lawgiving form of maxims is the only sufficient determining ground of the will: to find the constitution of a will that is determinable by it alone. (CpV 5:28)

To begin with a question, can this mystery will take its reasons from considerations other than the fact that its maxims have the form of a law? Certainly the lawgiving form of its maxims is the only sufficient *justifying* ground of choice. But that point aside, could this volitional capacity choose to act from some material principle? The way Kant poses the problem, clearly not: this will is ‘determinable alone’ through the lawgiving form of its maxims and no ground of choice other than the ‘mere universal lawgiving form can serve as a law for it’ (CpV 5:28, 29). It is then a pure will – a volitional capacity to act exclusively on purely rational considerations; and the FUL is its formal principle. Given this stipulation, what can we conclude about such a will?

We know that a will that bases its reasons exclusively on the lawgiving form of its maxims will not act from practical principles whose normative force depends upon empirically given interests. Interests of this sort arise through the experience of finding an object or activity agreeable in some way, as a result of an individual’s aptitudes and susceptibilities. Such interests ‘belong among appearances’ and can be given naturalistic causal explanations in terms of psychological principles. But the stipulation that the mystery will acts independently of such interests does not secure the independence from natural causation that Kant is after without some argument. Kant’s conception of rational agency commits him to the view that even interests of this sort influence choice in some sense independently of natural causal laws. The interests from which a material practical principle gets its normative force are brought to practical consciousness as rational dispositions – as acceptance of principles and values, or the tendency to view certain considerations as reasons, and so on – and they govern choice normatively, not causally. Let me explain.

Suppose that one of my ends is to increase my knowledge of nature and that I am attached to this end because I find it stimulating to learn about nature. It agrees with me. So I am imagining that this end is the basis of a material practical principle. There is a naturalistic causal story to be told about how this interest arises that will cite my aptitudes and dispositions, various experiences that I have had, and so on. However, although there is a causal explanation of the origin of this interest, it is taken up into practical consciousness in a certain form (and I use that term intentionally) – e.g. as a disposition to judge that increasing my understanding of nature is a worthwhile activity and to see reasons for action in opportunities to add to my

knowledge. Further, this interest influences my choices through my self-conscious application of various normative considerations and rational principles – for example, through judgments that I have good reason to pursue some opportunity to increase my understanding of nature, given that it is one of my ends. And as a rational agent, my judgments never ‘consciously receive direction from any other quarter’ (*G* 4:448). They are never directed by anything other than my own grasp of various reasons and my application of relevant rational norms to my circumstances. There is then a substantive sense in which rational choice on a material practical principle is independent of determination by causal laws: since it is governed by judgments about reasons and normative principles, it is governed other than by causal laws. Though the interest from which the practical principle gets its normative force has a causal explanation, choice on that principle is still normatively, rather than causally governed. It thus displays a not uninteresting form of freedom.

In the case of my decent though not pure willing, the rationality of judgment and choice blocks naturalistic causal explanation of the choice, but not of the interests on which the reasons for choice are based. But pure volition is not based on any such interests and takes its reasons only from the fact that a certain substantive principle instantiates the form of a law. What Kant seems to be getting at in Problem I is that for pure volition, all causal explanation is blocked, because the reasons for choice are not based on any empirically given interest (something given in appearances) for which a causal explanation can be given. When the ground of choice is the legislative form of a maxim, ‘the determining ground of the will is distinct from all determining grounds of events in nature in accordance with the law of causality, because in their case the determining grounds must themselves be appearances’ (*CpV* 5:28–9). So now suppose that I bring my activity under the duty of natural perfection and increase my understanding of nature out of a commitment to make good use of my natural faculties. Since my reasons for choice are now based only on the lawgiving form of my maxim, they are not based on any interest in the matter of choice that admits of the kind of causal explanation sketched above. To put the point another way, since I act on what is necessarily a reason for any rational agent (the lawgiving form of my maxim), my reasons are not taken from interests due to influences specifically on me. In this way the activity of a pure will (of an agent who exercises the capacity for purely rational volition) ‘must be thought as *altogether* independent of the natural law of appearances, namely the law of causality’ (*CpV* 5:29; italics added). It is governed by a fundamental normative principle that operates independently of any interests that

admit of naturalistic causal explanation. It seems to follow that the pure will satisfies the definition of transcendental freedom. It is governed by its own constitutive principle – a principle that it gives to itself in the sense that it arises from the nature of pure volition. And the capacity to follow this principle makes such a will ‘altogether independent’ of natural causality, since its reasons are not based on any interest that can be given a naturalistic causal explanation.

Though I cannot say for sure, this argument seems intriguingly close to succeeding.