Durham E-Theses

The sources of real agency: Kant on the metaphysics of freedom and life

RODRIGUEZ-GONZALEZ, ALVARO

How to cite:

RODRIGUEZ-GONZALEZ, ALVARO (2023) *The sources of real agency: Kant on the metaphysics of freedom and life*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/15367/

Use policy

This work is licensed under a Creative Commons Attribution Share Alike 3.0 (CC BY-SA)

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

The sources of real agency: Kant on the metaphysics of freedom and life

Álvaro Rodríguez-González

ABSTRACT

The role of Kant's Critique of Judgment in his overall theory of freedom has been largely confined to making room for the possibility of the realization of our ultimate moral goals in the world. While this is a central task Kant endeavoured to address with the third Critique, that work had a more overreaching impact in his general metaphysical picture of nature, which is instrumental in making sense of the core tenets of his theory of freedom. In this dissertation, I argue that Kant's solution to what I term the real agency problem, namely, how actions that are determined by the laws of nature can be nevertheless said to be freely caused by agents, is riddled with problems unless we incorporate the requirements brought by the power of judgment into it. I argue that, according to Kant, judgment compels us to posit a purposive ground of nature that accounts for the necessity of empirical laws, and in such a way that the causality of organisms can be made possible. By doing so, we gain a warrant to judge human beings as intelligibly grounded. Insofar as humans exhibit rational behaviour, this opens a space to our judging them as having an intelligible character which grounds actions as freely caused. Thus, the new insights provided by the third Critique are instrumental in understanding Kant's metaphysics of freedom.

The sources of real agency:

Kant on the metaphysics of freedom and life

Álvaro Rodríguez-González

Final dissertation presented for the title of
Master of Arts by Research in Philosophy

Durham University
Department of Philosophy
September 2023

CONTENTS

CONTENTS **3**

ACKNOWLEDGMENTS **4**

INTRODUCTION **5**

1. THREE PROBLEMS WITH FREEDOM **7**

§1. The purview of freedom 8

§2. Three separate problems 13

§3. The importance of real agency 18

2. KANT ON THE REQUIREMENTS OF FREEDOM **21**

§4. An absolute first beginning 21

§5. Why is an *automaton spirituale* not enough? 28

§6. Perspectival freedom 31

§7. Scofflaw freedom 36

§8. World-grounding freedom 39

§9. Action-grounding freedom 41

3. NO AGENCY WITHOUT EMPIRICAL UNITY **43**

§11. Agents in two realms 44

§12. The unity problem 49

§13. The scope problem 52

§14. The moral progress problem 54

4. THE INTELLIGIBLE GROUND OF NATURE **58**

§15. Making laws necessary 59

§16. What makes organisms special? 69

§17. How to encounter a free being 72

5. MAKING SENSE OF TIMELESS FREEDOM **79**

§18. Demystifying timelessness 79

§19. Whither go alternate possibilities? 83

§20. The epistemic boundaries of critical philosophy 86

CONCLUSION **90**

BIBLIOGRAPHY **91**

ACKNOWLEDGMENTS

There are many people whose help and support have been fundamental for finishing this dissertation. I want to express my gratitude to the audiences and participants at the several conferences and seminars where I have presented various stages of this dissertation: the Eidos Seminar and the first workshop of the Continental Research Group in Durham; the CHiPhi Work-in-Progress Meeting in York; and the Transcendental Idealism and Ethics conference in Trinity College Dublin. In particular, I want to thank, for their interest in conversation and their sharp objections, James Clarke, Pavel Reichl, Jim O'Shea, Patricia Kitcher, Tom Whyman and Andy Hamilton; crucially, this project would not have been possible without the support and keen help from my supervisor, Joe Saunders, whom I must thank especially. This dissertation has been the result of a year-long Master's by Research in Philosophy, fully funded by a scholarship for postgraduate study abroad from "La Caixa" Foundation.

INTRODUCTION

The famed conclusion to the *Critique of Practical Reason* is a rare poetic outlet for Kant, in which he avows his admiration for the two grand goals of the critical project. Amid sceptical threats, a solid ground must be found to save “the starry heavens above me and the moral law within me”. Securing the clockwork regularity of the empirical world and the unconditional bindingness of morality, that might well be the essence of Kant’s philosophy. If these are colossal feats on their own, they pose an especially difficult challenge when pursued side by side. The exigencies of natural science and ethics can easily clash when one expects so much as Kant did from both of them; particularly, they appear irreconcilable when the matter of freedom comes up. The notion that certain actions in the world should be understood to be under our control as free beings, and yet still belong to one and the same world, governed by unchanging laws, is a matter that has not ceased to haunt philosophers. Kant’s treatment of the problem, far from putting the debate to rest, moved baffled and fascinated philosophers alike into making sense of what threatened to be an intractable mystery.

As an acknowledgement of the uneasy coexistence of these two inextirpable interests of reason, Kant wrote the *Critique of Judgment*, with the declared intention of bridging the gap between the realms of freedom and natural necessity, and unifying them into the one world we inhabit. However, the idiosyncrasies of this complex text, with its analyses of art and biology, have meant a relative neglect in comparison to other works when it comes to studying Kant’s theory of freedom. In this dissertation, I take Kant’s promise at face value. I will argue that, in order to shed light on the many concerns that Kant’s readers have expressed regarding his unique theory of freedom, the *Critique of Judgment* must be given its due attention; it holds the key to understanding how the realms of nature and free agency may interact. The requirements that the possibility of empirical knowledge entails for the power of judgment, Kant argues, open the door to positing a determinate intelligible ground of nature. Given, further, that we as human beings are organisms – thus, mechanically inexplicable, per Kant’s doctrine –, we have a special warrant to judge that we belong to this substrate of nature. Once we have a way into positing a determinate ground of appearances where we, as humans, must feature, we find an apt place in our metaphysics to locate the special intelligible ground elicited by the demands of morality: the free ground of our actions. The demands of the power of judgment, thus, let us bridge what appears to us in experience and what morality demands that we find and realize in the world.

My argument will proceed as follows. In chapter 1, I will elucidate the problem around freedom that will be the focus of my research, which I term the *real agency problem*: how can we be free authors of phenomenal, law-bound actions in the world? I will argue that it is distinct from other relevant issues with freedom which Kant addresses, but still a central matter in Kant’s overall account. In chapter 2, I explore the requirements that Kant imposes to any acceptable account of freedom, and I argue for an “action-grounding” model. A free action is one which obtains in virtue of the right sort of intelligible ground. I explore other alternative readings and I highlight their drawbacks. Afterwards, in chapter 3, I contend that Kant’s notion of a free action requires that its author be an empirically unified being, a point that has largely been left unaddressed by the literature. Without it, I contend, key practices related to freedom – namely, the imputation of actions – cannot be sustained. In chapter 4, I lay out my core proposal. For us to find an empirical unity that has the right sort of intelligible grounds, while remaining within the boundaries of critical philosophy, we have to look into the *Critique of Judgment*. There, I contend, Kant argues that we need to posit an intelligible legal ground of nature in order to account for the necessity of empirical laws. Moreover, in order for us to make sense of organisms, we have to suppose that they are an objective part of this ground. Since we are organisms, we are entitled to assuming that we have an intelligible ground; as rational organisms in particular, we are the only ones that could plausibly have the right sort of ground for freedom. Finally, in chapter 5, I go over some potential objections, pertaining to the cogency of timelessness, the role of alternate possibilities, and whether this reading breaches Kant’s epistemic restraints.

1

THREE PROBLEMS WITH FREEDOM

A useful first step before exploring Kant's conception of freedom is clarifying what it is we actually mean to look into. Freedom is an equivocal term all through, one that spans from the political to the metaphysical, and which has historically deserved myriad qualifications[1]. This problem inherent to the topic is only compounded in Kant's case, mostly on two fronts. First, Kant makes a fair number of distinctions within freedom and its adjacent concepts: practical and transcendental freedom (KrV A 533/B 561; A 801-2/B 829-30), psychological freedom (KpV 5: 96), negative and positive freedom (GMS 4:446-7) or *arbitrium liberum* and *brutum* (KrV A 534/B 562) are some of his most notorious distinctions. How these interact with one another, and what their precise meanings are, have naturally been topics of lengthy discussion among Kantian scholars[2]. Second, Kant tackles freedom from his 1755 *Nova Elucidatio* all the way to the end of his career, with his 1797 *Metaphysics of Morals*. These four decades saw substantial changes both in emphasis and in doctrine regarding his treatment of freedom, and the exact extent to which these different texts conflict or agree with each other is highly disputed.[3]

In this preliminary chapter, I will narrow the aim of the rest of my enquiry by pinpointing one specific problem Kant's theory of freedom is meant to give an answer to, and by arguing for its significance. I begin (§1) by distinguishing three core problems Kant considers as falling under the umbrella of his theory of freedom, which I name the *volitional*, *teleological*, and *real agency* problems. I then (§2) go on to argue that, though closely interconnected, they are conceptually independent, and can merit separate discussions. I finally

---

[1] Some classical loci include Mill (2015 [1859]) and Berlin (1958).

[2] The distinction between practical and transcendental freedom has been particularly tricky, since the traditional reading finds an inconsistency in Kant's treatment of the pair *within* the first *Critique*. Practical freedom would rest on transcendental freedom according to the *Dialectic*, and it would be sufficient for morality and independent of transcendental freedom according to the *Canon*. See Allison (1982), Kohl (2014).

[3] (sources)I will focus throughout this work on Kant's critical (post-1781) and published theory of freedom, using pre-critical and unpublished texts only where they may help understand points left unclear by the latter oeuvre. I agree, furthermore, with the majority view that the Groundwork for the Metaphysics of Morals and the Critique of Practical Reason are a watershed for Kant's views on freedom, with the introduction of autonomy (Allison, 2020: 299; Timmermann, 2022). Nonetheless, as an interpretative stance, when in doubt, I will assume the reading that ensures the most consistent doctrine across works. Barred Kant's explicit disavowal of an opinion or evident contradictions, this method seems most apt to acquire as broad a picture as possible of a unified "Kantian theory of freedom". Otherwise, later works will take priority.

(§3) advocate, on textual and philosophical grounds, for the relevance of the real agency problem, which will be the focus for the following chapters.

§1. The purview of freedom

In discussing freedom, the stakes could not be higher for Kant. Freedom is the "keystone of the entire edifice of a system of pure reason" (KpV 5:3). Not only is it consubstantial with the possibility of moral standing, but it constitutes the sole pathway towards the enshrined objects of the old *metaphysica specialis*: the boundaries of speculative reason can only be extended to reach God and the immortality of the soul through freedom, by *fiat* of practical reason (KpV 5: 3-4; KU 5: 474**;** DO 8: 139). The heavy burden the notion of freedom bears in Kantian philosophy makes it, then, easy to lose sight of what "Kant's theory of freedom" may refer to at any given moment. In this section, in order to fend off those difficulties, I will identify three distinct problems[4] which are present in Kant's various discussions on freedom, and which delineate three complementary aspects of his overall theory.

As a starting point, it is hardly a matter of debate that freedom acquires its gravest relevance for Kant through the exigencies of morality. Prior to establishing the "supreme principle of morality" (GMS 4: 392), not even the possibility of freedom had been reached at (KrV A 558/B 586). It is only once the moral law is given to us that the reality of freedom is demonstrated in any way to us (KpV 5: 29-30, 47). It is well known, however, that this is not simply *one more* practical postulate, on the same standing as the belief in God and in the immortality of the soul[5]; freedom and the moral law entail one another, insofar as the moral law is the causality of a free will (GMS 4:446-7; KpV 5: 29). Freedom is not simply a requirement entailed by our moral commitments, as a more intuitive argument (for instance, from responsibility) would have it; a will that follows the moral law, according to Kant, is *nothing other* than a free will[6]; freedom and the moral law are inseparably interconnected.

---

[4] These are not meant to exhaust every matter Kant considers philosophically relevant around freedom. Importantly, for instance, I will not mention the key concept of Kant's doctrine of right: external freedom (MS 6: 230).

[5] The exact way in which the postulate of freedom differs from the other two, however, is not always clear. Kant himself groups them together as "hypotheses necessarily presupposed from a practical standpoint" (KpV 5: 132), but earlier on he had noted that "all other concepts (those of God and immortality), which as mere ideas remain without support in the latter [the idea of freedom], now attach themselves to this concept and with it and by means of it get stability and objective reality" (KpV 5: 3-4). Further, in the third *Critique* we find that "freedom is the only concept of the supersensible that proves its objective reality (by means of the causality that is thought in it) in nature, through its effect which is possible in the latter, and thereby makes possible the connection of the other two ideas to nature" (KU 5: 474).

[6] "In the concept of a will, however, the concept of causality is already contained, and thus in the concept of a pure will there is contained the concept of a causality with freedom, that is, a causality that is not determinable in

Moral worth, Kant argues, cannot come from any empirical object we seek in our acting, it belongs exclusively to the principles by which we determine our wills.[7] First of all, the worth of our actions does not depend on the felicity of the circumstances under which we act, as it would occur if value resided in the object we sought:

> "Even if it were to happen that, because of some particularly unfortunate fate or the miserly bequest of a stepmotherly nature, this will were completely powerless to carry out its aims; if with even its utmost effort it still accomplished nothing, so that only good will itself remained (not, of course, as a mere wish, but as the summoning of every means in our power), even then it would still, like a jewel, glisten in its own right, as something that has its full worth in itself" (GMS 4: 394).

Moreover, Kant contends that all such material principles for determining the will are ultimately reducible to the principle of self-love (KpV 5: 22), that is, to the strive for pleasure and for avoiding pain[8]. But then:

> "The matter of a practical principle is the object of the will. This is either the determining ground of the will or it is not. If it is the determining ground of the will, then the rule of the will is subject to an empirical condition (to the relation of the determining representation to the feeling of pleasure or displeasure), and so is not a practical law" (KpV 5: 27)

This is because practical *laws* are universal and necessary, whereas what causes pleasure or displeasure is subject-dependent and can only be discovered *a posteriori* (KpV 5: 21-2). These peculiarities in Kant's approach to ethics pose a special conundrum he needs to provide an answer to. How is it possible that our wills be determined by a principle that is independent from empirical conditions? That is, how can pure reason be practical? (KpV 5:3). Given that a will determined by pure reason alone is a free will, and the causality of a free will is the moral law given by pure practical reason, this question could be fairly rephrased as: *how is a free will possible?* This question, which ostensibly motivates – at least – the final sections of the first chapter in the *Critique of Practical Reason* (KpV 5: 42ff.), as well as, even more

---

accordance with laws of nature and hence not capable of any empirical intuition as proof of its reality, but that nevertheless perfectly justifies its objective reality a priori in the pure practical law" (KpV 5: 55)

[7] For a more detailed examination of Kant's argument against morality being grounded on material principles of the will, see Timmermann (2022, ch. 3).

[8] There are good reasons for finding this Kantian doctrine problematic (Sticker and Saunders, 2022). Some, however, have argued that utilitarianism (Korsgaard, 2008) or consequentialism in general (Stroppa, 2023) do rely on such a monistic conception.

clearly, the subsection "How is a categorical imperative possible?" in the *Groundwork* (GMS 4: 453ff.), is the first problem with freedom I wish to single out. Let us call it the *volitional problem* from now on.

That Kant does not rest his case on freedom by merely addressing the volitional problem is clear beyond a doubt in the second *Critique*. Once we discover how it can be the case that our wills be determined by a principle other than that of self-love – which is to say, a non-empirical principle (KpV 5: 29) – there still remains the matter of what it is a will so determined *wills*:

> "pure practical reason […] seeks the unconditioned for the practically conditioned (which rests on inclinations and natural needs), not indeed as the determining ground of the will, but even when this is given (in the moral law), it seeks the unconditioned totality of the object of pure practical reason, under the name of the *highest good*." (KpV 5: 108; cf. KU 5: 450)

We encounter thus a problem we did not have before. Per the demands of the moral law, we are obliged to strive towards the realization of a certain goal, the highest good. But unlike the determination of our wills according to the right principle, attaining the highest good is at odds with the possibility of a "stepmotherly nature". We somehow need a warrant to judge nature as amicable to the completion of our paramount moral goals. Kant's clearest discussion of this matter is to be found in the *Critique of Judgment*:

> "although the determining grounds of causality in accordance with the concept of freedom (and the practical rules that it contains) are not found in nature, and the sensible cannot determine the supersensible in the subject, nevertheless the converse is possible […] That which presupposes this *a priori* and without regard to the practical, namely, the power of judgment, provides the mediating concept between the concepts of nature and the concept of freedom […] for thereby is the possibility of the final end, which can become actual only in nature and in accord with its laws, cognized." (KU 5: 196)

How it is possible that nature, which certainly does not seem to be adequate in its laws for a goal such as the highest good (KU 5: 450), would nevertheless be so, is a further lingering puzzle. Kant finds it appropriate to speak of it as the problem of making "possible the transition from the domain of the concept of nature to that of the concept of freedom" (KU 5: 196), which is reasonable: if we explicate the possibility that our moral goals be effected in nature, we will have shown how it is that a concept made possible merely by the causality of freedom could take place. This is a problem regarding the realizability of freedom (as determinate moral legislation) in nature, and thus merits being the second in our list: the *teleological problem*.

At this point, we have encountered two distinct problems around freedom that, however, appear almost completely foreign to our common understanding of what is so pressing and haunting about our free wills. Jerry Fodor's oft-quoted remark is a colourful illustration of what is generally felt to behove a theory of freedom: "if it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching […] then practically everything I believe about anything is false and it's the end of the world" (1990: 156). For us to be free must mean that we are, in a significant manner, in control of our actions; that there is a substantial difference between what we *do* and what *happens* to us.

Kant too seems to believe that this is a relevant problem. Translating it into Kantian terms, the question is how it can be that actions are free when, as appearances, they must follow strictly the law of cause and effect. We will go over the full width of Kant's treatment of this point in the next chapter, but we can already point that this will amount to making sense of how actions can be, at the same time, phenomenally determined according to a natural law, and stand under an intelligible law noumenally. This is the crux of the Third Antinomy in the first *Critique*:

> "In respect of what happens, one can think of causality in only two ways: either according to *nature* or from *freedom*. The first is the connection of a state with a preceding one in the world of sense upon which that state follows according to a rule. […] By freedom in the cosmological sense, on the contrary, I understand the faculty of beginning a state *from itself*, the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature." (KrV A 532-3/ B 560-1)

The matter will be to establish how it can be that one and the same action can truthfully fall under both descriptions (KrV A 541/B 570), so it can be free while still residing in the empirical realm, the limits of which our theoretical cognition cannot trespass (KrV A 296/B 352-3). There is nothing we can do, after all, with actions that occur completely beyond the sphere of our possible knowledge[9]. This worry, moreover, is not limited to the not-quite-mature account of freedom found in the *Critique of Pure Reason*; Kant is insistent that the first *Critique* discussion is instrumental to making sense of his later developments:

> "if one still wants to save [freedom], no other path remains than to ascribe the existence of a thing so far as it is determinable in time, and so too its causality in accordance with the law of *natural necessity, only to appearance, and to ascribe freedom to the same being as a thing in itself*. This is certainly unavoidable if one wants to maintain both these mutually repellent concepts together; but in application, when one wants to explain them as united in one and the same action, and so to explain this union itself, great difficulties come forward, which seem to make such a unification unfeasible." (KpV 5: 95)[10]

Hence, it seems like there is a third problem Kant acknowledges as relevant with regards to freedom, which I will call the *real agency problem*. Let us recap, then, the three problems we have spotted throughout Kant's discussion of the general matter.

> **Volitional problem:** How can our wills be determined by a non-empirical principle? i.e., how can pure reason be practical?
>
> **Teleological problem:** How can the ultimate goals of morality be realized in the world? i.e., how can the legislations of nature and freedom be harmonious?
>
> **Real agency problem:** How can we be free authors of phenomenal actions? i.e., how can an action be both freely caused and remain under the laws of nature?

I have shown that these three problems appear at different points of Kant's overall treatment of freedom, but do they stand on the same level? It could so happen that some of these issues are, on closer inspection, pseudo-problems, reducible to a more fundamental question. In the next section, I will argue that, even though these three problems are deeply interconnected, such that they cannot be answered in isolation, they are conceptually independent in a relevant sense: solving any of them on its own will not immediately yield a solution to any other.

---

[9] Cf. GMS 4: 387. "[N]atural as well as moral philosophy can each have its empirical part, since the former must determine the laws for nature, as an object of experience, the latter for the human being's will, in so far as it is affected by nature".

[10] "[C]onsiderations of this kind, including those that are once more directed to the concept of freedom, though in the practical use of pure reason, should not be regarded as interpolations which might serve only to fill up gaps in the critical system of speculative reason (for this is complete for its purpose), or as like the props and buttresses that are usually added afterwards to a hastily constructed building, but as true members that make the connection of the system plain, so that concepts which could there be represented only problematically can now be seen in their real presentation." (KpV 5: 7)

§2. Three separate problems

We have already briefly touched on the peculiarities of the volitional problem as a problem for Kant, but it is important to revisit them in order to emphasize the core difference between that first problem and the other two. Indeed, at first glance it may appear that the volitional problem, with all its talk about the "causality of reason" and the analogies between the moral law and natural law (GMS 4: 446; KpV 5: 42) should give a sufficient answer to the whole ordeal. However, as I am about to argue, it does not and it cannot.

Note that the defining fact for both the teleological problem and the real agency problem is that they are concerned with states of affairs in the phenomenal world. The former asks how a certain state of things, one we are obligated to have as our goal, can obtain in the world; the latter, how a certain state of things that does obtain can be said to be under our free control. There are reasons, conversely, to believe that the volitional problem posits nothing of the sort, and that its being addressed is independent of anything transpiring amid appearances. While it is quite clear that the teleological problem is an independent problem, since it asks about the ultimate success of our moral goals – and Kant gives extensive separate discussion to that regard – it may be more doubtful that the volitional problem and the real agency problem are not the same; this is what I will devote most of this section to elucidate.

To begin with, we must understand why the volitional problem is a problem at all. Famously, Hume held that reason on its own cannot determine any volition in us: "'Tis from the prospect of pain or pleasure that the aversion or propensity arises towards any object: And these emotions extend themselves to the causes and effects of that object, as they are pointed out to us by reason and experience" (*Treatise* 2.3.3)[11]. If the role of reason goes no farther than letting us know the causal connections that ultimately resolve in an object of our volition, reason itself is motivationally inert.[12] Kant attempts to show that there is, indeed, such a thing as a principle of pure practical reason, namely, the moral law, expressed to us as the categorical imperative (GMS 4: 421; KpV 5: 30). We could face a problem before enjoining the bare possibility that we would act on the principles of pure reason alone if no such principle could exist[13].

---

[11] Moreover, Kant explicitly argues against the realist Pistorius, and in favour of formalism, in the second *Critique*. For a discussion of Kant's assessment of Pistorius's position, see Timmermann (2022, ch. 3).

[12] For modern critiques against this kind of scepticism, see Korsgaard, (1996, ch. 11; 2008, ch. 1).

But even if we venture beyond the mere existence in theory of principles of practical reason, the volitional problem poses a cogent matter with no empirical commitments; namely, how can it be the case that these are principles *we* are under? Or, how is it that pure reason is practical *for us*? I contend that Kant's arguments both in *Groundwork III* and in his discussion of the "fact of reason" in the *Analytic of pure practical reason* aim primarily at this, letting us isolate a problem about freedom independent from its link to any appearances. While it is outside the scope of this discussion to elucidate the precise nature of these highly contentious sections of Kant's philosophy, a very brief survey will suffice to show that nothing phenomenal hangs on the question they raise.

Starting with *Groundwork III*, I follow Saunders (2021) in reading Kant's argument as trying to dispel what doubts we may legitimately have regarding the possibility that reason (and, moreover, pure reason) be practical. Kant would be advancing a non-moral argument that, insofar as we experience ourselves as *rational beings*, and insofar as freedom has been shown to not be theoretically contradictory, we can regard ourselves as possibly being moved by reasons alone, which ultimately was the doubt besetting both kinds of practical reason. Kant's argument, then, gives reasonable grounds to hold (i) that rational beings in general must be free, and (ii) that, in being free, their pure reason alone must be able to determine their wills. This applies to us *qua* rational beings, and it is meant to extend without loss of generality to purely rational beings as well (GMS 4: 447). It is insofar as we can legitimately consider ourselves to be members of the intelligible world that these observations hold (GMS 4: 452).

If this is so, Kant has discovered a property of a certain kind of intelligible being, and he has given reasons for us to hold that we are among those beings, thus proving – granted that his argument works, which need not concern us here – that we are free and stand under the moral law. But, crucially, at no point does Kant concern himself with whether any phenomena will conform to the moral law and by which mechanism it would do that. Kant is explicit that, in the *Groundwork*, (i) he is abstracting from everything empirical (GMS 4: 390), (ii) that it could well be that nothing ever had conformed to moral law without hurting his arguments in the slightest (GMS 4: 406)[14], and (iii) that the sort of bindingness he is concerned with here regards not what occurs but what *should* occur (GMS 4: 408). By pressing the point that we are rational beings, and that rational beings are free and under the moral law, Kant gives an answer to the volitional problem, that is, to whether we stand under the moral law and how we get there (Kryluk, 2017)[15] without yet committing to any phenomenal happenings.

---

[13] Korsgaard (1996: 311) distinguishes content and motivational scepticism; similarly, Galvin (2019) argues that the fact of reason at times focuses on our awareness of the content of the moral law, whilst, other times, it refers to its being authoritative for us. Since the "motivational question" – to establish that we have an *interest* in morality – is treated separately by Kant to that of the practicality of pure reason in each case (GMS 4: 459-60; KpV 5: 71-89), and since it can be construed to have a bearing on the possibility of our actual phenomenal following of the law, depending on how we decide to approach it may fall on the side of the volitional problem or of the real agency problem. Due to this fuzziness, I have decided not to tackle the matter – important as it is – upfront. As we will see later, solving the volitional problem *demands* in a certain way solving the real agency problem; and the real agency problem is only elevated to the status of a real problem if the volitional problem has been dealt with. They are, thus, deeply interconnected though independent (in the aforesaid sense by which solving one does not do out with the need to solve the other), so it should not come as a surprise that certain matters may have an ambiguous relation with one and the other.

Much of the same can be said about the second *Critique*'s "fact of reason" route. Once again, it is important to emphasize that Kant's explicit purpose in this book is "to show *that there is pure practical reason*" (KpV 5: 3)[16]. We should then expect that if the fact of reason is to play a decisive role in the argument here, it will be by warranting the practicality of pure reason. The specific status of the fact of reason is, as expected, widely debated[17]; but all we need to find out for our purposes here is whether it has a role in answering the volitional problem as a distinct problem. Bypassing, then, concrete disputes around the fact of reason, we can look at the broad strokes of Kant's assessment. The controversial *Factum* enters the picture in order to ascertain "from what our *cognition* of the unconditionally practical *starts*, whether from freedom or from the practical law" (KpV 5: 29). Freedom would be a non-starter, since it cannot be the object of a possible experience, so, as preluded several pages earlier by Kant, the moral law will be the *ratio cognoscendi* of our freedom. In whatever form it may take, and whatever specifically *about* the moral law it refers to, "[c]onsciousness of this fundamental law may be called a fact of reason" (KpV 5: 31), and it must, somehow, take us to establish the reality of freedom.

The way in which it might do so is laid out in *Problem I* of the *Analytic* (KpV 5: 28-9), where Kant proves the first side of the biconditional between a will determined by the moral law and the free will; this is what we need in order to go from our awareness of the moral law to our awareness of freedom. This follows, Kant says, because a morally determined will is not determined by anything in nature; as such, a morally determined will is independent of all empirical conditions, and therefore, free. If Kant's argument works, and our consciousness of the unconditional practical law suffices to assert the reality[18] of a will independent of empirical necessitation, then he has answered the volitional problem. But has he not, in one stroke, also solved the real agency problem? After all, the famed gallows example says:

---

[14] Some authors have gone even further in claiming that action may not need any sort of confidence in the hospitability of the world to their realization. That is clearly not Kant's view, per the teleological problem, but it helps the case that there we could isolate such a problem. See Freyenhagen (2020).

[15] Kryluk sustains that Kant ultimately backs off from this attempt, but his distinction between the *how* and the *whether* of morality to characterize Kant's analytic and synthetic methods respectively is helpful here.

[16] "The function of the fact of reason in the Critique is to do nothing less than establish the central thesis of the text, that pure reason can be practical through a priori concepts" (Kryluk, 2017: 717).

[17] For recent commentary, see, among others, Ware (2014), Kryluk (2017) and Galvin (2019).

> "But ask him whether, if his prince demanded, on pain of the same immediate execution, that he give false testimony against an honorable man whom the prince would like to destroy under a plausible pretext, he would consider it possible to overcome his love of life, however great it may be. He would perhaps not venture to assert whether he would do it or not, but he must admit without hesitation that it *would be possible* for him. He judges, therefore, that *he can do something because he is aware that he ought to do it and cognizes freedom within him*, which, without the moral law, would have remained unknown to him. [emphases added]" (KpV 5: 30)

We could follow Galvin (2019: 33) in holding that "Kant's claim is not that such deliberations establish that one's will is as a matter of fact independent of empirical conditions, but rather that when I take moral considerations as authoritative, I conceive of myself that way". But even if our reading of the fact of reason is less deflationary, we can show that the gallows example does not solve the real agency problem. Let us say that this, in fact, proves that we *must* be capable of being truthful. First of all, it is left unresolved in *what sense* we are capable of this. It surely will not be a matter of successful phenomenal execution, since this lies outside of the boundaries of a critique of practical reason (KpV 5: 45-6; 66), contingent on the empirical situation as it will be (we may be tricked into giving a statement that will then be used against that person, etc.). What is more, it still says nothing about *how* it could be that we would be in free control of a phenomenal instance of lying, given that phenomena are subject to the laws of nature. Under the most generous reading of this solution to the volitional problem, we are given a warrant and expectation to solve the real agency problem. If it so happened, however, that there was absolutely no way to account for our freely controlling natural phenomena, then either we would have to resort to a deflationary solution to the volitional problem, or we would need to admit that, although there conceptually exists a solution to the practicality of pure reason, it is a chimera with unfulfillable requirements.

---

[18] The sense in which we take "reality" here can vary significantly, and it should not affect my argument. Ware (2014) argues that it is a "reality" founded on practical grounds, distinct from the theoretical standpoint that would motivate a sceptic; Kryluk (2017) considers that Kant is not addressing sceptics at all, rather showing that morality is self-justifying and cannot be grounded on anything external. But whichever sense it is, insofar as Kant rested content with his answer, it must have been the right sense of "reality" he was after when he set out the problem of proving that pure reason can be practical.

I submit, then, that the volitional problem and the real agency problem are independent, though closely interconnected, problems. Once the volitional problem is solved, we had better find a solution to the real agency problem; similarly, if the volitional problem is not solved, the real agency problem loses its relevance. Kant, of course, had already sketched an approach to the real agency problem, as we saw in §1; it is the point of the resolution to the Third Antinomy, which resurfaces in the *Remark to the Analytic of pure practical reason*.

A final objection to the independence of the real agency problem merits brief comment before moving on to the next section. It may be conceded that it cannot be reduced to either the volitional problem or the teleological problem on their own, and yet that, if those two are solved, then there is no space left for the real agency problem to cover. We may find a position akin to this one in Korsgaard's *Self-Constitution*, when she addresses our "attitude of trust towards the world", which she explicitly links to Kant's moral religion:

> "We cannot regard ourselves as agents, that is, as the causes of certain effects through our wills, if in fact our wills have no power at all to make our effects be the ones that we will. And yet we must regard ourselves as agents, that being our situation, and not negotiable, for to be human is to have no choice but to choose […] this means that, just as the speaker is forced to take up an attitude of trust towards her hearer, so the agent is forced to take up an attitude of trust towards the world itself". (Korsgaard, 2009: 87)

This passage may point to two different things. On the one hand, what type of assent backs our claim that we can freely control actions, and whether it is closer to knowledge or to faith. That is a very important matter that I will revisit in due time (§20), and which asks about the specifics of our response to the real agency problem rather than nullifying it. On the other hand, however, this could be construed as collapsing the real agency problem into the volitional problem and the teleological problem. If we must have our goals realized, and if we must hope for the world to be in such a way as to have our goals realized, then this may as well apply to every felicitous use of our free will into the world. But that is not so. The real agency problem is concerned specifically about *how* a naturally determined action can yet be said to be freely caused by an agent. If we merely focus on the teleological problem, which regards those goals that we *must* have, either we need to add in an account of how the ultimate goal that must be realized in the world has to work through our individual actions – or, what is the same, we need an answer to the real agency problem, since we still need to single out actions as freely caused by agents – or we erode all individual action from our view of freedom. In the next section, I will argue that this last option is unacceptable.

§3. The importance of real agency

One might still have some doubts that the real agency problem as I have characterized it is a proper problem in Kant's theory of freedom. In this final section, I briefly go over three reasons why we cannot simply eliminate the role of individual actions being freely caused in Kant's overall account. First of all, I argue that even a minimal version of pure practical reason will raise questions tackled by the real agency problem, on account of Kant's psychological determinism. Then, I countenance that it is key for Kant's own account of responsibility as well. Finally, I contend that, without it, Kant's theory would lose the bite required for any satisfactory philosophical account of freedom.

In §1, I surveyed some fragments from the first two *Critiques* where Kant seems explicit about the need to explain how actions can be both under the influx of natural causation and yet be the effect of transcendental freedom. This alone should make the case for the relevance of the real agency problem in Kant's account. However, one may be tempted to downplay its importance. After all, if Kant's ethics are so insistent about the ultimate irrelevance of successful execution for moral worth, maybe the results of the *Third Antinomy* are merely meant to open up a space for the possibility of a pure practical reason, and any mentions to freely caused phenomena are superfluous. Now, let us assume that this is so. We know that, at the very least, there must be a sense in which we regard our refusal to tell a lie as a possibility. This is as perfect as a duty gets (MS 6: 429); it does not require anything in particular to happen in the world, since it is a negative duty. However, we must remember that our mental states too are phenomenal; as such, they are temporal, and thereby subject to natural causality[19]. Then, our being free in refusing to tell a lie will require us to make sense of phenomena being both naturally determined and freely caused, which is precisely the real agency problem. Therefore, even a minimal version of the practicality of pure reason forces us to confront the real agency problem; the problem is completely general as it pertains to the interaction between a deterministic nature and a free ground of actions.[20] For us to refuse to lie, something must have caused that refusal, that something must have a cause, and so forth. In other words, even if the practicality of pure reason can be established independently of any and all empirical events, we cannot make sense of the least actual adherence to the prescriptions of the moral law without it having a phenomenal impact[21]; any possibility of abiding by the moral law, of making the autonomous law of reason have any bearing on our actions, requires having accounted for the real agency problem. It is in this sense that the volitional problem, once solved, begs for a solution to the real agency problem.

---

[19] This is extremely controversial (see Frierson (2014: ch. 1) for an overview of the debate). Many commentators, in fact, do not consider the inner sense to be deterministic in Kant's account. For reasons I will explore in §7, I consider Frierson to be correct in claiming that our mental states *are* deterministically ordered according to Kant. Let us assume, however, that the inner sense is not deterministic. Insofar as we have a will, we can cause effects through our representations in the external world (KpV 5: 8)**.** If these representations are indeterministic and free, then we have a free cause having effects in the actual world, which, if the unity of experience is to be preserved (KrV A 216/B 263), abides to natural laws. Therefore, even if we consider the inner forum to be unproblematically indeterministic, the question of how it can be that we cause things freely in the world despite its obedience to the laws of nature – the real agency problem – still holds.

Furthermore, refusing to give significance to the matter of phenomenal actions being freely caused would deal a fatal blow to the ethical purport of Kant's account of freedom. Indeed, being able to single out actions as freely caused is part and parcel of our practice of imputation, according to Kant:

> "An action is called a *deed* insofar as it comes under obligatory laws and hence insofar as the subject, in doing it, is considered in terms of the freedom of his choice. By such an action, the agent is regarded as the *author* of its effect, and this, together with the action itself, can be *imputed* to him, if one is previously acquainted with the law by virtue of which an obligation rests on these." (6: 223)[22]

Naturally, we encounter the actions we judge in experience and nowhere else. Thus, if we fail to make sense of the notion that phenomenal actions can be freely caused, then Kant's doctrine of imputation, central as it is to any workable ethics, becomes groundless and incomprehensible.

Finally, it simply seems unavoidable for a theory of freedom to, at least, attempt to tackle the matter of how actions in the world may be freely caused. If we dodge the issue, our theory arguably becomes irrelevant in our actual practices. Freedom would be a purely virtual philosophical concept, useless in attributing actions to agents (O'Connor, 2000: 26) and in telling us what relevant features distinguish certain beings *as we encounter them* as potentially free[23]. If freedom is to play a role in our practical philosophy, it cannot be completely detached from things as they happen in the world. Aside from the mounting textual evidence, if Kant wanted his theory of freedom to be successful at what we legitimately expect it to be, he simply could not have avoided the real agency problem.

---

[20] Does this undermine my previous argument that the volitional problem and the real agency problem are distinct? No. the volitional problem in any case makes it necessary for us to find a solution to the real agency problem without immediately providing one.

[21] For further discussion about the sense in which the prescriptions of morality must be possible for us even in situations of apparent unfeasibility, see Rodríguez-González (2023).

[22] Cf. "*Imputation* (*imputatio*) in the moral sense is the *judgment* by which someone is regarded as the author (*causa libera*) of an action, which is then called a *deed* (*factum*) and stands under laws." (Coll. 27: 288)

In sum, I have argued that Kant distinguishes, at least, three independent yet closely interrelated problems related to freedom. One of them, the real agency problem, asks how it can be that phenomenal actions, which are determined by natural necessity, can nevertheless be free. This will prove to be quite a challenge. It will be the task of the remaining chapters to investigate how Kant's philosophy goes about meeting it.

---

[23] See Saunders (MS: 12) for a discussion of this point in regard to so-called "two standpoints interpretations" of Kant's theory of freedom.

2

KANT ON THE REQUIREMENTS OF FREEDOM

The main selling point of Kant's theory of freedom may well be that it promises the strictest form of freedom while not forfeiting a deterministic view of nature that allows for the greatest successes in the empirical sciences. Wood famously called this stance the "compatibility of compatibilism and incompatibilism" (1984: 74). Such big promises, however, will need some substantial backing. It will be the task of this chapter to lay out the specific model by which Kant keeps to his word in responding what we have been calling the *real agency problem*[24], that is, how phenomenal actions can nevertheless be free.

I will begin (§4) by analysing the Third Antinomy in the *Critique of Pure Reason*, elucidating Kant's concept of transcendental freedom and how it can be possible without contradiction in our world. We will see that transcendental freedom is an absolute ground of determination for a causal chain, and one that can exist thanks to transcendental idealism, by intelligibly grounding our empirical character. Then (§5) I will argue that Kant is justified in requiring such freedom for our actions, on at least three grounds: it can bring moral normativity into the world of appearances, it is needed to ground the imputation of actions, and it alone can bear the strictures of Kantian ethics, making room for the possibility of moral worth. After this, I will address several influential interpretations of Kant's theory of freedom to conclude three conditions that a successful reading must comply with. (§6) Free empirical actions must stand in a real relation with their intelligible ground, (§7) they cannot breach the laws of nature, and (§8) our intelligible character must ground our actions alone, not the entire world. Finally, (§9) I recap my reading and I briefly remark that it does not commit me to any specific side of the debate on the ontological status of phenomena and noumena.

§4. An absolute first beginning

In order to understand the critical Kant's notion of freedom, it will help to start at the Third Antinomy. Let us briefly recall the general concept of an antinomy according to Kant. The categories are logical functions of judgment that bring intuited manifolds under the unity of apperception (KrV B 143); the diverse in sensibility is brought before us as an object of experience, according to Kant, through their synthesis in understanding, which ensures the validity of the categories of understanding – as the precise functions through which this synthesis is possible – in their application to objects of experience (KrV B 152). Now, the proper scope for the objective applicability of categories are objects of empirical intuition (KrV B 148), beyond which these pure concepts only serve to determine objects of mere *thought*, not of cognition (KrV B 146)[25]. Now, reason, says Kant, tends to seek the universal condition of whatever conditioned it may encounter (KrV A 307/ B 364). The image Kant makes use of when characterizing reason is that of the syllogism. Whatever we find conditioned in experience must be so conditioned because there exists a major premise under which, by means of fulfilling its condition, it can be subsumed. For example, by pressing this key on a piano I produce a B flat. I know this because it is a case of the general principle, "whenever this key is pressed, a B flat is produced". Reason, according to Kant, compels us to go further in seeking universal conditions *pro prosyllogismos*, that is, by treating the general principle we know as the conclusion of a more general syllogism; in this case, maybe, something like "whenever a piano key is pressed, a note is produced in accordance with the length and tuning of its corresponding string". And so on.

---

[24] From now on, if left unspecified, problems "around freedom" refer to the real agency problem in particular.

The subreption that generates what Kant calls a "natural and unavoidable illusion" of reason (KrV A 298/ B 354) occurs when we take the universal condition set as a task by the syllogistic chain built atop the conditioned in experience as if it were already a given. Specifically, in the case of antinomies, we assume that the absolute condition of the synthesis of appearances through categories is given to us (KrV A 407-8/ B 434). What makes the antinomies tricky is that they generate a contradiction within reason itself, because we always have two options in determining what this unconditioned in the synthesis of appearances may be:

> "Now one can think of this unconditioned either as subsisting merely in the whole series, in which thus every member without exception is conditioned, and only their whole is absolutely unconditioned, or else the absolutely unconditioned is only a part of the series, to which the remaining members of the series are subordinated but that itself stands under no other condition." (KrV A 417/ B 445).

In the First Antinomy, which corresponds to the categories of quantity, we find ourselves grappling with whether spacetime is finite or not. The idea is that any bounded region of space and time (and every appearance must fall under such a discrete region, per the axioms of intuition (KrV A 161/ B 202)) is conditioned by the surrounding space in one case, and by the preceding time in the other (KrV A 413/B 440). Reason would be, so to speak, content if it found a universal term in that series of conditions, which can either be a member of the series (a limit beyond which there is no more space or time, i.e., spacetime is finite); or the series itself (which is then infinite). However, it can be proven that both are wrong and lead to contradiction: these proofs are the thesis and antithesis comprised by the First Antinomy (KrV A 426/B 454 ff.)

---

[25] Of course, not just those objects that as a matter of fact are in our perception, but all those that are legally connected to our actual experience (KrV A 218/ B 266).

The idea of freedom arises in the Third Antinomy, corresponding to the categories of relation, and thus seeking unconditional unity in *causality* amid appearances. The problem here has a long history[26] traditionally framed along the lines of "everything that happens has a cause, and, either the chain of things that happen needs to finite and thus there are unmoved movers, or the chain is infinite", followed by the problems entailed by one or the other alternative. Let us, however, take a closer look at Kant's own exposition of the matter, so we are not carried away by the received connotations of "unmoved movers". Indeed, Kant's talk of "freedom" in what appears to fall under the classical banner of the *kalam* argument may seem disconcerting at first. Kant, however, is no longer concerned at this point with the creation of the universe or any such thing; that would fall under the First Antinomy, as it regards an absolute beginning *of* time[27]. Here, rather, we are to think about the conditions of the causality that unites appearances. When A causes B, Kant gathers, that is due to a rule by which B as an occurrence must follow from A's obtaining. Now, this rule's holding may depend on a further cause or not. If the former, then it is a case of "causality in accordance with laws of nature" (KrV A 445/B 473). Otherwise, it is transcendental freedom (KrV A 446/B 474). The first option – freedom – pinpoints a member of the chain as the absolute condition of its subsequent effects; the second one – universal determinism – resolves to consider the whole, infinite causal chain as the unconditioned, thus leaving every single member of it as individually conditioned.

---

[26] Perhaps most prominently, Aristotle (*Met.* Λ.6.) and Thomas Aquinas (*Summa Theologiae*, Pars Prima, q. 2 art. 3).

[27] Indeed, Kant attributes the analogy between the two antinomies to the defender of the antithesis, or "omnipotence of nature" (KrV A 449/ B 477). The defender of the thesis rather than being accused of postulating an unmoved mover is characterized as asserting the existence of "fundamental properties (fundamental powers)". First movers are mentioned (KrV A 450/ B 478), but they are ostensibly a subset of these fundamental powers; if there is only one uncaused cause in the world, it will be the unmoved mover, but so far Kant has not presented any argument as to why there should only be one.

If something is determined to cause an effect by an external condition, that something is unfree (at least, unfree in causing that effect); on the other hand, if something can cause an effect through no other condition, but through “absolute causal spontaneity beginning from itself”, then it is transcendentally free. Freedom, in the transcendental sense, is inextricably linked to the specific issue the antinomies are preoccupied with: seeking the absolute condition of what is conditioned in experience. But while in *kalam* arguments we are concerned with finding a *creator*, the first being who was not itself created but rather created everything else, in the Third Antinomy the question is whether every member of a chain of causes and effects was *conditioned to* cause something or whether there could be a member that was the absolute beginning of the said causal chain. Say a murderer comes to my door and asks for the whereabouts of his victim. Following the advice of John Henry Newman[28], rather than telling him a lie, I take it upon myself to knock him out. In his last conscious seconds, he tries to stay on two feet looking for something to grasp onto, but instead tosses a flowerpot, which wakes my dog up, causing him to attend the scene not too happy about the noisy stranger loitering about his home. All this slapstick revolves around the following causal chain:

> I knock out the murderer → The murderer tosses a flowerpot → The flowerpot shatters → My dog wakes up.

The relevant question at this is not whether I was “of woman born”[29], or whether I have existed for all eternity[30]. None of the subjects of these causal events – myself, the murderer, the flowerpot, etc. – came into existence during this episode. Rather, the relevant question is whether I was determined by anything else *to knock out the murderer*[31], or whether my action was the absolute beginning of this causal chain[32]; this is the sense in which transcendental freedom is one of the two possible inferences of reason when dealing with the ultimate conditions of causality. As Kant's own example in the Remark on the Thesis shows:

---

[28] I found this example originally in Korsgaard (1996: 157 n. 10).

[29] Kant says more than once that “if [the preceding state] always existed, [it] could not have produced any effect that first arose in time” (KrV A 532/B 560). This, however, refers to *states* of the agents in question. If it had always been the case that I knocked out a murderer, then it could not transpire *just now*.

[30] This is not to say that it could not be the case that only the creator of the world should be free, as Kant himself acknowledges (KrV 449/B 478). But, as he is quick to clarify, that is not the only way things can be. It could also be, unelicited as such a step would be at this point, that we are free, since “here we are talking of an absolute beginning not as far as time is concerned but as far as causality is concerned” (KrV A 451/B 479).

[31] Kant’s language is quite clear, he does not focus on the free agent in the exposition of the thesis and antithesis so much as in the freely caused happening: “a causality must be assumed through which something happens *without its cause being further determined by another previous cause*” (KrV A 446/B 474; emphasis added).

[32] One may object that this relies on an idiosyncratic reading of Kant’s theory of causality as powers-based rather than an events-based one. While I am sympathetic to Watkins’s (2005) argument in favour of reading Kant as a powers theorist, the point can easily be translated into either model. If we adopt an events model of causality, the question boils down to whether “my knocking out the murderer” was caused or uncaused; if we adopt a powers model of causality, the question is whether I was determined to exercise my power to knock out the murderer or not.

> "If (for example) I am now entirely free, and get up from my chair without the necessarily determining influence of natural causes, then in this occurrence, along with its natural consequences to infinity, there begins an absolutely new series, even though as far as time is concerned this occurrence is only the continuation of a previous series" (KrV A 450/ B 478)

Now, we have merely set up the issue. There is an argument against the thesis as much as there is one against the antithesis. How do we go on from here? Is it possible that I could be completely spontaneous in beating that noxious fellow, or was I completely determined by external factors into doing so? The conflict in the case of the First Antinomy is solved by declaring both the thesis and the antithesis to be false; neither a finite nor an infinite spacetime could be objects of a possible experience, they are mere thought-objects. We can only fit an indefinite empirical regress in our experience, by which we know that we can always keep looking for more space and more time (KrV A 519/B 547). This is the only available response because the First Antinomy is concerned with what Kant calls a *mathematical synthesis* of appearances, that is, the members of the series of conditions can be regarded as conditions and as conditioned *qua* homogeneous members. This region of space is bounded, and thus conditioned, by its surroundings because they are all likewise parts of one space. As such, the conditional chain needs to stay squarely in the phenomenal realm (KrV A 529/B 557), and, therefore, the answer must conform to what may be given as an object of possible experience.

However, in the Third Antinomy that is not quite so. Causality is a *dynamical* category, which synthesizes heterogeneous relata (KrV A 530/ B 558); A causes B not by being the same kind of thing as B. From there, Kant claims that his phenomenon/noumenon distinction will suffice to give a more positive response to the antinomy around freedom: transcendental idealism allows us to hold both the thesis and the antithesis to be true; everything is determined by another cause, and yet there could be freedom – the former, as it pertains to sensible (phenomenal) conditions; the latter, as it pertains to intelligible (noumenal) ones (KrV A 531/B 559).

How can this be? To begin with, Kant reaffirms what he established with the Second Analogy. Every appearance, insofar as it is in time, and thus must stand under an objective succession, must also have a cause in time (KrV A 199/ B 244). Now, for every empirical cause, its causality – its effectively bringing forth an effect – is something that also arises in time, and it is thus as determined by an empirical antecedent as anything else in experience. In that sense, everything that we may perceive as a cause will be determined *to so be* by something prior:

> "The law of nature that everything that happens has a cause, that since the causality of this cause, i.e., the action, precedes in time and [...] cannot have been always must have happened, and so must also have had its cause among appearances [...] is a law of understanding, from which under no pretext can any departure be allowed or any appearance be exempted" (KrV A 542/ B 571)

As such, and as already shown in the general discussion of the antinomies, we cannot have in experience anything like an absolute beginning (KrV A 533/ B 561). When we describe the effects of any being in the world, we have to enquire into their empirical characters, that is, the general and constant rule by which they produce certain effects when they are affected by certain conditions (KrV A 549/ B 577). If appearances were things in themselves, that would be it[33]; there would be no further space for freedom as absolute beginning (KrV A 536/ B 564). Nonetheless, per transcendental idealism, they are not (KrV A 540/ B 568). When we stick to empirical explanation, we simply abstract from the fact that appearances must presuppose a ground that is not itself an appearance without annulling it:

> "For if we follow the rule of nature only in that which might be the cause among appearances, then we need not worry about what sort of ground is thought for these appearances and their connection in the transcendental subject, which is empirically unknown to us." (KrV A 545/B 573)

Kant, then, proposes the following experiment. Let us suppose that there are beings with a faculty to absolutely begin a state, what would that look like?[34] As far as we can experience their actions, they would still need to have an empirical character, and, what is more, if we could investigate them in due depth, we would be able to predict their every move (KrV A 550/ B 578). But if we have a reason to venture as to what ground a certain occurrence may have beyond appearances – which we do not always have (KrV A 546/ B 574) –, then we can posit an intelligible character to these beings. These are, likewise, rules of causality, but whereas empirical characters link preceding states to effects in time, intelligible characters determine the causality that connects appearances together – the empirical character – itself:

---

[33] Strictly speaking, if appearances were things in themselves, the antinomy would be unsolvable and there would be no nature or freedom.

[34] "...supposing also that it is in any case merely invented, if one assumes that among natural causes there are also some that have a faculty that is only intelligible..." (KrV A 545/ B 573). Also, some paragraphs later, "Now let us stop at this point and assume it is at least possible that reason actually does have causality in regard to appearances" (KrV A 549/ B 577).

> "Is it not rather possible that although for every effect in appearance there is required a connection with its cause in accordance with laws of empirical causality, this empirical causality itself, without the least interruption of its connection with natural causes, could nevertheless be an effect of a causality that is not empirical, but rather intelligible[?]" (KrV A 544/ B 572)
>
> "...the intelligible character, which is the transcendental cause of the former [the empirical character], is passed over as entirely unknown, except insofar as it is indicated through the empirical character as only its sensible sign" (KrV A 546/ B 574)
>
> "...the intelligible character, of which the empirical one is only the sensible schema..." (KrV A 553/ B 581)[35].

Every action remains, as appearance, neatly integrated into the thoroughgoing unity of experience, but some effects can be regarded as nevertheless freely caused, and, thus, as having an absolute beginning in a "different relation" insofar as the empirical character that grounds the fact that the effect follows from the cause is, *itself*, grounded in "certain conditions that [...] would have to be regarded as merely intelligible" (KrV A 545/ B 573). For example, say Shylock seeks revenge against Antonio. As appearance, we can investigate Shylock's empirical character, which may be, for instance, modelled as a conjunction of conditional statements:

> **Shylock's empirical character**: (1) if Shylock is pricked, then he bleeds; and (2) if Shylock is tickled, then he laughs; and (3) if Shylock is poisoned, then he dies; and (4) if Shylock is wronged, then he revenges; and...

These are all true causal statements about Shylock. Of course, (1) through (3) are not voluntary actions, and so it would appear that the explanation at the level of appearances here suffices[36]. Conversely, we do regard Shylock to be free in case (4). We, then, assuming that we are authorized in regarding him as free, should posit an intelligible character to Shylock as an agent. At this point, Kant has not developed anything specific about what these characters look like; on the contrary, "why the intelligible character gives us exactly these appearances and this empirical character under the circumstances before us, to answer this surpasses every faculty of our reason" (KrV A 557/ B 585). So all we can say is:

---

[35] On the significance of Kant's use of the notion of schema here, see Grenberg (2010).

[36] With caveats. See §§15-16.

> **Shylock's intelligible character**: Shylock is (intelligibly) *such that* if he is wronged, then he will choose to revenge; and...[37]

If we look at Shylock's act of seeking revenge, then, we can see Kant's model play out. On the one hand, his revenge was perfectly determined by natural laws, prompted by the fact that he was wronged. On the other hand, however, that his revenge would follow from his being wronged is part of who Shylock is as an intelligible being. What this may mean cannot be ascertained as of now, but it shows a way in which this temporal occurrence can nevertheless be attributed to him as absolute starter of the causal chain. It is in virtue of Shylock's being "such that he will choose to revenge", which is a fact about him as an intelligible being, that he indeed does when he is so prompted. This is a "condition outside the series" (KrV A 552/ B 580) and, if we have any credence to make sense of the idea that Shylock can be "such that he will choose to revenge" in an intelligible sense, then his revenge will be absolutely determined by his intelligible character while remaining, as a sensible action, within the boundaries of possible experience.

§5. Why is an *automaton spirituale* not enough?

In the last section, I went over Kant's treatment of freedom in the Third Antinomy to elucidate what exactly Kant understands by transcendental freedom, as well as the specifics of his solution to the apparent contradiction between the necessity of natural causality and such free faculty. Transcendental freedom is a capacity to absolutely begin a state, that is, it is a causality by which an agent becomes the absolute condition of a causal chain. This cannot exist amid appearances, but by virtue of transcendental idealism – appearances being *mere* appearances, and thus allowing for a further, intelligible ground – we can locate such a faculty outside of the chain of natural causality. In particular, if we have the warrant to attribute transcendental freedom to a being, we can postulate that it exhibits an intelligible character the 'effect' of which is the empirical character that gives natural regularity to his phenomenal actions. Up until now, the discussion has considered transcendental freedom hypothetically. In this section, I will address whether we have such a warrant in our case.

---

[37] A clear objection here is that we cannot use the category of causality and dependence for an intelligible object. However, (i) we actually *can* and *must* use the pure categories of understanding in *thinking* intelligible objects, which is as far as our access to such objects goes (KrV B 146); (ii) we are not attributing cause and effect to the intelligible ground itself, we are merely saying that, however the object may be – which we do not know – it has to ground, specifically, the causal law that applies to Shylock phenomenally. Thus the purposefully vague language in describing his intelligible character, we have no choice on that matter (KrV A 546/ B 574).

We have already seen (§2) that Kant, one way or another, holds that our awareness of the moral law authorizes – or, better, forces – us to regard ourselves as free beings. However, there is a big gap between this rather general claim to freedom and the very specific model Kant lays out in the first *Critique*. Nevertheless, it can hardly be disputed that Kant retains his radicalism regarding freedom in the second *Critique*. In what may be one of the most commonly quoted sections of his work on freedom, Kant decries attempts at settling for a "comparative concept of freedom" as a "wretched subterfuge" – most notably, the Leibnizian *automaton spirituale* –, which leaves us with the "freedom of a turnspit" (KpV 5: 99). We need, then, an argument to the effect that "the abolition of transcendental freedom would also simultaneously eliminate all practical freedom" (KrV A 534/ B 562)[38], so we may understand why Kant is adamant that morality bears such a high price tag. I contend that there are, at least, three reasons for Kant to believe so.

The first one is Kant's own argument to that effect in the resolution of the Third Antinomy: we need it for moral normativity in general. If all there was were natural causality, which encompasses any purely "internal determination" of states of affairs through our psychological states[39], our judgments could merely refer to what things are, but it makes no sense to seek an "ought" in nature (KrV A 547/ B 575). This is so because it can be the case that something ought to occur, yet it "have not occurred and perhaps will not occur" (KrV A 548/ B 576). By introducing an intelligible level of affairs, which somehow grounds part of what transpires in the empirical sphere but follows laws that are completely different from those ruling over appearances, Kant claims, we can understand what it would mean for something to be deontically necessary despite not ever occurring. This is Kant's own point, but it is dubious that the argument works. After all, it can be the case that we need a different standpoint from the mere naturalistic picture of the events that occur in the world in order to attribute normativity to some of them, so to speak "stepping outside" of things as they merely transpire, but by no means do we need transcendental freedom in particular in order to do so; on the contrary, in and of itself, transcendental freedom is a very strange candidate to assume this role. Once we survey the other reasons, Kant's point may become more cogent; for now, let us bracket it off.

---

[38] As mentioned in footnote 2, Kant's opinion on this matter in the *Critique of pure reason* is controversial; however, there is ample consensus that, from the *Groundwork* on, this is his view.

[39] See footnote 19. More on this in §§6-7.

Another relevant consideration why Kant would require that we have transcendental freedom is the imputation of actions. As he puts it in the remark on the thesis of the Third Antinomy, "[t]he transcendental idea of freedom [...] constitutes only that of the absolute spontaneity of an action, as the real ground of its imputability" (KrV A 448/ B 476), remark that resurfaces in his later work[40]. The point here seems to be as follows. In natural causality, there is no real ground for imputing an action to a single agent. After all, we can never find something like an ultimate cause in the phenomenal sphere. In the famed case of the malicious liar in the first *Critique*, for instance, why is *he* instead of his upbringing, or his company – or, why not, God[41] – imputable for his lie? However, if he has transcendental freedom, he is the absolute beginning of the causal chain by which he lied, insofar as his intelligible character grounds that empirical character by which, faced by the circumstances, he chose to be untruthful. Responsibility, in a sense, dissolves away in the thoroughgoing unity of experience if there is no sense in which agents are the paramount originators of their actions[42].

We can detect a similar train of thought in Kant's theory of private property. Mere phenomenal possession is accidental, it does not provide a sturdy enough link between owner and property, which is supposed to withstand unlawful appropriations (MS 6: 247); only an intelligible relation honours the original acquisition as legitimate and distinguishes the rightful owner from other claimants. Likewise, the link between agents and actions cannot simply be that of sharing a causal chain; the agent must somehow have a relation to that action that goes beyond the accidental happenstance of natural causation. Say that some prankster pushes me in front of a dart board, and a player, *naturally*, hits me. A complete natural causal history of my injury includes the past states resulting into the dart being thrown *and* those resulting into myself being pushed. But it is quite clear that we should regard the prankster, not the dart player, as the proper originator of the causal chain that ends in my getting struck by a dart[43]. Additional agents coming in and playing incidental roles in the causal chain are as irrelevant for the ultimate ground of attribution as a friend holding my phone for a second is for her to claim it as her own.[44]

---

[40] E.g., RGV 6: 52n.

[41] On this issue, central to the *Freiheitsstreit* stirred after Kant's critical theory of freedom, see §19.

[42] Thus Kant's concern in the second *Critique:* "if […] these determining representations have the ground of their existence in time and indeed in the *antecedent state* […] they are […] therefore under the necessitating conditions of past time, which are thus, when the subject is to act, *no longer within his control*" (KpV 5: 96).

[43] Compare this to modern "actual-sequence views", like the one espoused by Sartorio (2016: 18). If I just wandered into the playing area, or if the player threw the dart towards me on purpose, the *actual* source of my getting hit by a dart would be quite another than in this example.

Finally, Kant's specific moral commitments require us to be fully spontaneous agents. Recall that acting purely out of respect to the moral law is what makes an action morally worthy (GMS 4: 398). Crucially, this pure respect to the moral law is not something we may intuit even in our inner forum (RGV 6: 38); it is not a specific mental state, which is determined by a further empirical state, that can confer moral worth to an action; rather, it is the intelligible ground on which such mental state can be said to rest. The only way in which we could even expect to attain moral purity is, then, (i) if our actions can be grounded in a non-empirical principle, (ii) if our actions can be grounded absolutely in such principle, rather than have it be a merely concurrent cause[45]. Having said this, it may make more sense now why Kant thinks that transcendental freedom is the key to having moral normativity in the world. A morally worthy action is one that stands in a peculiar relation with an equally peculiar ground of determination, and this is so because Kant rejects absolutely the contingency entailed by potential phenomenal grounds of normativity (GMS 4: 390): whether an action succeeds or not, for instance, cannot be determining to its being morally worthy or not, the worthiness of our acting is, and must be, at least in principle a clear-cut matter. Thus, only by being somehow bound by a non-empirical causal law can we make sense that such a strict standard of normativity should be relevant for our world.

In conclusion, then, Kant clearly requires transcendental freedom – that is, absolute intelligible grounding of our actions –, rather than any other kind to feature in his practical philosophy. Only transcendental freedom can open up a space for the strict notion of moral normativity Kant has, secure the imputation of actions from phenomenal interference, and make it at least possible in principle that we could achieve moral worth.

§6. Perspectival freedom

So far, mentions of "intelligible grounding" have abounded in my reading of Kant's account of freedom, but how should we take this? There are two broad families of interpretations of

---

[44] This may be a reason why some current philosophers of action are wary of "stories" where the agent herself fails to settle anything in her acting, and everything is reduced to a causal process linking a belief and a desire to an action; the agent as a whole may have a claim to initiating *something* for an action to count as such. See Steward (2012: 39; 62).

[45] "And indeed one regards the causality of reason not as a mere concurrence with other causes, but as complete in itself, even if sensuous incentives were not for it but were indeed entirely against it; the action is ascribed to the agent's intelligible character" (KrV A 555/ B 583)

Kant's theory of freedom that have tended to downplay the metaphysical import of free action[46], whose heritage we can trace back to the work of two authors. On the one hand, there is a noticeable strand of Kantians influenced by Lewis Beck's *The Actor and the Spectator* (1975)[47]. On the other, we have those who use Donald Davidson's anomalous monism, as presented in *Mental Events* (1970), as a model broadly applicable to Kant. While these two have sprouted into fairly distinct family trees in their flair, general approach, and philosophical commitments, I contend that they agree on a key point. Under these readings, a naturally determined action is not made free by standing in a real relation with an intelligible ground; rather, it can be considered to be free to the same extent that it can be considered to be naturally determined insofar as it falls under one or another appropriate description. Two different conceptual orders, none more fundamental than the other, can be made to coexist.

I will not attempt to summarise Beck's and Davidson's theories, since brevity would, in this case, no doubt be a disservice. Rather, I will simply point out some salient aspects their heirs chose to adopt in their readings of Kant. Both of them agree that human agency allows of two different modes of description, one based on reasons, one oriented to explanation[48]; these are not neatly translatable, their respective concepts and grammar do not have clear correspondence rules:

> "There are no strict psychophysical laws because of the disparate commitments of the mental and physical schemes. It is a feature of physical reality that physical change can be explained by laws that connect it with other changes and conditions physically described. It is a feature of the mental that the attribution of mental phenomena must be responsible to the background of reasons, beliefs, and intentions of the individual. There cannot be tight connections between the realms if each is to retain allegiance to its proper source of evidence." (Davidson, 1970: 222)

---

[46] Following Nelkin's (2000: 566) characterization, we could regard these as "two-standpoints" readings: they ostensibly regard that we assent to two contradictory propositions regarding human freedom, albeit from different stances or conceptual layouts. Since two-standpoints is a broader overall reading of Kant's ontology, however, and not every two-standpoint defender follows the same strategy in making sense of Kant's theory of freedom (see §9), I will use the more neutral "perspectival" to refer to them.

[47] This is, needless to say, a gross simplification of the genealogy of these interpretations, which would need to include, among others, the towering presence of Wilfrid Sellars; for the sake of simplicity, however, I have found it useful to focus on these two particular works, since their influence is unquestionable and generally straightforward.

[48] They, of course, did not invent this distinction, a key one in the early 20th century's *Geist-/Naturwissenschaften* debates. We find it in, to name but two, Max Weber and Wilhelm Dilthey.

> "On a specific occasion when Spectator I says the child is writing, Spectator II says muscles 71, 83, and 99 behave in such and such a manner. If this were a translation of what the first spectator says, there would be a rule for its production from what the first spectator says, and this rule could be followed again and again. But it cannot be; the net time the one says "the child is writing", the other cannot just look up this rule of translation" (Beck, 1975: 39).

Their difference, for our purposes, is one of emphasis. Davidson presses his point about the token-token identity of events of unlike types in order to accommodate the anomality of the mental within a causally closured world (1970: 224). Beck, meanwhile, appears to be more concerned with the shortcomings of behaviourist reductionism (1975: 55).

Davidson and Beck's Kantian connection is not particularly covert, and it comes as no surprise that we would find Davidsonian and Beckian[49] readings of Kant's theory of freedom. As for the former, I would like to bring attention to the proposals of Meerbote (1984) and Hudson (1994); as for the latter, to the very influential interpretation brought forward by Christine Korsgaard (1996). The three, due to their heritage, belong to one same grouping I will call the *perspectival reading*: what it means for an action to be free is not to stand in a peculiar relation to its ground, but to simply fall under the appropriate description in a reasons-oriented language. Actions *qua* free and *qua* unfree do not belong to the same conceptual cutout of reality.

Meerbote and Hudson both argue that Kant makes a relevant distinction between determining and non-determining judgments when speaking of natural causality and teleology, respectively[50] (Meerbote, 1984: 60; Hudson, 1994: 46). Any of our actions is token-token identical to a natural event that is, thus, subject to the determinism of causality, but they can be regarded as intelligibly grounded when the non-determining descriptors proper of normative judgments are deployed onto them (Hudson, 1994: 46). Korsgaard, meanwhile, downplays the usefulness of token-token identity to emphasise the peculiar standpoint that deliberation brings with it. When we deliberate, Korsgaard argues, we must take whichever causes bring us to action, not merely as causes determining us, but as reasons. Even if we have been programmed to act and think a certain way, the way rational agency works is for us to regard our decisions "as springing ultimately from principles that we have chosen" (1996: 163). Otherwise, we would not count them as decisions to begin with.

---

[49] Not to be confused with Beck's own account of Kant's theory in his critical works.

[50] On a textual basis, this is a grave confusion, anyway. The teleological judgments that are non-determining for Kant are those in which we consider a being as exhibiting natural finality (KU 5: 361), but whenever we do act according to representations, and thus teleologically, we are one special kind of natural cause among others (KU 5: 196n.). Nevertheless, Meerbote and Hudson's accounts are problematic even if we disregard this. Cohen (2009b), as a contrast, argues that teleological judgment can provide a *model* for explanation in the human sciences.

A case could be made that these approaches would succeed in addressing last chapter's volitional problem; they do show that, when we act, we consider ourselves in a manner distinct from the mere succession of natural events. But they certainly cannot solve the real agency problem, let alone how Kant conceives of it, and that is so for three reasons.

First of all, it is very hard to see why transcendental idealism is needed at all if Kant's point is simply that we have different conceptual frameworks to describe people as natural beings or as deliberators.[51] Say – ignoring the contradictions from the Third Antinomy – that natural causality is absolutely fundamental and a feature of things in themselves. It would be an ultimate matter of fact that we are not free. Yet, the language of reasons and non-determining pro-attitudes would not become moot. We can, at any point, construct a conceptual framework to overlay on top of mechanistic reality, and it might well be useful for us even if the world is metaphysically unwelcoming to a more robust sense of freedom.[52] Korsgaard would be quite right in saying, then, that determinism does not matter (1996: 162). But Kant's view is rather another: "if one takes the determinations of the existence of things in time for determinations of things in themselves (which is the most usual way of representing them), then the necessity in the causal relation can in no way be united with freedom" (KpV 5: 94).

Secondly, Hudson claims to be honouring Kant's invective against the *automaton spirituale* insofar as freedom, according to him, "is found in the applicability of an *intelligible* cause, independent of natural causes" (1994: 52). But if the only thing an intelligible cause means is a certain type that does not neatly map with the concepts we use in our general understanding of nature, (i) it is difficult to understand how this is not "to evade this by saying that the *kind* of determining grounds of his causality in accordance with natural law agrees with a *comparative* concept of freedom" (KpV 5: 95-6), since, at the end of the day, all we are doing is rearranging phenomena according to principles of action and noting that they do not fall in line with the regularities of nature[53] (though they remain "in accordance with them", since their causal efficacy is parasitic on natural laws), and more importantly (ii) nothing is done to answer Kant's real concern, that "since time past is no longer within my control, every action that I perform must be necessary by determining grounds *that are not within my control*" (KpV 5: 94). If there is no real intelligible ground of natural causality, there is an objective sense in which, regardless of our reasons-based idioms, the determining ground of our action is *outside* ourselves. This is precisely one of the reasons why we saw in the last section that Kant requires specifically transcendental freedom in his account; if a shift in description were enough, Kant's worries about the real ground of the imputation of our actions would be far less poignant.

---

[51] Korsgaard seems to waver between this strictly perspectival view of freedom, as in the quote mentioned before, and acknowledging noumenal grounds: "since we must suppose that there are some undetermined first causes, or free agencies, which generate the appearances, we must suppose that things which exist in the noumenal world are free" (1996: 201). If she indeed believes this, then my argument here does not apply to her; Frierson (2003, 2010) seems to double down on this route, defending a two-standpoints grounding relation. However, she ultimately chooses mutism about the relationship between phenomena and noumena over an account such as the one I will introduce (1996: 203). As I will argue now, I believe it unavoidable to admit somewhat more specifically that our free actions must be intelligibly grounded.

[52] As Cohen (2009a) argues, we only need a comparative or practical concept of reason for this.

Finally, if all there is to free action is the deliberative stance, or non-determining pro-attitudes, Kant's claim that we can never know for certain our moral worth becomes either a merely empirical assertion or utterly mysterious. If we are free, and thus can be morally worthy in acting from respect to the law alone, by simply considering ourselves as rational agents, what exactly is the opacity that brings Kant to categorically deny any certain knowledge about the true motives of our heart (GMS 4: 407; RGV 6: 38), and to require nothing less than an eternal afterlife to approach moral purity asymptotically (KpV 5: 122)? Sure, we often deceive ourselves, but it would be quite the anthropological statement to say that we *never* put ourselves "under the idea" of acting according to duty. However, if moral worth is inaccessible through the mere appearance of our grounds of action and it resides, rather, in something of which we can only know the "sensible sign" (KrV A 546/ B 574) then it makes sense to understand this opacity as systemic and irremovable.

Perspectival readings may make for a more palatable theory of action all things considered, but they cannot keep up with the specific and strict requirements Kant has in mind for freedom; he needs those appearances we call free actions to stand in a grounding relation with an intelligible causality, not merely to have something like an "intelligible description" apply to them. This is a matter that cuts through different ontological commitments with regards to transcendental idealism (Korsgaard being a proponent of two-standpoints, whereas Hudson defends two-aspects, for example), and leaves our requirement for a real ground, for now, as not necessarily linked to one ontological position or another.

---

[53] Certain physical laws rule over us as natural beings, say that they are such that 50% of the time we encounter a situation that falls under the concept of "an opportunity to lie to our friend" it results into an event we label as "telling the truth", and 50% of the time under "telling a lie". There are no natural laws, assuming they could not be probabilistic, that link these sets of states. However, for any specific case, all there is is an instance of one or the other natural law, and it is *just* the fact that our overlaid concepts can apply in those cases – which, evidently, will depend on some feature about those natural events – that accounts for this causality obtaining. At no point does the intelligible cause become the ground of effects, it is only a non-determining sorting of already causally efficient phenomena.

§7. Scofflaw freedom[54]

Meerbote and Hudson's approaches relied on a certain reading of the *Critique of Judgment* that seemed to open a space in Kant's system for a non-determining sort of judgment about our actions. I have argued that this – aside from being textually dubious – does not suffice to do away with a real intelligible ground of actions. There is, however, another line of interpretation put forward by Hanna (2007) that tries to use peculiarities from the third *Critique* to address what seems like an impossible puzzle in Kant's account. Namely, he argues that there is no contradiction between free action and the causality of nature because the Antinomy of Teleological Judgment derogates the universality of natural causation and, in particular, because we, as living beings, do not behave in a deterministic fashion[55].

Hanna argues that, in the third *Critique*, Kant realizes that we do not have an immediate warrant to assert the existence of "specific empirical laws of nature 'all the way down'" (Hanna and Moore, 2007: 121), and that there are parts of nature – such as ourselves, as living beings – that are not subject to mechanistic laws. Because of this, there would be a "nomological gap" open for rational beings to create "one-off" laws that, while permitted by the general laws of mechanism, are not necessitated by them. Thus, free agency manifests itself by, quite literally, creating its own laws in the world.

As a general point, before addressing Hanna's argument more closely, I must emphatically insist that there really cannot be a Kantian indeterminism at the phenomenal level without forfeiting the core thrust of the *Critique of Pure Reason*, and this includes both our outer and inner fora (A 34/ B 51). Our mental states, as much as external objects, are subject to the form of time, and are to that extent appearances. But insofar as we can determine an objective succession of occurrences in time – and we can even for our mental states, as explicitly established in the Refutation of Idealism (B 275) –, then these occurrences must be subject to the law of cause and effect, since it is this very law that makes sense of the difference between a subjective and an objective temporal succession (A 192/ B 237). Kant could hardly be more explicit about the extent to which he holds phenomenal determinism to apply, as we have already seen in his remarks about the possibility of predicting the actions of humans, and in general in his discussion of causality (A 199/ B 244; see Allison, 1990: 31).

---

[54] I borrow the phrase from O'Connor (2000).

[55] Versions of this approach have also come from Kant's remarks in the *Metaphysical Foundations* that there cannot be such a thing as a psychological science. My general argument works against any kind of scofflaw, so I will not consider them in particular. For a closer look at this, see sections 1.2. and 1.3. in Frierson (2014).

For us to hold that the Antinomy of Teleological Judgment abrogates this would mean that Kant considered the third *Critique* to deal a lethal blow to one of the crown jewels of his theoretical philosophy, with no explicit mention to that regard. Of course, if no other reading of the antinomy were possible, we would have no other choice but to admit a very substantial difference between a pre- and post-*Critique of Judgment* Kant, but that is not the case. As it has been argued recently (McLaughlin, 2014; Geiger, 2022), mechanism is not synonymous with natural causality. Without entering into specific details yet, mechanism seems to refer to part-to-whole dependence as opposed to the whole-to-part dependence proper of teleological products. Even though, as Zuckert (2007) has argued, time plays an important role in understanding the different judgment structures involved in teleology and mechanism[56], this does not mean that teleologically created beings in any way bypass the flow of time. Further, even though this is a great matter of debate[57], with influential readings to the effect that organisms are properly theoretically inexplicable from a *causal* perspective (Ginsborg, 2004), some other interpreters regard that there are two levels of explanation at play in organisms, one of which is still strictly mechanistic and causal (Illetterati, 2014; Geiger, 2022). If this is so, the case for an organism-based disregard for natural laws is further undermined.

Now, Hanna is not strictly a “scofflaw”, insofar as he does admit that our actions as organisms should follow laws and be constrained by those of mechanism, but by means of non-necessitated, one-off empirical laws we bring into the world. This is, however, still problematic. To begin with, and as Indregard (2018) observes, Kant himself regards our empirical character as standing “through and through in connection with other appearances in accordance with *constant* natural laws” (KrV A 539/ B 567 [emphasis added]). Moreover, though this is a matter of great controversy, there are strong arguments to believe that the Second Analogy requires general laws, not just a general connection of occurrences to preceding states (O’Shea, 1997). But even leaving this aside, what is a one-off law, exactly? If we understand it as Moore does in his reply to Hanna, as a “maximally specific” law (Hanna and Moore, 2007: 126), then it is only a one-off law contingently; it is a law that holds universally but which, as it so happens, only has a single event in the history of the universe fall under its scope. Why would it matter whether a law has only applied once or a million times to establish whether our actions have been free?

---

[56] “As we have seen, Kant articulates the structure of purposive causality, by contrast to that of efficient causality, as a symmetry of ‘‘ascending’’ and ‘‘descending’’ relations (V: 372). This symmetry in purposive relations provides Kant with a structure that describes the ‘‘specific’’ unity found in organisms. But this very symmetry of (non-intentional, reciprocal) purposive relations in time drives purposive causality into conflict with the law of causality and the nature of time, as Kant conceives it in the CPR.” (Zuckert, 2007: 136). This, I gather, does not mean that in organisms things do not follow the objective succession of time; rather, that teleological *judgment*, required by the unity of the contingent present in organisms, operates in a way incompatible with those objective successions of mechanistic judgment.

[57] See Gambarotto and Nahas (2022) for an overview of this debate.

The key to Hanna’s proposal is that we, somehow, are the ones to create or generate these one-off laws. Now, it can be that we do so simply inasmuch as our intelligible character grounds our empirical character, but then we are back at my reading, which requires no one-off laws and no nomological gaps. The other option is that the law only became one the moment we ‘issued it’; we would *literally* bring a new law into existence in the specific moment of our action. But this is extremely hard to fathom. Let us say I am choosing between a vanilla and a chocolate ice cream, given conditions a, b, c, etc. I decide for the vanilla ice cream, thus closing a nomological gap and creating the law that “if conditions a, b, c, etc. (which will never again repeat, and had never before happened) take place, then I will choose a vanilla ice cream”. This law was not true before I actually chose the ice cream; we can then imagine an identical world that leads to the opposite law becoming true. What makes the law true, then?

If it is the mere fact that what happened, happened, then laws are mere empirical generalizations and they can hardly fit the criteria of universalizability and necessity that Kant attributes to them[58]. If my ice cream law can toggle one way or the other just by a contingent, groundless fact by which I chose one or the other, then it is no more a law than one that prohibits murder unless the killer chooses to commit it. If the law is grounded in something, on the other hand, then either the ground preceded the law in time, and so the worlds leading to a vanilla or a chocolate law were not the same after all, or it did not. In this latter case, I can only picture that an intelligible ground could do the job. But an intelligible ground is not under the conditions of time, and although this does not necessarily mean that they can have nothing to do with time whatsoever (§18), they cannot come into existence (A 541/ B 569), and thus we cannot say that this law only became true when we acted; it became applicable for the first time, but it was already ‘in the books’.

---

[58] Hutton (2021). I go over the status of empirical laws for Kant in more detail in §15.

Finally, even I we managed to make sense of one-off laws in a Kantian context, Hanna uses Kant's third *Critique* realization of the contingency of empirical laws with regards to the conditions imposed by transcendental laws in a way that cannot work. Kant, indeed, brings this issue up, but only to note that a solution must be found – empirical laws *cannot* be contingent, so we will have to find a way to regard them as necessarily following from a more fundamental ground (KU 5: 180) (§15). We cannot use what Kant regards as a problem in need of solution, and ostensibly one he thought to have solved, as the steppingstone of our theory of freedom.

In sum, then, freedom is not to be found in bypassing the laws of nature. Kant does not allow indeterminism, and seeking ways to relax the requirements of empirical laws leads to philosophical and textual problems. Kant's theory of freedom is one that embraces the ironclad rules of nature.

§8. World-grounding freedom

So far, I have argued that Kant's theory of freedom requires us to posit a real relation between our phenomenal actions and their intelligible ground, and in such a way that they respect the empirical laws of nature. This, however, still leaves the door open to different possible readings that, despite complying with these core principles, result in dramatically different views of Kant's theory. The one I want to address here is proposed by Wood in his 1984 *Kant's Compatibilism.* Wood indeed takes it as a given that "Kant does *not* hold in general that freedom is incompatible with causal determination or even necessitation of the free being's actions" (1984: 82). Nonetheless, the way in which he understands this to play out is, to say the least, peculiar:

> "A particular timeless choice of my intelligible character affects the natural world by selecting a certain subset of possible worlds, namely, those including a certain moral history for my empirical character, and determining that the actual world will be drawn from that subset of possibilities." (1984: 91)

Now, Wood is keenly aware of the *prima facie* difficulties such a world-grounding account of freedom has. For one, it seems as though we are to be held responsible for everything that has ever happened, so far as our "timeless choice" determined an entire course of the world's history (Wood, 1984: 92). His defence against objections of this sort is twofold: (i) Kant's only intention is to show that the notion of freedom is consistent, and (ii) insofar as we actually *do not know* what goes on at the intelligible level, we cannot be properly blamed for actions other than those we have empirical reasons to impute.[59] This strategy may strike us as dubious; after all, it seems that, even if we do not have knowledge about the intelligible realm, our metaphysical model of freedom should be as adequate as possible in explaining those features of freedom we care about, which is what we attempted in §5. Unless there is a very good reason to understand transcendental freedom as an atemporal choice of a possible world, it seems that we are better off without it.

Here is why Wood may have thought that he needed to attribute such all-encompassing powers to freedom. In the second *Critique* quote Wood provides just before stating his interpretation, Kant says that, for any unlawful action a rational agent performs, "this action, with all the past which determines it, belongs to a single phenomenon of his character, which he gives to himself" (KpV 5: 98). This is, no doubt, a challenging and mysterious passage, but is Wood's reading entirely justified? As I read it, it is a restatement of what we already know. Due to the results of the Third Antinomy, an action can be granted an absolute intelligible beginning without denying its determination by the past as appearance. All Kant ever says is that *our actions* are to be regarded as consequences of "the determining ground of [our] causality as a *noumenon*" (KpV ibid.). Our action, not devoid of the past which determines it, but "with it", is what can be regarded as absolutely beginning from our character. It belongs to a "single phenomenon" of our character insofar as it is considered to be an absolute beginning in the causal chain, as regarded from the standpoint of the intelligible character, which has the empirical character as its effect without being itself determined.

Still, there may be an intuitive sense by which we may want to regard the entire past as determined by the intelligible character. After all, if everything in appearances is determined, and we control a certain occurrence, we must also control its cause, and the cause of the cause, and so forth. This might be the insight behind Wood's maximalist solution, and thus of his reading of Kant's fragment. But if we recall the results of the Third Antinomy, this complete control of causes going all the way back to the beginning of the universe is not necessary. We already saw (§4) that an absolute beginning in a causal chain is not the same as an absolute beginning in time; Kant makes room for such beginnings taking place amidst other phenomenal occurrences. This can happen because the empirical character that regulates what effects follow from what causes may find its ground in an intelligible character, which, then, is seen as the proper ground of determination of whatever empirical effects follow from there. If we remember Shylock, his committing revenge is under his control, not because he caused every single phenomenon that led to it, but because he, as an intelligible being, grounds the empirical law by which he commits revenge when wronged. We "give our character to ourselves", and from there, after the past has transpired, our action follows as something we are in control of, since it is our intelligible character that grounds our character. O'Neill's note to that effect is helpful as ever:

---

[59] See also Pereboom (2007).

> "…all naturalistic explanations – even the most impressive explanations of some future neuroscience – are conditional explanations of the "appearances of men's wills". In a certain sense they are incomplete, for they can never explain that any natural law should take the form that it does" (O'Neill, 1989: 68).

The intelligible ground of our actions finds a space precisely as the ground of those empirical laws that rule over our actions, but only there. We do not need to promote our noumenal self to the rank of a cosmological choice.

§9. Action-grounding freedom

In the last section, I argued that Wood's world-grounding approach to Kant's theory of freedom should be a last resort if no other option remained, and I sketched why I consider that there is no real need to read Kant as holding such a view. With that, we have all we need to recap and give a more fleshed-out proposal as to what Kant understands by a free action.

We act freely in the transcendental sense if we can be regarded as the absolute ground of determination of our action (KrV A 446/ B 474). This is to say, if we can be regarded to not have been determined into causing such effect, if we are the absolute beginning of the causal chain that resulted in such action. Since our actions are appearances, however, they are subject to the conditions of time and determined by the past (KrV A 199/ B 244; KpV 5: 98). It is only thanks to transcendental idealism that we can consider an absolute ground of determination outside of this causal chain: an intelligible ground of determination. In particular, we can posit that our empirical character – the empirical laws that govern our actions – is the "sensible schema" (KrV A 553/ B 581) of an intelligible character, which is under no conditions of time, and thus is not itself determined by anything in the past. This model is required in order to secure the entrance of normativity into the world of appearances, to ground the imputation of actions, and to make sense of the possibility of moral worth. Further, this is a real relation between a posited intelligible character and our empirical self, one that maintains the laws of nature in place, and which does not apply universally to all phenomena in the world, but only to those actions we judge as free (KrV A 546/ B 574)[60].

Does this commit us to any specific views on the ontological status of phenomena and noumena? All that is required by this reading is that our view of Kant's base metaphysics allows an intelligible grounding of phenomena. If one's ontological commitments do not permit this, then my reading will be unacceptable; however, if my arguments are successful, this reading is textually backed and sound with regards to the problems Kant sets out to solve, so it lends, at least, some credence to casting doubt on such an ontology. Nevertheless, this reading can be amicable to a wide array of such stances. Frierson (2003, 2010) defends a two-standpoints reading of Kant's philosophy, but he finds room in it for a grounding relation between noumena and phenomena:

> "From a *metaphysical* standpoint, the free will is a noumenal ground of the appearances […] From a *practical* standpoint, one considers the free will as grounding the appearances not as an abstract truth, but as the realized condition of deliberation" (Frierson, 2003: 132).

If two-standpoints, as the deflationary reading it is, can accommodate this grounding relation, it should not be too problematic for other, more metaphysically hefty readings, to do so. Thus, an action-grounding model of freedom seems to be the way to go in understanding Kant's solution to the real agency problem.

---

[60] How we do this is an important question I will get back to in chapter 4.

3

NO AGENCY WITHOUT EMPIRICAL UNITY

In the last chapter, I argued that Kant understands free actions to be intelligibly grounded in such a way that, though phenomenally determined, they can be said to have the agent *qua* noumenon as their absolute cause. This, however, leads to a challenging consequence: it makes any attempt at finding freedom in experience fruitless, since all that makes actions free is strictly beyond possible experience. If we wondered about the mere possibility of freedom[61], we would not need to worry all that much. We understand how conditioned causal chains in the phenomenal sphere leave room for an ultimate noumenal grounding, and we need not concern ourselves with anything else playing out in experience. However, if our goal is to solve the real agency problem, and have freedom actually play a role in our moral practices, we need something more.

In this chapter, I will argue that, in order to give a satisfactory answer to this issue, we specifically need to attribute actions to empirically unified beings who are intelligibly grounded in the right way. If we do not do so, I will claim, we will face three unpalatable results. I start (§11) by presenting my reading for Kant's account of action attribution, surveying relevant textual evidence in support of it. Afterwards, I present three problems that arise if we do not accept that actions should be attributed to intelligibly grounded, empirically unified beings. First (§12), the *unity problem*: nothing about actions, as phenomena, can tell us that they should be attributed to one and the same noumenal self. Then (§13), the *scope problem*: our own moral awareness does not have enough content for it to determine which specific phenomena count as our action. Finally (§14), the *moral progress problem*: we must be able to radically change our moral characters at the noumenal level while still being responsible for our past actions. These three possibilities, which would be deleterious for our considering phenomenal actions as caused by free beings in a meaningful sense that looks after our moral practices, are dispelled by having freedom entail an intelligibly grounded empirical unity of the agent.

---

[61] I use the term "mere possibility" in contrast with real possibility, which Kant explicitly states we cannot theoretically assert about freedom (KrV A 558/B 586). I do not use "logical possibility" instead only because it could be argued that freedom would not cease to be *logically possible* if, say, appearances were things in themselves.

§11. Agents in two realms

Frustratingly enough, Kant is not explicit about the exact relation that holds between our phenomenal actions and the presumptive intelligible ground that makes them free. We know that we do not act freely when the determining grounds of our action escape our control, as it would be the case if all there existed was temporally determined phenomena (KpV 5: 94). We are free insofar as we can be said to possess a different kind of causality, one that pertains to us as noumena (KpV 5: 98). In particular, since we must be autonomous, as demanded by morality, our wills must be able to determine our actions independently of alien causes (GMS 4: 446).

All throughout these discussions, there appear conspicuous references to the noumenal character of "our" wills in general, where the "we" the sentence refers to is left as though it needed no further discussion (GMS 4: 447), to properties of "rational beings" across realms (GMS 4: 453), and to considering "the same being" as appearance and thing in itself (KpV 5: 95). It seems reasonable to suppose that whatever we must posit in the intelligible sphere pertains or adheres to a certain being in the empirical realm[62], giving us a sense in which we can call them "the same", however heterogeneous these two spheres are. The fact that Kant seems to assert it without much argument (GMS 4: 452) indicates that he probably found that to be obvious[63]. Little attention has been paid, however, to the kind of being that we should expect to find at the phenomenal level, such that it can be "the same" as a noumenal self[64].

It is hard to understand what a noumenal self may be other than a unity[65], since nothing "starts" or "ceases to be" in noumena (KrV A 539-40/B 567-8). Kant, as a matter of fact, says that "reason is present to all the actions of human beings in all conditions of time, and is one and the same" (KrV A 556/ B 584), as the "persisting condition of all voluntary actions under which the human being appears" (KrV A 553/ B 581). However, things get – shockingly – trickier when we try to surmise the consequences this may have for its phenomenal counterpart. The equivalence between sensible and noumenal selves gets muddled from the start by the fact that phenomena, by themselves, cannot simply express the sort of unity we may expect of intelligible entities. For one, as syntheses in space and time, they are always composite[66]; we only find those kinds of underlying unities as regulative ideas of reason (e.g. KrV A 771/ B 799). Further, any such synthesis of the manifold of intuition into a unity depends on understanding deploying an empirical concept that can do the work (KrV B 162). It appears, then, that in order for us to track what our noumenal selves are supposed to be correlative to, we need some sort of empirical unity we might find in experience.

---

[62] See McCarty (2009: 120).

[63] For an alternative interpretation, Marshall (2013) considers phenomena and noumena in general to be partial "qua-objects" of the same overarching entity.

[64] Some exceptions include Munzel (1999) and Frierson (2014).

[65] Even though unity is a category, we can think objects, though not cognize them, through unschematized categories (KrV A 245/ B 302).

What that kind of unity may be, will have to wait until the next chapter. For the moment, I will present a provisional characterization of Kant's account of free action that relies on *some* sort of empirical unity for the agent, and I will then proceed to argue that such unity is an indispensable requirement for Kant's theory of freedom:

(Intelligibly Grounded Agents) A phenomenal action[67] a_P is a free action if and only if:

(i) An empirically unified agent *A* is causally responsible for a_P,
(ii) *A* has the right sort of intelligible ground,
(iii) *A* is causally responsible for a_P in virtue of *A*'s intelligible ground.

It is only natural, when we reflect about the conceptual backdrop of our practices, to assume that *we* must be their relevant subjects, that is, the physical and actually existing beings that have desires, needs and beliefs, and who can suffer and have their interests contested. Based on Kant's near-constant appeals to us *qua* rational beings, or intelligible selves, though, one may be inclined to think that he did not mind our status as physical entities that much[68]. Nevertheless, as we have already seen, Kant simply cannot disregard the phenomenal in constructing his practical philosophy: it is in the sensible world that we act, after all (KU 5: 196).

Let us start by being more precise as to what an "empirically unified agent" is. In its barest form, we can start with the notion of an "object" (*Object*) as presented in the transcendental deduction: "that in the concept of which the manifold of a given intuition is united" (KrV B 137). Something is an object for us insofar as we can synthesize the manifold in which it consists, which we only do through concepts, via our understanding (KrV B 146).

---

[66] Per the Second Antinomy, we can never suppose that we have reached an indivisible object in experience, even though we cannot think either that there should be *no* such objects.

[67] I specify "phenomenal action" to distinguish it from the actions of God, which are free, but do not require an empirical agent of any sort. The status of divine action is complicated, however, and the sense in which it is free may not be the same as that in which our actions are. See Insole (2013).

[68] See, for instance, Michalson's (1990: 69) remarks on "Kant's latent resentment against the body, his philosophical chagrin that pure reason must cohabit with sensuousness".

Now, as it will later become apparent, this cannot simply be an illusory or gerrymandered object of our thought, but it needs to be *persistent*, remain in existence for a certain time[69]. Finally, it needs to be an *agent*, that is, it needs to be causally responsible for a number of events in the world (KrV A 204/ B 249). We are looking, then, for a synthesized manifold in experience that can be said to persist and effect changes into the world. Let us, then, see if we can find – implicitly or explicitly – this account in Kant's own words.

The first condition of (Intelligible Grounded Agents) will not be hard to find. First of all, we already consider ourselves as empirical unities for theoretical reasons. Kant showed that rational proofs of the substantiality of the soul are paralogisms, which fail because:

> "the proposition 'I think', insofar as it says only that **I exist thinking**, is not a merely logical function, but rather determines the subject (which is then at the same time an object) in regard to existence, and this cannot take place without inner sense, whose intuition always makes available the object not as thing in itself but merely as appearance" (KrV B 429-30)

Only through experience, and, more specifically, through the relation of our inner sense to external, persisting objects (KrV B 275-6) do we talk of ourselves as existing, whereas the pure "I think" of apperception abstracts from any content of experience, including what we know about ourselves. So far as it makes sense for us to speak of ourselves, then, it must be as empirical objects. We can even push this point further if we follow Kitcher (1982) and Longuenesse (2007) in considering that the synthetic unity of apperception requires a unified *empirical* self[70]; if that is so, then our existence as empirical unities would even come to be a condition of experience.

---

[69] This could lead us to think of substances, and, in some sense, it needs to be so. Kant says: "For according to the principle of causality actions are always the primary ground of all change of appearances, and therefore cannot lie in a subject that itself changes, since otherwise further actions and another subject, which determines this change, would be required" (KrV A 205/B 250). However, empirical substances in and of themselves do not seem to warrant any intelligible ground in particular; the route I will propose is one where this ground is warranted from the beginning, as an exigence of our power of judgment (§§15-16).

[70] In a follow-up paper, Kitcher (1984) argues that, even though a Kantian philosophy of mind should commit to these empirical selves, Kant himself thought of the "I think" as having neither a phenomenal nor a noumenal reference. I believe, nevertheless, that this is not a problem for the point being pressed here. It may be so that Kant thought that the operation "I think" had a problematic status in relation to the phenomena/noumena distinction, *and* that we nevertheless need to think of ourselves as persisting, unified empirical beings to relate our temporally distinct mental states to the same subject. I see that my strawberry plant has just blossomed, and a couple of days afterwards I see that the petals have started falling and strawberries are starting to grow. The representation "I think" must be able to accompany every representation involved in this, and it may so be that it is not phenomenal; but unless *I nevertheless am* an empirical unity, I cannot attribute the two temporally separate representations to myself as a subject, which is Longuenesse's argument.

Moreover, insofar as we are agents in the world, that is, beings who act, Kant holds that we must be empirically persistent:

> "For according to the principle of causality actions are always the primary ground of all change of appearances, and therefore cannot lie in a subject that itself changes, since otherwise further actions and another subject, which determines this change, would be required. […] For that the primary subject of the causality of all arising and perishing cannot itself arise and perish (in the field of appearances) is a certain inference, which leads to empirical necessity and persistence in existence, consequently to the concept of a substance as appearance." (KrV A 205-6/ B 250-1)

This is further confirmed by Kant's account of empirical character in the resolution of the Third Antinomy. The empirical character is the law of the causality of a subject of the world of sense (KrV A 539/ B 567). If the way our free actions appear in the world is as effects ruled by an empirical character, and there must be a *subject* that *has* the said character, then, together with the theory put forward in the Second Analogy, there is little room to doubt that phenomenal agents must be empirical unities.

The second condition, we have already basically argued for in the last chapter. Let us however focus on two points that may have been left unclear in our discussion so far. First, that it is the *agent* who must "have" such ground. If we go back to the Third Antinomy, Kant states that it is the very subject who had an empirical character that must *also* have an intelligible character:

> "for a subject of the world of sense we would have first an **empirical character**, through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws […] Yet second, one would also have to allow *this subject* [emphasis added] an **intelligible character**, through which it is indeed the cause of those actions as appearances, but which does not stand under any conditions of sensibility and is not itself appearance. The first one could call the character of such a thing in appearance, the second *its* [emphasis added] character as a thing in itself." (KrV A 539/B 567)

Secondly, it has to be the *right* sort of ground. The previous fragment appears in the short epigraph titled "The possibility of causality through freedom unified with the universal law of natural necessity" within the resolution of the Third Antinomy. However, Kant says, this is not a solution of the problem of freedom yet. In the next sub-section, Kant calls it a "silhouette of a solution" (KrV A 542/ B 570) and goes on to characterize freedom as a causality of *reason* (KrV A 547/ B 575), made possible by the metaphysical backdrop he had set up[71]. It is true that, in 1781, when Kant wrote this, the relation he pictured between freedom and reason was more extensive than what he would admit from the *Critique of practical reason* forward (Timmermann, 2022: 48n.; cf. Korsgaard, 2009), which can still be partly seen in *Groundwork* III (Saunders, 2021). Nevertheless, the basic idea holds even more so after the second *Critique*; *pure practical reason*, through its moral law, takes up the role that Kant previously gave to reason in general (KpV 5: 93).

Why does Kant find it apt to characterize the relation between empirical and intelligible selves as one of grounding? In the *New Elucidation*, Kant defines "ground" as "that which determines a subject in respect of any of its predicates" (NE 1: 391). Let us recall why exactly Kant brought an intelligible character into the picture in order to make room for freedom (§4). If we only had an empirical character, that is, if phenomena were things in themselves, actions would always be ultimately determined by factors outside ourselves (KpV 5: 95), since natural causal chains extend indefinitely towards the past. However, if we admit an intelligible character, we can have *it* determine the empirical character, which then determines individual actions (KrV A 556/ B 584; KpV 5: 99[72]). Our intelligible character is the ground of our free actions.

This takes us straight into the third condition we were after. In the Introduction to the *Metaphysics of Morals*, Kant defines free actions – deeds – as those where the agent is regarded as their *author* inasmuch as we consider that they have freely chosen them (MS 6: 223). This makes sense. As rational beings, we have an intelligible character that serves as ground for free actions, but not every effect for which we are the cause is a free action. If I stumble, fall to the ground, and accidentally crush a model of the *HMS Victoria* someone had left lying around, even though I am causally responsible for the miniature wreckage, the ground of that causal relation lies outside my intelligible character, and, thus, it is not a free action.

---

[71] Which makes Beck (1963:192) and Bennett's (1984) objections against the cogency of Kant's first *Critique* account of freedom hard to understand. There are scope problems with freedom, as I will address in §14, but it is not that freedom could be ubiquitous just because it is noumenal. Freedom is the causality of reason, which needs to have certain conditions: namely, it supposes an absolute causal origin in a way that only a noumenal ground can accommodate. Bennett points at the very serious epistemic problems freedom entails – which I cannot address in this chapter – but his objection is beside the point: "although the Kantian theory says that our untutored opinions on freedom might be right, it offers no way in which that could be other than sheerly fortuitous" (1984: 107). But it *cannot* be fortuitous; there may well be other absolute causal origins in noumena, but only where we are authorized to speak of rational action can we posit an intelligible ground that is, properly speaking, freedom. A different question is that which Fichte raises in his letter to Reinhold from August 29, 1795: how do we ascertain who is a rational being?

[72] "[…] that entire chain of appearances, in relation with that which can always be of interest to the moral law alone, depends on the spontaneity of the subject as thing in itself […]"

There is ample evidence, then, in support of Kant's adherence to (Intelligibly Grounded Agents). Condition (i) establishes that phenomenal action, in general, requires empirically unified agents. Condition (ii) locates the possibility of freedom in an intelligible ground rational agents have. Finally, condition (iii) sorts out those actions that are attributed to the agent in virtue of this precise ground, as those that are, properly, free. But some may try to resist this conclusion. It may be too metaphysically costly[73], or have unsurmountable epistemic problems, by forcing us to give quite some content to noumena and their relation to phenomena. Following Bennett, one might say: "I do not dispute that insofar as Kant had a single doctrine about noumenal freedom it was [this] one […] But I protest that this kind of possibility is not worth establishing" (1984: 106-7). Perhaps Kant thought this much about free agency, but maybe he should not have. In the following sections, I will argue that, if we try to remove empirically unified agents with intelligible grounds from the picture, three fatal problems arise against Kant's theory. If I succeed in my argument, and yet the model still seems unpalatable, that may be bad news for Kant's account of freedom.

§12. The unity problem

Moral agents need to extend in time somehow. We sometimes undertake lengthy, even lifelong projects, and we must take responsibility for our past actions. If there is no way for me to say that *he and I* went out to have dinner yesterday, *you* booked some train tickets a week ago, and *she* enrolled in a university course almost a year ago, there would be so much as an agent in the world. Kant thought as much, as we saw in the importance he gave to freedom as a condition of imputability (§5), and is further evidenced in the *Critique of Practical Reason* and in the *Religion*. Firstly, it is presupposed in the postulate of the immortality of the soul. If we are not the same agents over time, it makes no sense to have as an article of rational faith that we should keep on living after death in order to progress towards moral perfection (KpV 5: 122). Secondly, when discussing his theory of justification[74], Kant notes that we must imagine God to judge our *whole lives*, since we have no insight over our moral disposition other than how we conduct ourselves overall (RGV 6: 67). Once again, this would be incomprehensible if we did not persist as the same moral agents over time.

---

[73] Although even two-standpoints supporters like Frierson (2003; 2010) agree to speak of a grounding relation between intelligible and empirical.

[74] In the theological sense.

Say, then, that you want to do without an intelligibly grounded empirical unity in reading Kant's theory of freedom. There must be something about our actions in the world, through time, that can trace them to ourselves as *free* – that is, intelligible – agents without resorting to our empirical selves' having an appropriate ground for that matter. How could we do it? Note that I am not raising the general epistemic worry about linking the bridge between concrete phenomena and their noumenal substratum; that, after all, would apply to my proposal as well, since I have not yet answered how we may judge an empirical unit as having any intelligible ground in particular. The problem, rather, is as follows: how can different phenomenal actions pertain to the same intelligible agent if not through an empirically unified mediator, who provides their common intelligible ground?

An option may be a return to Allison's Incorporation Thesis. If we do not hold that reason is "literally" efficacious, and, rather, what is meant by that is that only those incentives we incorporate as reasons can determine our wills (1990: 51-2), we may want to think that this "I" that "takes" incentives "as reasons" (37-8) is all the identity we need, in a similar move to the unity of apperception. I will tackle whether this works for our own agential unity in the next section; here, I will note a different problem this view has. How does this model translate into attributing actions to others?

We already noted before[75] that there is a major challenge in judging other people to be rational and have the intelligible ground they must have (cf. Saunders, 2016). Allison's approach, however, would make the challenge utterly intractable. Let us say Henry is baking a cake and, at the same time, chatting on his phone. From his perspective, he has incorporated maxims that result in his doing both those things, but there is no such clue from our side. We cannot reduce the incorporation of maxims to "rationalized" causation, on pain of resorting to a "wretched subterfuge" (KpV 5: 95). What we can *see* is that a certain (empirical) human being, called Henry, is causally responsible for the cake and the chatting, but this seems to be completely irrelevant for his setting of maxims[76]. If it were not, we would be judging an empirically unified being as (i) causing certain effects, (ii) having a certain intelligible capacity capable of determining said effects, through the incorporation of maxims, and (iii) as having determined as a matter of fact certain effects on the basis of that intelligible capacity. But that is exactly (Intelligibly Grounded Agents), since a capacity that can determine effects *is a ground*. Thus, either we deflate agential identity to a mere internal function, and we lose any clues whatsoever as to what makes external actions have authors, or we add in causality and grounds of determination to the picture and go back to our old model. We cannot cut down on our metaphysical costs turning the unity of the self into a functional unity and still claim that we can attribute actions in experience to agents as *free* agents.

---

[75] Cf. footnote 60.

[76] It may be objected that I am forgetting the difference Kant draws between choice and mere wish (KpV 5: 143n.), which already incorporates our perceived causal powers. But how can we do that if we do not regard ourselves as empirically unified beings who have a certain empirical character, with capabilities associated to it?

Another option that may fare better would be as follows. People, let us say, have agential unity, not in virtue of an empirical unity that sustains an intelligible ground, but because of a structure of maxims that defines their particular conduct as rational beings[77]. We might be able to infer the structure of maxims behind actions in the world[78], as it were, retaining the unique "fingerprint" of their agents. Let us say that somehow this is not an empirical unity of the sort I am defending. This could potentially work for many actions if we accept that we can unambiguously individuate agents through their structure of maxims. If maxims can be arbitrarily specific this may well be the case. However, a quirk of Kant's commitments regarding freedom makes this unsatisfactory as a general answer. We can only properly infer the maxims of action when we are dealing with *evil*. As Frierson points out, "[f]rom the failure to *act* in certain ways, one can infer the status of a person's disposition" (2003: 106; cf. RGV 6: 38). Given that we should especially care about being able to attribute good actions if there were such, it is an unfortunate deficiency of the model that it can only attribute evil. If a solution were to be sought in appealing to the "veil of benevolence" (MS 6: 466; Guyer, 2016: 160), we would gain the possibility of thinking of others as acting on good principles, but we ostensibly lose the capacity to individuate them through their maxims, since such veil only obscures what flaws we may perceive in others, without giving any more content as to how they are good people in particular. A veil, benevolent though it may be, cannot give insight into the intelligible identity of people.

In conclusion, it seems extremely hard to attribute actions to others if not through their causal responsibility, which is what we are acquainted with as a matter of fact. If we cut off causal ties because we cannot accept the idea that an empirical unity would be required for our moral enterprises, or that such unity should have such a thing as an "intelligible ground", it will be a nearly desperate task to find what phenomenal actions, taken as standalones, may have in common that noumenally unifies them.

---

[77] I want to thank James Clarke for this suggestion.

[78] See Esser (2008).

§13. The scope problem

We can now address the point we had pending last section. Maybe we cannot attribute actions directly to others through the Incorporation Thesis, but perhaps we do not need to do it. If we manage to attribute actions to ourselves by doing so, we may hope that a mix of analogy arguments and veils of benevolence will overcome the third-person problem. By doing this, we will encounter the second problem with denying empirical unity with an intelligible ground: the *scope* problem. That is, from our moral awareness of freedom alone there is no way to determine which phenomena correspond to us as agents.

We have already seen that we cannot really ascertain what maxims we act on, since our true motivations remain obscure to us, but let us say, for the sake of argument, that we do. We freely incorporate a maxim when we act, and it is thereby as good as an intelligible ground. Now, all we get by doing this is that certain maxims are determinations of our wills, but what happens with phenomenal actions, which is what we are after? Freely adopted maxims cannot ensure the success of an action, which is a point Kant gets back to constantly; only the maxim, and not the realization of is object, which is a contingent matter, has moral worth (GMS 4: 399-400). In that case, there is a gap between maxims and actions that precludes attribution via the incorporation thesis. Our practical identity is restricted to a set of a maxims the actualization of which into actions cannot but remain problematic if we do not have a causally efficacious empirical being to whom we attribute both maxims and powers[79] to act on. Suppose I have a weak will, whereas you are admirably disciplined and consistently get done whatever you resolve to do. We both may have a maxim to finish a paper by Friday, but you will actually have a paper, and I will have an uphill weekend ahead. If all we have to attribute actions with are our maxims, though, how do we tell?[80] Both the finished and the unfinished paper are compatible effects of the maxim "I shall finish my paper by Friday". My maxims do not suffice to attribute successful actions to myself, since they can always fail.

---

[79] I follow Watkins (2005) in understanding Kant's theory of causality as a powers theory, but not much of this argument hinges on that. Whatever the specifics of Kantian causality, it is the subject who provides an empirical character which grounds the effected changes. If we do not have any such subject, it is hardly comprehensible that its maxims should turn into effects in the world.

[80] Things get complicated by the fact that Kant seems to imply in the *Religion* that "frailty" of the will is a result of evil maxims as well. This, however, will hardly make the case for a practical identity approach to agential unity easier. If we understand maxims as policies of action we freely undertake, does it make any sense to understand weakness of the will as a chosen maxim? If we follow the *Religion* account, something *else* needs fixing.

McCarty (2009) gives an argument extracted from Kant's *Lectures on Metaphysics* for the determination of personal identity on practical grounds alone, by which "the phenomenal agent who can consciously say "I do" must be self-identical with a noumenal agent acting independently of both divine and natural causal determination" (2012: 120-1). On a similar line, Palatnik (2022) argues that we can assent to certain otherwise unwarranted[81] beliefs by considering them as structural to the practical point of view, via a special form of *acquisitio originaria*. I see three ways of construing either proposal. The first one is to accept (Intelligibly Grounded Agents) regardless. We would have an empirical unity to begin with, and we assent to its being rightly grounded through the practical assent of our "I do" or through the requirements of the practical point of view.

Otherwise, two options ostensibly remain open. One is to just say that our awareness of the moral fact immediately warrants a phenomenal agential unity. Such an inference, however, as Longuenesse (2007) notes, is a paralogism. We cannot immediately move from the "I do" of practical reason to a substantial unity amid phenomena any more than we can do it with the "I think" of theoretical reason. Nonetheless, and this is the last option, practical reason does have tools that theoretical reason lacks. If something is a necessary condition for the consecution of our unconditional moral goals, we are authorized to assent to it as an article of rational faith, as we can see in the *Canon* of the first *Critique*. So we may want to use an argument of this form:

(i) I have the duty to X.
(ii) In order to fulfil any duty, I must have agential unity.
(iii) Therefore, I (can assent to the belief that I)[82] have agential unity.

This seems promising, but, on closer inspection, it leads us to a hopelessly vague conclusion. To begin with, it is not too much of a stretch to suppose that many of our specific duties, rather than the general ones we find in the *Metaphysics of Morals*, already presuppose that we have an agential unity, and so, if perhaps the argument may serve to give as a warrant for our assenting to *that* unity being morally relevant, it does not help us in finding it on the first place. Let us try a more fleshed out example to show this.

(i) I have a duty of beneficence towards others.

---

[81] She focuses on the third-person problem, but extrapolating to the more basic issue of our own agential unity seems like a natural move.
[82] On the different types of transcendental argument, see Stern (2000).

(ii-a) In order to fulfil my duty of beneficence, I must be the creator of the universe, thus securing that others achieve happiness through my grounding of all natural laws.

(ii-b) In order to fulfil my duty of beneficence, I need nothing but the capacity to set the right maxim of action.

(ii-c) In order to fulfil my duty of beneficence, I need to be the same person as that phenomenal being called "Álvaro" from January 2023 to June 2023. Henceforth, I need to be the same person as that other one called "Alberto".

(iii) Therefore, etc.

The middle term of the transcendental argument has, as we can see, a lot of leeway in the specific form it may take. It can range from the most maximalist version (ii-a)[83], where we look for *assurance* in the resolution of our moral duties – which does not veer that far from Kant's own moral argument for the existence of God –, to the most minimalist version (ii-b), akin to the virtuous atheist discussed in section 87 of the third *Critique*. On what grounds exactly do we decide between these when all we have is a transcendental argument from the requirements of our duties? What is more, if we were to determine that both those options are far too extreme, what about our general duties precludes something like (ii-c), where our moral awareness corresponds to different physical persons across time? Unless we already have a clear empirical reference for the "I" that speaks in our transcendental argument, this cannot be determined. Therefore, we need empirical unity prior to any of these arguments from practical reason.

§14. The moral progress problem

The unity and scope problems speak to a foundational conceptual challenge in Kant's theory of freedom. If we do not presuppose from the beginning that a certain empirical unity has a particular noumenal ground, there will be nothing about the phenomena in and of themselves that will be able to tell us that they should be noumenally unified, nor will there be enough information from the mere "interest" practical reason takes on phenomena (KpV 5:99) to establish who we are as sensible selves. Without a presupposition of empirical unity, we have to face the threat of misalignment, so to speak, from both ends.

---

[83] See Wood (1984), Pereboom (2007).

There is a further kind of misalignment, however, that threatens to undermine the possibility of attributing actions if we do not accept the role of empirical unity. Let us grant that the scope and unity problems could be solved, and that we could somehow tell which phenomena follow from our intelligible selves. Even so, we could not do without empirical unity. In the *Religion,* Kant develops his picture of human agency with his famed theory of radical evil. Insofar as evil is imputed to us, it must consist in a freely adopted maxim that somehow runs contrary to morality (RGV 6: 21). We are not diabolical beings, because the moral law is always an incentive for us, but it is, Kant remarks, obvious from experience that evil exists (RGV 6: 32-3). Thus, he says, evil consists in a freely chosen subordination of the moral law to our sensible interest in self-love (RGV 6: 36). Now, insofar as our actions, despite being free, need a ground of determination, Kant argues, this evil maxim is fundamental, and cannot be “extirpated”. That is, we cannot simply choose to discard our evil maxim, since that decision would need to, itself, be based on a maxim – we would need to already be good in order to do that.

This has some worrying implications. Evil is prevalent and inextirpable, and yet, we have a moral duty to overcome it. The only way we have of conceptualizing this moral necessity, Kant says, is through “a kind of rebirth” (RGV 6: 47). Somehow, we must believe in the possibility that we shed off our old intelligible characters and become “morally another being” (RGV 6: 74). If this is so, however, our attempts to ground the attribution of actions in our mere intelligible selves is undermined *ipso facto*. By becoming new moral beings, we can, at most, take on a voluntary penance for misdeeds past, but in what sense are they *our* misdeeds? However virtuous we are in doing so, and as Kant makes explicit in his critique of the traditional doctrine of original sin, guilt is not transferrable. If moral progress is to make sense, it must be *us* who become better, and, thus, our past injustices must remain ours. If we only have our intelligible selves to attribute actions to ourselves, moral revolution amounts to a clean cut from our past, and identity falls with it.[84]

---

[84] An option that I have not covered would be to understand our overarching identity as that of a *noumenal* substance; this resonates well with Kant’s doctrine of the immortality of the soul (KpV 5: 122), and it has been put to the forefront by interpreters like Insole (2013: 114ff.). This is in all likelihood an element of the overarching picture of Kant’s theory of freedom, but it cannot abrogate the need for empirical unity. Substances, under Kant’s conception, do not subordinate their accidents, rather they exist in the way set by that which inheres onto them: “In regard to substance, however, [accidents] are not really subordinated to it, but are rather the way substance itself exists” (KrV A 414/ B 441). Further, there is an issue with individuating noumenal substances other than through their determinations – how would we do that without space and time for merely numerical differentiation? Thus, if all we have to characterize these persisting substances is what morality gives us, which is our intelligible

Even if we depart from Kant's rigorism and accept that our intelligible selves may be more complicated than a single fundamental maxim, perhaps returning to the idea of an individual structure of maxims or a life-project, Kant seems to be right in saying that examples of evil abound. We do not need any stretch of the imagination to come up with instances of twisted practical identities that, in order to conform to the strictures of the moral law, could not do with a mere reform, but would properly require a revolution.

Although this commitment to radical moral change will be challenging no matter what, (Intelligibly Grounded Agents) can accommodate it. Since what is required for attributing a free action is an empirical we judge to have the right *sort* of intelligible ground, a radical change in the content of maxims will not be a problem so far as it still is a free will, and so far as the empirical being is still the same one. Michalson (1990) showed concern regarding the possibility of grounding personal identity in empirical continuity, as potentially deleterious for freedom, since such unity would only be maintained through deterministic causal chains. Firstly, that is not exactly so; it is most likely that something like personal identity could not be solely based on the Second Analogy. The other two Analogies, as they regard persistence and the community between the different parts of our bodies, would probably play a role too. But, most importantly, so long as such causal chains are at all points grounded in the right sort of intelligible substratum, Kant does not seem to have any problem with its being deterministically conjoined in experience (RGV 6: 74). On the contrary, it must be so[85], or the unity of experience will be threatened.

I have argued that the most plausible interpretation for Kant's account of the relation between our intelligible selves and actions in the world is through an intelligibly grounded empirical unity. Even though Kant does not make it entirely explicit, it seems to be his working assumption throughout his discussion of our agency. Further, I have argued, it seems like Kant has no choice given his philosophical commitments but to accept this sort of intelligibly grounded unities. Without it, we would face three potentially fatal problems: we could not tell that different actions belong to one and the same intelligible self, we could not trace an empirical counterpart to our moral awareness, and the necessity of moral revolution for our

---

character, then they will provide no bedrock through moral change; we need a further sort of unity which we may then find grounded in a noumenal substance.

[85] Indregrad (2018) argues convincingly that a change in intelligible character must entail an empirical change as well. But that does not mean we cannot have empirical identity. He himself gives us two things that do not change through moral revolution: the laws that make up the framework of our empirical character, and the subject who possesses the powers the degrees of which vary.

progress out of evil and into good would make action attribution impossible. There may still be reasonable doubts, however, that such an empirical unity in possession of an intelligible ground can actually be made sense of. To argue that it can, will be the task of the next chapter.

4

THE INTELLIGIBLE GROUND OF NATURE

In the last chapter, I argued that Kant's theory requires empirically unified beings to be causally responsible of those actions we judge to be free. It is these beings, that we already judge to be free, who have the proper sort of intelligible ground required to make some of their actions distinct from the usual happenings of nature. However, this suggestion brings with it a considerable challenge. If we only have knowledge of appearances, and no investigation on them, however thorough, could ever yield any theoretical cognition about things in themselves, how can we ever judge an empirically unified being to have an intelligible ground? It will be the aim of this section to give a possible solution to this puzzle: the *Critique of Judgment* expands on Kant's theoretical philosophy in such a way that we can find a warrant to posit intelligible grounds for empirical unities, yet never attain proper cognition of them.

First (§15), I will provide an interpretation of Kant's account of empirical laws in the introduction to the third *Critique*. There, I will argue, Kant delves deeper into a problem introduced by the Appendix to the Transcendental Dialectic of the first *Critique*. In order to solve it, I contend, regardless of our interpretation of what a law of nature is according to Kant, we need to posit a supersensible[86] ground of nature from which particular laws may necessarily follow. Then (§16) I briefly go over Kant's philosophy of biology. According to Kant, organisms entail a special kind of inexplicability related to their existence as objective wholes, which elicits that we judge them as noumenally grounded as well. This, I countenance, authorizes us to count organisms among the objective junctures that the universal legal unity of nature grounds. Finally (§17) I use this extension of Kant's theoretical philosophy to respond to sceptical concerns raised against the tenability of noumenal freedom given the limitations of our knowledge under transcendental idealism, arguing that the noumenal ground of organisms can bridge the gap between our experiencing humans as empirical unities and judging them to have the right sort of ground. Since we are organisms, and we exhibit rational behaviour, we are authorized to posit that we have a supersensible ground such that makes our rationality possible.

---

[86] I will use "intelligible" and "supersensible" indistinctly (KrV A 257/ B 313).

§15. Making laws necessary

The Second Analogy has engendered strong emotions, for better or for worse, in most Kantian commentators. Whether it relies on a "*non sequitur* of numbing grossness" or not (Strawson, 1966), how strong of a conclusion Kant intended it to have (Allison, 1996), or whether it is a reply to Hume or not (Watkins, 2005), these are matters that still stir no small controversy. A point, however, that remains largely undisputed is that this principle of understanding was not, and could not have been, Kant's last word on the laws of nature. The Transcendental Deduction already warned that:

> "The pure faculty of understanding does not suffice, however, to prescribe to the appearances through mere categories *a priori* laws beyond those on which rests a **nature in general**, as lawfulness of appearances in space and time. Particular laws, because they concern empirically determined appearances, **cannot be completely derived** from the categories, although they all stand under them." (KrV B 165)

Coming from this background, we find one of the most relevant fragments of the extremely dense published introduction to the *Critique of Judgment*. Kant insists there that the particular empirical laws we surmise rule over appearances are left undetermined by the general law of cause and effect which is constitutive of experience:

> "But there is such a manifold of forms in nature, as it were so many modifications of the universal transcendental concepts of nature that are left undetermined by those laws that the pure understanding gives *a priori*, since these pertain only to the possibility of a nature (as object of the senses) in general, that there must nevertheless also be laws for it which, as empirical, may seem to be contingent in accordance with the insight of **our** understanding, but which, if they are to be called laws (as is also required by the concept of a nature), must be regarded as necessary on a principle of the unity of the manifold, even if that principle is unknown to us." (KU 5: 179-80)

Two important questions arise from this quote: (i) What exactly is the *necessity* empirical laws are missing from the principles of understanding? (ii) What can this unknown principle of the unity of the manifold be?

Kant's general point about the contingency of empirical laws is simple enough. There is nothing about the principle that "everything that happens must have a cause" from which we can derive laws the likes of "if you heat up water, it boils". As far as our understanding and its a *priori* legislation over experience in general are concerned, such a law is contingent; it is not a condition of possibility of experience that water boils when heated. Clear as this is, though, it does not immediately warrant the claim that laws *need* to be necessary. Why does Kant believe as much?

A precision is in order before carrying on. There is an obvious sense in which causal laws *have* to be necessary according to Kant (Hutton, 2021): cause and effect must be linked, per the Second Analogy, by a relation of necessity such that, if the cause obtains, then the effect necessarily follows. However, Kant seems to be pointing at a different sort of necessity in the *Introduction*. This general, 'formal' sort of necessity is not one that pertains to empirical laws or that is left undetermined by the understanding at all; the Second Analogy prescribes a necessary connection between appearances, that is what causality *is* according to Kant (KrV A 188/B 234). What are we missing, then?

Let us say that you fill up a pot with water, and you do two things in order to get it boiling: you turn on the hub to heat the pot up, and you ask it real nice to get the water boiling. All the understanding is prescribing to nature is that, when the water does boil, it must do so (i) caused by something, and (ii) in conformity with a law, but it has no say on *which* law it will end up being – this is down to experience (O'Shea, 1997). So, which is it? Sure, the experiment required to ascertain whether merely asking the pot to make water boil works is unusually simple, but the point remains that no mere perception is ever going to completely disambiguate what laws actually obtain in our experience; we could only ever, as Kreines (2017) remarks, get asymptotically close to an answer. And yet, understanding demands no less than a necessary connection between appearances. Kant's point then starts making sense. If empirical laws stood alone in front of us, with nothing that made them necessary, they would be undistinguishable from Humean regularities: increasingly better descriptions of the patterns we have encountered, but ultimately unbecoming of the title of a "necessary connection". On the other hand, if there exists a principle in virtue of which it is necessary that – say – if water is heated, then it boils, then there is a clear sense in which whispering to our pots does not cut nature at its joints. We may recall Kant's comments from the *Appendix* to the Transcendental Dialectic, which are tightly connected to the problem that concerns us here, when he wonders about our authorization to follow regulative principles of reason:

> "For then reason would proceed directly contrary to its vocation, since it would set as its goal an idea that entirely contradicts the arrangement of nature. Nor can one say that

> it has previously gleaned this unity from the contingent constitution of nature in accordance with its principles of reason. For the law of reason to seek unity is necessary, since without it we would have no reason, and without that, no coherent use of the understanding, and, lacking that, no sufficient mark of empirical truth" (KrV A 651/ B 679).

If we do not posit that there is a matter of fact as to empirical laws, which are the only possible instantiation of the principle of causality demanded by understanding, Kant is telling us that the rules under which we judge nature would lead us completely astray. The way in which Kant understands this "matter of fact" to which we asymptotically approach at all is through a common principle from which laws follow necessarily. If we are right that the rule "water boils when it heats up" accurately traces the joints of nature, then it is necessary that it does so, which is what our understanding demanded.

Why should it be a "principle of the unity of the manifold" that makes it necessary, though?[87] Laws could be made necessary by the individual essences of things without requiring an overall unity. However, if this were so, how would we assure that the laws are compatible and do not breach the unity of experience? This is a topic that preoccupied Kant since pre-Critical times. We find in the *Inaugural Dissertation* that:

> "*The substances which constitute the world are beings which derive from another being, though not from a number of different beings; they all derive from one being*. For suppose that they are caused by a number of necessary beings; the effects, of which the causes are free from any reciprocal relation, would not be in interaction. Therefore, the UNITY *in the conjunction of substances in the universe is a corollary of the dependence of all substances on one being*." (ID 2: 408)

And it coheres with Kant's *Appendix* notes on the systematicity of empirical concepts as well (KrV A 693/ B 721); same as we need to posit a unity of nature with regards to the empirical concepts we deploy on it, such that all concepts are species of an overarching genus and so forth, so need empirical laws to derive from an ultimate though inscrutable unity. This may look like little more than a rationalist affectation, but it speaks to an issue that modern metaphysics still reckons with. Power theorists, for instance, need to be careful about the

---

[87] A way to read that sentence could be as referring to the manifold of experience that the particular law synthesizes, and, thus, as referring to the necessity to posit real grounds, albeit individual, for laws. This seems very unlikely, however, since Kant had in that same paragraph used the term "manifold" to refer to the diversity of forms in nature.

"problem of fit" – different powers are to somehow align with each other as to not breach the unity of nature (Williams, 2019: 86). Kant's solution may ultimately be deemed inappropriate, but there is no reason to discard it as unacceptable from the get-go.

We have now elucidated in what sense empirical laws demand a necessity that understanding alone cannot provide. Now, it remains to be explored what this mysterious ground, the principle of unity of the manifold, could be. It will come as no surprise, given my overall project, that I take it to be a noumenal ground. Before putting forward what could be a controversial piece of metaphysics right in the middle of Kant's theoretical philosophy, though, I must acknowledge and address the diversity of opinions that exist in the literature around what Kant even means by an empirical law. It will be my goal to show that we will be hard pressed to avoid noumenal grounds whatever our stance on Kantian empirical laws is.

I broadly follow Messina (2017) in distinguishing two big families of readings around Kant's metaphysics of laws. On the one hand, *top-down* readings hold that the necessity of laws ultimately comes from an operation of our intellect, be it through deployment of the principles of understanding (Friedman, 2014) or by means of the systematising operations of reason (Guyer, 1990a). On the other hand, *bottom-up* readings take Kant to rely on the natures and essences of objects to ground the necessary laws that govern their powers (Watkins[88], 2005; Kreines, 2017; Massimi, 2017).

Let us start with the former, which would appear to be most hostile to such a metaphysically loaded companion to a quite central notion in Kant's philosophy. Even if we regard empirical laws as tracing back their necessity to either understanding or reason, we cannot ignore Kant's apparently undisputed claim that understanding alone is not going to give us the matter of particular laws, only their form. We still need to make sense of Kant's plea for a "principle of the unity of the manifold". We know that this principle is regulative, since it pertains to reflective judgment (KU 5: 180), and, much like those found in the *Appendix* (KrV A 666/ B 694), the principle cannot determine the specific content that the laws will have. However, Kant places a great importance on the particularities of transcendental idealism for the possibility of such a thing as regulative principles, which will showcase the merits of a noumenal ground.

---

[88] Though Watkins is one of the greatest exponents of a natures and powers-based account of Kant's philosophy of causality (2005: 244), he has proposed that laws in general, under Kant's conception, are prescribed by a competent authority – in this case, our understanding (see, on this controversy, Sethi (2023)). Not much will hang on whether Watkins qualifies as a bottom-up or a top-down theorist, even though he has clearly been most influential among the former camp.

Starting with the *Appendix*, Kant raises the issue, as we saw earlier, of how we can claim validity for principles that are not constitutive of experience, and that could thus run contrary to what nature actually has to offer. This requires some sort of deduction for the use of the ideas of reason, though not a transcendental one, since they do not legislate in experience (KrV A 669/ B 697). How do we do this?

Let us assume appearances were all there is. How exactly would our laws be different from Humean regularities then? Understanding, indeed, demands that we judge certain states of affairs as *necessarily* following others (and that is the sole mode of necessity we can cognize (KrV 227/ B 280)), but, as Kant ceaselessly makes note of, that still is not enough to fill in the gaps as to what the laws of nature actually are. And yet, if there is nothing undergirding appearances that may take on the role of making it true that a certain law *is* a law instead of a mere accidental regularity, then it is very hard to understand what Kant is going on about with regard to the necessity and universality of laws. Speaking of the maxims of reason, Kant says:

> “This law [the logical law of the *continuum specierum*] must therefore rest on pure transcendental and not empirical grounds. For in the latter case it would come later than the systems; but it really first produced what is systematic in the cognition of nature […] one can see clearly that the laws judge the parsimony of fundamental causes, the manifoldness of effects, and the consequent affinity of the members of nature in themselves reasonably and in conformity with nature, and these principles therefore carry their recommendation directly in themselves and not merely as methodological devices.” (KrV A 660-1/ B 688-9)

The principles elicited by the requirements of our intellect are not just useful fictions. Insofar as they relate to an actual need of understanding, without which we have no claims of objectivity at all (KrV A 654/ B 682), they have to at least possibly trace real regularities in nature. It is for this purpose that transcendental idealism opens up the door to “assuming a supreme ground” (KrV A 676/ B 704). Granted, we cannot claim knowledge about it, but insofar as we need regulative principles to fulfil what understanding itself demands from experience, we are obliged to posit that they actually refer to *something*:

> “I can have a satisfactory reason for assuming something relatively (*suppositio relativa*) without being warranted in assuming it absolutely (*suppositio absoluta*). This

> distinction is pertinent when we have to do merely with a regulative principle, which we recognize as necessary, but whose source we do not know, and *for which we assume a supreme ground merely with the intention of thinking the universality of the principle all the more determinately*" (KrV A 676/ B 704).

This "subtle yet important" distinction makes it (somewhat) clear in what sense the epistemic boundaries are kept under check while resorting to a ground of appearances to secure an objective regulative use for the maxims of reason. We do not determine any object, nor its real possibility, by means of these maxims, if anything because the ideas they rely on as grounds cannot be the object of any possible experience. There is no contradiction in holding both that we need to assume "a supreme ground" and that ideas "should not be assumed in themselves", but "only as analogues of real things" (KrV A 674/ B 702). If we could assume them *in themselves*, then we would be able to determine the ground on its own, and ostensibly derive from the ground itself conclusions about what *has* to obtain in experience. Rather than that, we always have to stick to the data of appearances, but such data cannot be all there is and express the systematicity, necessity, and so on, that we need it to express. Kant's meta-metaphysical point here is crystal clear: our metaphysical suppositions in theoretical philosophy are only warranted to the extent that experience requires them, and they can *never* usurp the place of observation:

> "For that [systematic unity] we posit a thing corresponding to the idea, a Something or a real being – by this fact it is not said that we would extend our cognition of things with transcendental concepts; this being is grounded only in the idea and not in itself, hence only in order to express the systematic unity which is to serve us as the standard for the empirical use of reason, without settling anything about what the ground of this unity is, or about the inner property of such a being on which, as cause, it rests" (KrV A 674-5/B 702-3).

As Kant points out repeatedly, what we know of these indeterminate objects is what they must *do*, not what they are. This distinguishes them from the actual objects of our knowledge, but it also makes sure that they are not mere thought-entities; they have to serve as "substratum" (KrV A 697/B 725), as possible indeterminate grounds of experience:

> “What this original ground of the world’s unity is in itself ought not have been thought through this, but rather only how we ought to use it, or rather its idea, in relation to the systematic use of reason in regard to things in the world.” (KrV A 697/B 725)[89].

In short, as it pertains to regulative principles in general, Kant is consistent about their reliance on a “ground” or a “substratum”, despite not being able to cognize it as a determinate object at all; and he is deeply preoccupied with securing a *sui generis* sort of validity for these principles, which would seem superfluous if they were nothing else than methodological fictions: indeed, he is explicit in that they *are not*. But one can still be rightly sceptical that this would involve anything to do with an *intelligible* ground in particular. Could it not be just an unknowable ground we posit in appearances, but which has nothing to do with the one already entailed by transcendental idealism? This multiplication of the grounds is not parsimonious or particularly helpful[90], but, even then, once we move to the *Critique of Judgment*, the appeal to the supersensible becomes unavoidable.

We have already seen that the regulative principle Kant is after with empirical laws in particular is such that it must (i) ground their necessity, as to distinguish them from mere accidental generalizations, and (ii) be a principle of unity, that may “ground the possibility of the systematic subordination of empirical principles under one another” (KU 5: 180). Kant quickly characterizes this principle as that of the *purposiveness of nature*: nature must be considered “as if an understanding contained the ground of the unity of the manifold of its empirical laws” (KU 5: 181). Why is this? Ideally, we would want a general principle from which all the particular laws would derive, but, due to the peculiarities of our intellect[91], this is impossible. Since our understanding is discursive, our concepts operate by marks – the more general a concept is, the fewer distinct marks it has and, thus, the poorer it is (L 9: 98), and vice versa:

> “Our understanding thus has this peculiarity for the power of judgment, that in cognition by means of it the particular is not determined by the universal, and the latter therefore cannot be derived from the former alone; but nevertheless this particular in the manifold of nature should agree with the universal (through concepts and laws),

---

[89] Cf. DO 8:136; FM 20: 279.

[90] Moreover, consider: “Thus if one asks (in respect of a transcendental theology) first whether there is anything different from the world which contains the ground of the world order and its connection according to universal laws, then the answer is: Without a doubt.” (KrV A 695-6/ B 723-4)

[91] McLaughlin (2003: 215) is critical of this strategy. I side, however, with Zuckert (2007) and Geiger (2022, ch. 4) in considering that, given Kant’s independent and consistent characterization of our intellect as discursive (that is, as proceeding from parts to wholes), this is not an unwarranted move.

> which agreement under such circumstances must be quite contingent and without a determinate principle for the power of judgment." (KU 5: 406-7)

It then seems that the solution to the problem that beset us – grounding the necessity of empirical laws despite their appearing as contingent to our understanding – is strictly beyond our powers, since we cannot regard a particular case as being necessitated by the more general case, or a part as being determined by the whole. We have access, however, to an analogy that allows us to bypass the limitations imposed by our lack of an intuitive understanding[92]: final causality (KU 5: 405-6).

Take a favourite among philosophers' props, a watch. It depends for its existence on the different parts that make it up: there was no watch at all before putting the gears, the dial, etc. together. In that sense, the parts being assembled as they are is a contingent matter; it happens that they are together and in that sense they form a composite object, but the "watch" has no existence over and above the parts that would make their arrangement necessary by any means. Except, this is not quite true[93]. Watches do exist "as wholes" in a specific way: in the head of watchmakers (KU 5: 372). When a watchmaker crafts a watch, the arrangement of parts stops being contingent: the watch as an actual object has its existence grounded on a *concept*, and so it does not depend on its parts coalescing of their own accord to form it[94] – it exhibits purposiveness (KU 5: 360). Purposiveness, then, is the second best thing after intuitive understanding. Had we got such an understanding, from the very concept of this individual watch we could derive with necessity what parts go in its making; we do not, but we do not need to, since humans make watches through concepts. In the case of nature as a whole and its legal unity, we have absolutely no credentials to claim insight on *whether* it is conceptually grounded, let alone *what* this concept should be, but positing such a ground is the only way in which, given that we cannot understand how the most general of laws should ground the more particular ones, we can conceptualize the necessity of the empirical laws of nature. The ground of nature is then assumed to be a legal layout, as if the grand plan of an understanding different from ours.

---

[92] We can interpret Kant's move either as finding a substitute for intuitive understanding, or, following Gardner (2015: 31), as nevertheless thinking nature as an object of intuitive understanding. These two approaches need not be in conflict, we may have to *think* nature as an object of intuitive understanding *and* approximate it through teleology and analogy to our own understanding. It could further be argued, like Heidegger does (1929, §5) that discursive understanding being a feature of a finite intellect, any creative understanding we should posit as the ground of nature would be intuitive.

[93] On the kind of inexplicability that artifacts share with organisms, see Ginsborg, 2004.

[94] Which would usually be the case: "This [physical-mechanical] conception of matter entails that the material parts of these objects have equal claim to being considered as individuals in their own right, or that these material objects, or the events/motions they undergo, might just as well be taken to be parts of some broader material interactions or objects" (Zuckert, 2007: 107).

How can this be, if the material world follows mechanistic laws, and, thus, (at least, external) appearances inexorably stick to parts-to-whole relations? Here, we finally reach the conclusion we were after: transcendental idealism makes this possible. There is much debate on what the conflict from the Antinomy of Teleological Judgment actually *is*[95], but we can go straight to what ultimately is Kant's solution to it. The last section of the *Dialectic of Teleological Judgment*, "On the unification of the principle of the universal mechanism of matter with the teleological principle in the technique of nature", gives a clear answer to how Kant regards that these two principles – the legal ground of nature and the mechanic laws we cognize a priori – must be made compatible[96]:

> "The principle which is to make possible the unifiability of both in the judging of nature in accordance with them must be placed in what lies outside of both (hence outside of the possible empirical representation of nature) but which still contains the ground of both, i.e., in the supersensible, and each of these two kinds of explanation must be related to that." (KU 5: 412)

The ground of appearances, i.e. the supersensible, gives us the space required to make sense of the compatibility of these two principles. We cannot straightaway postulate that the supersensible looks like a purposive system of laws, but (i) we cannot help but regard nature from a purposive stance, for the reasons already expounded[97], (ii) only by appeal to the substrate of nature that transcendental idealism already brings with it can we posit *something* that would take on the role of a legal ground for nature. Thus, even if we regard the necessity of laws as proceeding from an operation of our intellect, the presuppositions brought about by the regulative principles we need in finding those laws and regarding them *as* laws still entail an intelligible ground of the legality of nature.

If this is so with the more deflationary top-down reading, the argument for a noumenal ground in bottom-up ones will be much more straightforward. For one, under some readings, the "unknowable essences" that are said to ground the natures of causally powerful objects may already be noumenal[98] (cf. L 9: 61). But let us suppose that, despite being unknowable, these grounds are not intelligible themselves. The task of getting to know the natures of empirical beings may be asymptotical (Kreines, 2017), but it is always constrained to experience after all. However, as with all indefinite tasks in Kant, we have already seen (§4, and earlier this section) that hopes of completion cannot lie in experience itself; after all, experience can only tell you how far you have reached in your approach to such knowledge, not that you have a warrant for the unconditioned goal of such a process. There are, as we have seen, good reasons to locate such a *focus imaginarium*, using Kant's own lingo, in the noumenal ground of appearances.

---

[95] Quarfood (2014); Juárregui (2020); Geiger (2022).

[96] See O'Shea (2013) on this thesis as the main difference between Kant and modern teleological stances like Dennett's.

[97] I follow Zuckert (2007) and Geiger (2022) in regarding that the *Critique of Teleological Judgment* forms a cohesive whole with the *Introduction*, and thus does not exclusively refer to organisms, but to the general problems presented at the start of the book. I will have more to say about organisms in §16, but it seems warranted, if anything, by a principle of charity that Kant would attempt to respond the points he set out to answer in the introduction, instead of veering to a tangentially related topic and leaving it at that.

But even if we insisted that these grounds do not need to overstep the phenomenal/noumenal boundary, recall that Kant had two requirements that empirical laws had to obey: they must be necessary, and they must be grounded in a principle of unity, as to not breach the unity of experience. While particular grounds, perhaps underlying the empirical character of substances, might find a place amid phenomena, a unified ground from which all laws derive necessarily is certainly beyond possible experience: our understanding, once more, cannot determine objectively what specific members fall under the umbrella of a wider genus by merely starting from the genus. Only through analogy with the final causality we exhibit as beings with a will can we *think* – not cognize – such unified legal layout, and, as Kant says, such purposiveness of nature only finds room as an entailment of the grounds of appearance transcendental idealism forces us to admit.

In conclusion, then, regardless of whether we find the source of the necessity of particular laws in the operations of our intellect or in real grounds, the ultimate requirements imposed by what Kant understands by an empirical law and how empirical laws must fit together into a system of nature take us to the realm of the supersensible. We have to posit an intelligible ground of which we know nothing more than what it does: it grounds the legality of nature in such a way as to secure the necessity and unity of particular empirical laws, through a purposive image of the legal layout of nature made compatible with the mechanism of pure physics. A last objection we have to face before moving on, however, has been raised by Zuckert (2007) against Guyer's (1990a, 1990b) rather pessimistic outlook on the *Critique of Judgment*. According to Guyer, Kant's third *Critique* assessment is a partial retraction from the promises of the *Critique of Pure Reason* regarding the legislating powers of understanding; after the *Critique of Judgment*, we have to regard the necessity of empirical laws as a happy accident:

---

[98] This is, perhaps, the case in Massimi (2017).

> "We must presuppose that they are systematic, but we also recognize that it is a lucky accident that they are. Kant does not explicitly retract the first *Critiqu*e's doctrine of transcendental affinity and the entire metaphysical picture it implies, the picture on which we unfailingly impose complete order on the utterly plastic material furnished to us by remarkably cooperative things in themselves. But once he has linked the ideal of systematicity so closely to such fundamental requisites of the possibility of experience itself, an admission like this comes pretty close to the surrender of such a metaphysical model of our relation to reality" (Guyer, 1990a: 37)

Zuckert charges against this reading, which she identifies as a "thin utility" picture of purposiveness (Zuckert, 2007: 38) for being ultimately too weak and uninformative. The principle of judgment Kant is after in the third *Critique*, Zuckert argues, is an *a priori* principle that plays a central role in the actual formation of our empirical concepts (2007: 48). I completely agree with Zuckert that merely pointing at a noumenal ground we must assume is out there, and of which we can have no knowledge, would be a quite disappointing result for the *Critique of Judgment*. Furthermore, I think her reconstruction of purposiveness, as well as Geiger's defence of a theoretical usage of aesthetic judgment (2022, ch. 5), are invaluable in understanding this complex and often neglected piece of the critical edifice. However, I do not see why any of this should be in detriment of our need to assume that the ground of nature is a certain way. On the contrary, doing so moves transcendental idealism to the forefront of Kant's conception of regulative principles, making it clear that he is not a mere fictionalist. It is because we have to assume that appearances have a ground unknown to us that the demands of understanding, reason and the power of judgment can be expected to be met, and these principles, in all their inner *a priori* complexity, can have a certain kind of objectivity, even if they do not reach the rank of objective knowledge that is reserved to constitutive principles.

§16. What makes organisms special?

In the last section, we rather lengthily explored the problem Kant encounters with regards to particular laws of nature and concluded that their required necessity and unity compels us to posit a ground to the world *such that* it can undergird a purposive legal layout for mechanically ordered appearances. However, I have barely mentioned the purported main topic of the

Critique of Teleological Judgment: organisms as natural ends. This has been on purpose, to advance a relevant point about the goals of the *Critique of Judgment*. As we have seen, we can easily read the resolution to the Antinomy of Teleological Judgment as referring to demands of empirical laws in general, not just to the anomalous objects of experiences that organisms are. I follow Geiger in considering that:

> “our judgments of organisms *reveal rather than ground* the a priori transcendental principle of the purposiveness of nature for our cognition. The oddity of the argument from the organism is worth underscoring: It is an argument from a contingent empirical fact to an a priori principle.” (Geiger, 2022: 81).

Let us unpack this. Organisms, for Kant, are a tangible culmination of all the problems we had to face regarding the unity and necessity of empirical laws. They are phenomenal objects in time[99] that therefore cannot but be causally regular (*contra* Ginsborg, 2004; with Zuckert, 2007; Illetterati, 2014; Geiger, 2022). However, whereas the distinction between wholes and parts among the objects of mechanics is not completely meaningful, “[o]rganic objects seem […] to be identifiable non-arbitrarily as single, unified objects” (Zuckert, 2007: 108).

Consider a boulder flying off after being thrown by a catapult. It is ostensibly an empirical substance with certain properties. The boulder reaches its target and, on impact, breaks. What has actually happened to the boulder? In a sense, it has ceased to exist; the object that had a certain set of properties when airborne no longer obtains in the world. At the same time, however, the loss does not appear to be terribly tragic. The boulder does not strike us to be so much a whole as an *aggregate* of whatever crystals or molecules that happened to be held together, and no longer are (cf. ID 2:390)

But there exist certain beings that appear to us as wholes foremost and from the beginning. They synthesize their parts into a whole despite our judging those parts as contingent. We judge them as contingent because we can only validly judge wholes to depend on their parts for their existence, as they do. A non-arbitrary whole would be such that the parts depend on the whole for their *existence*. I do not think we should be too harsh on Kant for thinking it hard to conceive that the *parts* of something should depend, for their existence, on the whole – *whatever* is a whole aside from its parts? And yet, there appear to be such beings:

---

[99] See footnote 56.

organisms (KU 5: 370-1). An elephant uses its trunk to reach the tree leaves that it will digest and transform into new cells that will make up its trunk. A lizard gets its tail caught between two stones and it will regrow it. A chicken lays an egg that will turn into a new chicken. Insofar as biological facts are part of our experience of the world[100], there are parts that do not coalesce together arbitrarily like the pebbles a boulder breaks into; rather, there are parts that seem to be bound by a necessity we cannot perceive as such.

Let us go back to particular empirical laws for a second. Water should necessarily boil when we heat it up, and not when we ask the pot to make it boil. That this is the case *necessarily* (granted that we got it right) can only be made sense of, Kant says, by assuming that somehow the ground of nature *makes it* necessary. Note that the only thing compelling us to posit legality in the ground of nature is the fact that we need *some* law, we need it to be necessary, and this is our best shot at it. We can, however, imagine, unlikely as it is, that we are wrong about water being a natural kind. Let us summarize the argument for the purposiveness in general of nature we sketched in the last section:

(i) Empirical laws need to be necessary.
(ii) In order for them to be necessary without breaching the unity of nature, they all have to follow from a common ground.
(iii) The only way for us to conceptualize that empirical laws should follow from a common ground is if this common ground is as if created by an understanding, that is, if it is a purposive ground.
(iv) Therefore, we must posit a purposive common ground for empirical laws.

Which empirical laws are actually grounded purposively is always an open question, up to scientific investigation. With organisms, on the contrary, the route towards purposiveness is rather different:

(i) There are beings whose parts, according to our judgment, depend on themselves as wholes as much as themselves as wholes depend on their parts, and which have not been created by an intelligence. (KU 5: 373)
(ii) The only way for us to conceptualize that the parts of a being could depend on the whole (see §15) is if they are ends.

---

[100] And it seems that they are, even if – as it could so happen – they have to employ a different "grammar" than other provinces of our knowledge. See, on this, Thompson (2012).

(iii) There are beings we must conceptualize as ends despite not having been created by an intelligence (KU 5: 374), which we call natural ends.

(iv) We can only judge natural ends by remote analogy with the ends of human art, as existing by virtue of a purposive ground. (KU 5: 375).

(v) Therefore, we must posit a purposive ground to nature such that it makes these beings possible. (KU 5: 376)

Whereas purposiveness for nature in general guarantees the necessity of any laws may happen to obtain, it also has to guarantee the possible existence of *those organisms* in particular that are otherwise inexplicable for us. This does not get us any farther than the regulative principles we discussed before; as Kreines notes, we are never certain that natural ends actually exist, we are only forced to think so by the demands of our intellect (Kreines, 2012) (KU 5: 399). However, organisms "first provide objective reality for the concept of an end that is not a practical end but an end of **nature**" (KU 5: 376), they are– quite literally – the living and breathing image of an empirical being that instantiates a natural kind present in the purposive ground of nature. They are the model for what it means for a concept to objectively apply in appearances; we cannot help but consider that the laws that regulate them are actual joints of nature:

> "[...] the characteristic regularities exhibited by an organism are part of the natural order. Such biological laws as the law by which oaks produce acorns and acorns in turn produce oaks are genuine laws of nature, and grasping them is part of what is required if we are to achieve a full understanding of nature's workings" (Ginsborg, 2004: 61).

§17. How to encounter a free being

We are now ready to give a proposal of how we may judge ourselves and others as free beings, that is, as empirically unified beings with the right sort of intelligible ground. Let us recall the requirements for free agency we laid out in §11:

(Intelligibly Grounded Agents) A phenomenal action a_P is a free action if and only if:

(i) An empirically unified agent A is causally responsible for a_P,

(ii) A has the right sort of intelligible ground,

(iii) A is causally responsible for a_P in virtue of A's intelligible ground.

The problem we closed off chapter 3 with was the dubious prospects of judging any empirical being as having an intelligible ground in virtue of which it could be causally responsible of actions. Now, however, things look more plausible. Empirical laws – which make up empirical characters – need to be grounded in a way as to allow them to be judged purposively, to understand how they may be necessary despite being particular laws (KU 5: 180). Whatever this ground contains is unknown to us, but we at least know that it has to account for the existence of organisms, which is inexplicable for us otherwise (KU 5: 376). As organisms ourselves, we are thereby authorized to judge ourselves and our fellow humans as having a certain intelligible ground in virtue of which we are causally responsible for actions. In §13, I argued against Palatnik's (2022) proposal of a moral *acquisitio originaria* that could justify our treating others as free beings, but her argument was not unsound; rather, I contended, it was too hasty. Without the right sort of empirical unity, the demands of morality are woefully vague as to determine who can be free, but since humans are organisms, that is, intelligible grounded empirical wholes, there is a vacancy open for us to encounter the *right* sort of intelligible ground, rational nature[101]. We can *now* use the warrant from practical reason, since we have found a sort of empirically unified being that is amicable to free grounding.

Two main objections stick out. First, under what pretence do I derive individual intelligible grounds from the general purposive ground of nature the third *Critique* argues for?[102] Kant, after all, tries to minimize the role of the said ground as much as possible, calling for usage of mechanical explanation as far as it can reach (KU 5: 411). Moreover, he puts forward as a biological hypothesis for the origin of species a sort of "proto-mother" from which all existing living beings could descend through mechanical means (KU 5: 418-9), thus reducing the explanatory role of teleological grounds to its minimal expression. All of this is true, but it does not pose a problem in the least for the model expounded here. Temporally speaking, all of our determinations as organisms may find their beginning in a proto-mother of all life on Earth. However, how must this mother be in order to quench our explanatory demands, and what is the problem she is set up to solve? Kant tells us:

> "And yet ultimately he must attribute to this universal mother an organization purposively aimed *at all these creatures,* for otherwise the possibility of the purposive forms of the products of the animal and vegetable kingdoms cannot be conceived at all. In that case, however, he has merely put off the explanation, and cannot presume to

---

[101] "Reason is thus the persisting condition of all voluntary actions under which the human being appears." (KrV A 553/ B 581).

[102] I want to thank Pavel Reichl for raising this concern.

> have made the generation of those two kingdoms independent from the condition of final causes." (KU 5: 419, emphasis added)

Kant held this idea consistently, even before the critical period. In his essays on human races, he already explained the variety among humans in terms of pre-existing "seeds" (*Keime*) in primordial humans that had yet to be differentiated into races (VR 2: 434). What needs explanation in organisms are the particular and actually existing wholes that exhibit formative power (KU 5: 374) and their forms. The purposive ground of nature can be no less specific in its content than the individual organisms it is meant to make conceivable. We cannot help but judge chickens – each and every chicken! – as an undoubtably real part of this world, the seeds of whose behaviour lay dormant but real for millions of years before they actually came to be.

This may sound bizarre and speculative, but consider that this applies equally in the case of non-organic laws. Water boils when heated because it is a necessary law of nature that it does, and this law of nature, insofar as it has a noumenal ground, did not come into existence together with water – it was "already there" before the first molecule of water was created[103]. When water boils, we do not question that it instantiates the law it does, and the fact that the molecules in a pot of water ultimately trace their causal origin to the fusion of elements in stars is indifferent to this. The same goes with organisms, with the difference that we cannot help but include them in our ground of nature. They belong to an overarching unity of nature, but the ground they instantiate is *theirs*, and it grounds *their* actions[104]. There is no contradiction between the ground of nature being one, and it being specific enough in the particular laws derived from it as to take on the role of particular ground of an empirical being – indeed, this seems the only way about it if it is meant to solve the problems Kant set it out to solve. To use Frierson's (2005: 19; also 2014) notation for the grounds of causal relations:

… Heat → Boiling water
↑
Water's disposition to boil (it instantiates the appropriate law)

And:

… Chicken → Egg → Chicken …
↑ ↑
Purposive disposition of chickens to produce new chickens

---

[103] For an argument in favour of this reading, see §7.

[104] If they did not belong to a common ground, we would face the same problems regarding the unity of experience as we did with empirical laws in general.

These dispositions, or real grounds, as members of the overarching legal layout of nature, are outside the causal chain, but the kinds they track (water, chicken, etc.) are always individually instantiated inside of it. Thus, the overarching ground of nature contains and lends necessity to particular grounds, both for organisms and law-abiding empirical behaviour in general.

This model, however, seems to lead us straight into a second objection. It may well be that, as organisms, we can judge humans to have an intelligible ground and find room for freedom therein, but how are they different from *any other organism*, since they all have an apparently equal claim to such supersensible ground? Joe Saunders (2016) raises this objection explicitly[105]:

> "Imagine that you are a super-scientist who can predict the behaviour of everything under the sun […] you go the supermarket, and have to decide what type of things you should eat. You rule out the furniture because it has little nutritional value. What remains are a variety of vegetables, animal products, living animals and humans. How can you differentiate between these different types of things? […] Of course, we think that these different types of things possess different moral status. And a huge part of this, for Kant, is due to the fact that human beings are rational agents. But how are we supposed to have access to this? How are you, the super-scientist that you are, supposed to pick out the parts of nature that are rational agents, given that all of nature is entirely determined by natural necessity?" (2016: 173)

The problem Saunders raises can take the form, if I am reading correctly, of a dilemma. Either the peculiarities of organisms are enough for us to judge humans as free, but then we have to consider every organism to be free; or we have to take phenomenal data into consideration when ascribing free will, but this ought to be impossible, since all things phenomenal are determined by natural necessity. I contend that the second horn of the dilemma is based on a false presupposition. Saunders rightly notes that experience *itself* can never get us beyond the freedom of a turnspit (2016: 172), but this does not mean that the data of experience can have no imprint whatsoever on the sort of grounds we have to posit via the *a priori* principle of reflective judgment – in fact, as Geiger notes in a quote mentioned earlier, what is shocking about the role of organisms is that they reveal an *a priori* principle through particular, eminently contingent experiences (2022: 81).

---

[105] In his 2016 paper, Saunders is preoccupied only with the problem of recognizing *others* as free, accepting that we may have a privileged first-person access to freedom. If the argument I put forward in chapter 3 is solid, however, we can probably extend this issue to our own freedom as well.

As I understand it, the move by which the principle of purposiveness of reflective judgment authorizes us to posit a ground for organisms can be broken down into two steps, which we may call 'formal' and 'material' respectively[106]:

> **Formal step**: whenever I experience a being that I cannot help but judge as such that in it "everything is an ends and reciprocally a means as well" (KU 5: 376), I judge it to be a natural end, and I must posit a supersensible ground of nature that may make natural ends possible *in general*.
>
> **Material step**: whenever I judge a being as a natural end, I must posit that, whatever it is in experience that warranted that judgment, must be made possible *in particular* by the supersensible ground of nature that makes natural ends possible (KU 5: 419).

When we experience a chicken as a natural end, we are doing two things. On the one hand, the fact that it is a chicken is not immediately relevant to our positing a supersensible ground of nature; *any* natural end would trigger the need to posit such a ground. On the other hand, however, the ground we posit cannot be just anything. The experience that made us posit said ground was, as it happens, that of a chicken. Following the typical formula we have already seen for intelligible grounds (§15), the ground we posit must be, although inscrutable in its inner determination and real essence, *such that* the behaviour of chickens is rendered possible.

We can now understand how we would go about judging humans and not a shrub or a sponge as free beings. In the formal step of our judgment, we experience humans as natural ends *überhaupt* and thus open the space for an intelligible ground whatever it may be – this is the point where we are equal to all other organisms. But in the material step, we demand that the ground account for the kind of things humans do: sing, have opposable thumbs, speak a language, cook, brandish weapons, write love letters, and, presumably, use reason and act according to maxims. This, I surmise, is where we can make use of Kant's "sensible signs" (KrV A 546/ B 574) and "traces" (KpV 5: 85). Rational behaviour, Kant is telling us with this, is conspicuous. We can usually tell when someone is drunk or sober, or when someone is too young to make right judgments – that is not what is problematic in Kant's account (KU 5: 196n.). What is not conspicuous is that these behaviours we observe should find their headquarters in a noumenal ground, but this is precisely what the requirements of judgment allow. The formal step of our judgment of purposiveness opens the gates to the intelligible ground, and the material step 'incorporates' those behaviours that were in need of purposive explanation into that vacant intelligible level.

---

[106] Compare to Kant's characterization of symbolization: "the power of judgment performs a double task, first applying the concept to the object of a sensible intuition, and then, second, applying the mere rule of reflection on that intuition to an entirely different object, of which the first is only the symbol" (KU 5: 352). In the *Prize Essay*, he speaks of symbolization in the same terms as he did *suppositio relativa* in the Appendix to the Transcendental Dialectic (FM 20: 279). For a closer look into the role of symbolization in judging organisms, see Breitenbach (2014).

We can never attain apodeictic certainty that the ground of our actions is of the right sort, and Kant is very much aware of this (RGV 6: 38). I will go over this point in detail in §20, but I can admit already that there will always be a level of uncertainty involved in our attributions of freedom. Insofar, however, as we have a practical interest in our freedom and that of others, which we do, this model suffices to make sense of how we can be in free control of our phenomenal actions. The experience of rational action does not give a strong enough warrant for us to claim the objectivity of freedom, Kant is clear enough in that regard after the second *Critique*. But we know that, if freedom is real – which we must suppose, given our awareness of the moral law –, then it will appear as rational behaviour, since free actions are, by their own nature, rationally grounded. Only humans, then, are prime candidates for free agency.

Since we are organisms, we have no choice but to judge that we have an intelligible ground; since we exhibit rational behaviour, we must consider that such rational behaviour is due to said intelligible ground. With that, once we gain practical assent for the reality of freedom, we have all the pieces necessary to judge particular actions as freely caused, insofar as (i) we are in causal control of them, (ii) we are judged as having the right sort of ground, since we are rational animals, and (iii) those actions were due to that ground we have as free beings. The experience of humans as rational animals gives us a determinate space amid phenomena which free beings can occupy, since, *if* there are free beings, they will be intelligibly grounded in such a way as to make rational behaviour possible. Given that we are compelled by morality to affirm that there are free beings, we can practically affirm that human beings, when exhibiting the signs of rationality, are free. Once more, that does not mean that the empirical signs of rationality let us infer freedom, it is rather the other way around: since we have to assert freedom, we can only posit it for a select set of beings, that is, those with apparent rational behaviour.

This approach, further, does away with the temptation of privileging the first-person stance over all others – a quirk typical of Beckian/Rawlsian[107] readings of Kant's theory (§6). Since we can only judge a phenomenal action as freely caused through causal attribution to a unified phenomenal being (§11), our own actions are only legitimately deemed as free actions when they are a result of our powers as intelligibly grounded physical beings. Our judgment, then, applies once and for all to humans in general[108], and if scepticism strikes, we ourselves fall with our peers. This solution is avowedly anti-solipsistic.

In conclusion, then, the requirements brought forward by the power of judgment help us solve the outstanding issue of how to judge empirical units as intelligibly grounded. We need to posit a supersensible legal layout of nature to account for the necessity of empirical laws, and it must be so as to make organisms possible. Humans are organisms – empirical wholes – who exhibit rational behaviour. Since that is what we can expect from a freely grounded being, we are authorized to judge that the freedom we know to be real through our awareness of the moral law grounds the peculiar beings we are.

---

[107] See Rawls (2000: 322ff.)

[108] Humans in general have the right sort of ground to act freely; not every human does always.

5

MAKING SENSE OF TIMELESS FREEDOM

A free action is caused by an empirically unified being with, and by virtue of, the right sort of intelligible ground. We can attribute those grounds to humans because we are forced by the requirements of judgment to regard organisms as objective part of the legal substrate of nature, and because humans exhibit the sort of rational behaviour that a free being in the world would have. In this final chapter, I want to address some concerns that may remain after my exposition of the matter. First, (§18) I go over the core issue regarding the timelessness of freedom. If freedom is atemporal for Kant, how can it be relevant to our very temporal practices? I will argue that, by understanding the atemporal aspect of freedom as the ground of our actions, there should be no issue with timelessness, and it can even bring some reasonable practical implications. Then, (§19) I address the question of alternate possibilities in Kant's theory of freedom, and how moral change may occur in the model I have presented. Finally, (§20) I reply to potential concerns that I have overstepped the boundaries of transcendental idealism by making such substantial claims about the noumenal ground of nature.

§18. Demystifying timelessness

The atemporal character of Kantian freedom is the bane of sympathetic readers, and a prime target for critics of all extractions. However, all things considered, and given what we have seen so far, I will argue that the timelessness of freedom is neither mysterious nor deleterious of our actual practices.

For a start, we should fend off any mythical imagery of a disembodied choice prior to time itself that then somehow shapes one's empirical life. The atemporal aspect of freedom has nothing to do with what we as agents perceive and voluntarily do, and everything to do with the metaphysics of causality in transcendental idealism. Kant himself is very careful to note that, when we speak of any timeless "deeds", we must not mistake them for our usual free actions:

> "Now, the term "deed" can in general apply just as well to the use of freedom through which the supreme maxim (either in favour of, or against, the law) is adopted in the power of choice, as to the use by which the actions themselves (materially considered,

> i.e. as regards the objects of the power of choice) are performed in accordance with that maxim” (RGV 6: 31).

There is a decisive difference between the “fundamental deed”, and all the other deeds that are grounded therein. If my reading so far is right, this fundamental deed *is* the “right sort of ground” in virtue of which we causally produce a phenomenal action. Insofar as our actions, uniquely, can be causally attributed to this fundamental maxim (through the grounding relation discussed in (§4)), our actions respond to an order different to that of mere sensitive necessitation: they follow the moral law, or they subordinate it to pathological impulses. The moral law is nowhere to be found in the natural world, our rational nature brings forth a new sort of ground into the world’s order. The grounds of the actions of other animals, meanwhile, are not rational, and thus remain closer to mere natural laws than to free action.

We can then see that the actions we actually choose do not need to be themselves timeless. They are temporal decisions whose causal efficacy is undergirded and made possible by a ground that is atemporal and strictly unknowable to us. Thus, our entire phenomenology of free action remains resolutely temporal. It is in this sense that I partially agree with Allison when he says:

> “Although reason, according to this picture, is not literally an efficient cause of action, free actions are not regarded as uncaused. It is rather that the act of incorporation is conceived as the genuine causal factor and reason "has causality" only in the Pickwickian sense that it provides the guiding rule.” (1990: 51)

I agree to the extent that the timeless ground we must posit provides the legal link between the phenomenal occurrences that correspond to our timely bound decisions, it does not irrupt into appearances as an alien force – *that* would be incomprehensible[109]. Now, I would nevertheless go a step further than Allison. As Freyenhagen notes (2008), Kantian freedom cannot do with just “atemporal principles of reason”, in the sense that they hold generally as rules for rational action. After all, we can act under an evil or a good maxim; these are not completely general principles. In my reading, the atemporal aspect of freedom *really* grounds the causal link of actions, we have to posit a *real ground* in virtue of which alone our powers obtain, and depending on what this ground happens to be, we will produce different effects.

---

[109] “Now the action, insofar as it is to be attributed to the mode of thought as its cause, nevertheless does not follow from it in accord with empirical laws, i.e., in such a way that it is preceded by the conditions of pure reason, but only their effects in the appearance of inner sense precede it. Pure reason, as a merely intelligible faculty, is not subject to the form of time, and hence not subject to the conditions of the temporal sequence.” (KrV A 551/ B 579).

But can we really make sense of a timeless ground of temporal appearances? Saunders (2022) considers that possibility, but he observes some issues with it:

> "On the one hand, this might be a relatively weak claim: speculation about the mere possibility of positing some time-like ordering in order to overcome the difficulty of thinking of things-in-themselves being timeless. But if that is the case, then the proposal […] merely shifts the bump in the rug, and we now face the puzzling questions of how some other ordering can begin—or relate to—a temporal ordering? On the other hand, Walker's solution could involve a more substantive claim about an actual time-like ordering that things-in-themselves have, which seems to move us to second horn of the trilemma, where we accept that things-in-themselves could be in time, or have a time-like ordering." (2022: 282-3)

There is no discussion that the sort of relation that should hold between an atemporal ground and temporal phenomena is not precisely easy to conceptualize. Even more so if we are Kantians and believe that all knowledge, properly speaking, ultimately relies on intuition in time (KrV A 31/ B 46). If this matter is unsolvable, and the possibility that something that begins could do so in virtue of a ground that does not begin is inconceivable, then we should not just strike down Kant's theory of freedom, but transcendental idealism in general. How inconceivable is it, though?

There are many metaphysical relations in which a relatum features some novelty with regards to the other one (this, in fact, being what makes the relation interesting), and which we must take as primitive because of the role they play in our overall metaphysics. A prime example is, of course, causality itself: Kant himself (in a distinctively Humean move[110]) claims that we can have no notion of how something could be the ground for the existence of a whole different thing (KrV A 206-7/B 252), yet that is a constitutive principle of experience. More relevant for our purposes here, though, are contemporary truthmakers. In their classic form, a truthmaker is a *real* entity that has some bearing on the truth value of a proposition (Armstrong, 2004: 16)[111]. Even more so than between cause and effect, which are at least the same kind of ontological entity, the relation between the existence of something and the truth value of a proposition is highly mysterious. Nevertheless, it seems that we can make sense of it. Can we not make sense of a grounding relation between atemporal and temporal relata?

---

[110] *Enquiry*, 72.

[111] The analogy is relevant given the similarity in the role truthmakers and Kantian real grounds play (cf. Asay, 2023; NE 1: 391).

Let us shed some light on the question of timelessness using empirical laws as an example. As we have already seen, it seems that they do not come into existence when they are instantiated (§7, §15). If that is so, we can imagine an experiment. I put a snowball into a balloon, and I throw it for a mate to catch. It forms a parabola and my friend catches it at 09:00 AM. After that, he puts it in boiling water. At 09:01 AM the snow starts melting, and by 09:03 AM it is turning into steam. The expanding steam fills the balloon, which pops at 09:05 AM. This is a thoroughly temporal process. However, the idea that different laws of nature[112], which hold necessarily regardless of time, were instantiated at different points of this process, is not that outrageous a proposition. At 09:00 AM, the balloon forms a parabola in virtue of the laws of Newtonian motion; at 09:01, snow melts in virtue of the relevant thermodynamical laws; at 09:03, the balloon fills up in virtue of the law of ideal gasses. The specific phenomena that instantiate the laws are bound in time, that is for sure; but the laws held before they came into existence, and they will keep holding after they are gone – they are not, themselves, under conditions of time, even though they ground phenomena which are. If a free action is an action grounded by the right sort of intelligible ground, specific in scope as this *sui generis* ground may be, what further conceptual difficulty is there in holding that it grounds our powers for producing actions despite being atemporal itself?[113]

Finally, timelessness might not just be a metaphysical quirk needed for the overall apparatus of Kantian freedom to work. Recently, Korsgaard has used a notion of timelessness in order to respond to some concerns around her notion of moral standing the details of which need not preoccupy us at this moment. Timelessness may then be a conceptual tool with sharp practical consequences:

---

[112] I am using laws of nature instead of, for example, properties as the grounds of these phenomena because it appears to be the closest to Kant's own description. However, we can just as easily regard properties as atemporal (Rychter, 2008; Tugby, 2022: 44).

[113] Kant makes some enticing remarks in the *Inaugural Dissertation* regarding how he pictures the relation between the atemporal ground of reality and the temporal medium in which we experience it: "…the possibility of all changes and successions […] is to be found in the concept of time […] But that, of which the states flow, only endures if it is sustained by something else. And, thus, the concept of time, as the concept of something unique, infinite and immutable, in which all things are and in which all things endure, is the *phenomenal eternity* of the general *cause*." (ID 2: 410). If I am reading this correctly, the idea is that the eternal legal layout of noumenal nature is expressed in experience as the ceaselessness of time. What is a single, immutable ground of nature from the point of view of noumena, amid phenomena appears to us as the fact that every occurrence in our world shares one same overall and all-encompassing time as its medium (cf. KrV A 32/ B 47).

> "Since I believe that the subjects of lives and their moral standing are atemporal, I believe this: once you exist, once our life begins, you have a moral standing that is itself atemporal. That means in effect that you can be wronged by actions that take place either before or after your life." (2018: 89)

This is not itself an argument for or against Kant's need for atemporal grounds, but it shows that they *can* help ground some substantial ethical claims, rather than being a merely abstract and virtual cog in our metaphysical machinery. Thus, in conclusion, I find the objections around the conceptual untenability of timeless freedom unconvincing. It is no less understandable than other metaphysical relations that are not shrouded with as much mystery as Kant's theory of freedom, it is a key relation for transcendental idealism in general, and it can even have tangible, reasonable practical consequences.

§19. Whither go alternate possibilities?

A second concern that my proposal may fuel relates to the sort of freedom we end up with, if a free action is characterized by having a peculiar ground rather than anything else. There are two sides to this concern. First, one internal: part of my argument in favour of this image of freedom relied on the requirement for moral change espoused by Kant in the *Religion* (§14). Can I live up to my own claims and account for moral change? Second, one external: is this really what Kant meant by freedom? Are there alternate possibilities under Kant's view?

Let us start with the first point. It may seem that, by finding our intelligible ground in our experience of humans as organisms, we are locating freedom in an irredeemably static and unchanging essence, opposite from what freedom is about. Imagine, however, the following. You are a soldier at Agincourt. Among many other things, you are a physical being with mass, and as such you are pulled down by the force of gravity. Say a longbowman sees you and strikes you on the spot. It is still true that 'you' are a physical being with mass, pulled down by the force of gravity, but you have also lost a very good deal of powers: for instance, the capacity to breathe. You have stopped instantiating the ground that accounted for your powers as a living human being[114], but all throughout you have never stopped being grounded by the general nature that accounts for the universal laws of physics.

---

[114] This is why I have some doubts regarding Watkins's (2005) and Indregard's (2018) insistence that empirical grounds are immutable. However, I will grant them that they can account for this and not pursue any further discussion about this particular point.

Same as we, physical beings, can be living humans or dead humans depending on relevant circumstances, we can see how a living human being can either instantiate a good or a bad moral character. What we are authorized to judge is that we have the right sort of ground for free action in general, but this ground is broad enough that it can remain unchanged while our intelligible character changes: rational human beings *can* be either evil or good, both are within their possibilities as physical beings who nevertheless have morality as an incentive (KpV 5: 84; RGV 6:), as opposed to other animals or to holy beings. Remember that if we had an intuitive understanding, we could derive from our determination as physical beings that we are humans and whether we are going to be good or evil (KU 5: 406-7). Since we do not, we posit that there is some purposive matter of fact as to our moral character encompassed under our more general determinations. We, however, cannot know when or whether we have stopped instantiating an evil ground and started being grounded by a good fundamental deed – this is consistent *both* with Kant's general remarks on the intelligible ground of nature (L 9: 61, KU 5: 180), *and* with his notion of the opacity of our internal motives (GMS 4: 407; RGV 6: 38)[115].

Now, on to the second problem: does Kant need alternate possibilities, and does this reading allow them? It is a long-standing debate in philosophy of action whether the principle of alternate possibilities[116] is a necessary condition for freedom (Frankfurt, 1969; Mumford and Anjum, 2015). Similarly long-standing is the question whether Kant adhered to it or not. Interpreters abound both in the affirmative (Hudson, 1994: Insole, 2013: 81-2) and the negative camp (Pereboom, 2007: 542), the latter often maintaining that the only *ability* proper we have as free agents is that to do good (Timmermann, 2022; Ware, 2023), not to choose between good and evil. This is not just an interpretative quandary for modern readers, but a problem that has baffled philosophers ever since Kant published his first *Critique* (Gardner, 2017). So, what is it?

---

[115] Is it even possible that a human grounded by an evil character should transition into one with a good character? To understand the progressive character in appearance of this noumenal change of heart, Frierson (2003: 124) speaks of a "will in revolution", what I interpret as a good ground of action taking over in the context of the cumulative evils of life still present in the dispositions of that person. Timmermann (2022, ¶28) focuses on the status of the moral revolution through what I think is an extremely apt comparison to Kant's stance on the French Revolution: it is "a transfer of sovereignty [from inclination to reason] that backfires for those who initiate it" (2022: 85). A more recent example occurred in 1976, in Spain, after the death of dictator Franco. The parliament of the old regime used its still standing powers to pass the *Political Reform Act*, dissolving itself and calling for a constitutional process that should set up a new form of government. Similarly, be it unwillingly like Louis XVI or partially willingly like the Francoist *Cortes* (maybe overcome by the pains of an unabashedly selfish life, or fearful of consequences to come), an evil person may take a series of decisions, even from a bad motive, that eventually result in the downfall of their evil character, such that they stop instantiating their erstwhile twisted ground of actions. Unlike in politics, however, how this happens, and whether it has happened, we resolutely cannot know.

[116] Originally posed in relation to a puzzle with moral responsibility, it reads "a person is morally responsible for what he has done only if he could have done otherwise" (1969: 829). *Mutatis mutandis*, we can extend it to freedom of the will.

I am going to conveniently leave the debate over whether freedom is a two-way or a one-way ability aside, and focus on what immediately concerns my proposal, which was the classic matter of dispute during the *Freiheitsstreit*: can we have freedom if our actions are determined, even if what determines them is noumenal? It may be so that we would not want this for our own theory of freedom, but what did Kant think of it? If he regards that free action cannot be determined at all, then my reading has a problem. However, it seems that Kant did not think so:

> "There is no difficulty in reconciling the concept of *freedom* with the idea of God as a *necessary* being, for freedom does not consist in the contingency of an action (in its not being determined through any ground at all), i.e. not in indeterminism […] but in absolute spontaneity. The latter is at risk only with predeterminism, where the determining ground of an action lies *in antecedent time*" (RGV 6: 52n.)

Now, admittedly, this footnote refers to the possibility of divine freedom, which may reasonably have different requirements than human freedom (Insole, 2013: 81-2). However, the big problem around Kantian determinism, having left behind the issue of phenomenal predeterminism, is precisely whether we are not determined to act by the all-encompassing ground that is God (KpV 5: 101), and whether this matters. The matter then is quite analogous. If God's determination of our actions – which is, after all, the sort of intelligible legal grounding we have been talking about all along – is a form of predetermination, then there are reasons to think that such intelligible ground is deleterious for freedom. If it is not, it begs the question why it should matter now. Indeed, God exists outside of time, so there can be no such succession in the intelligible ground, and no predeterminism (KpV 5: 102)[117]. Insole, for instance, proposes that we regard God's creation as a noumenal *community* of transcendentally free beings, not subordinated to each other, but in reciprocal determination (Insole, 2013: 181-2)[118].

---

[117] "…it would turn out quite differently if the beings in the world as things in themselves existed in time, since the creator of substance would also be the author of the entire mechanism in this substance."

[118] Is this consistent with my reading of the intelligible ground of nature as a legal layout? We can see how the two pictures could coalesce through the notion of a world Kant presents in the *Inaugural Dissertation*: an intelligible world, according to the pre-critical Kant, is a whole of substances that interact in virtue of a common ground they share (ID 2: 407), namely, God as their creator (ID 2: 408). We can, then, have a community of intelligible substances behave in conformity with a divine legal layout. Adding noumenal substances into the picture does not necessarily trump or hurt my law-focused reading.

I will not intend to brush off this centuries-long issue in a couple of paragraphs[119]. There seems, however, to be reasonable grounds for holding that Kant regarded determination by the right sort of intelligible ground (i.e. divine creation) as ultimately unproblematic. Even if this fails in the end, the fact that this has been a problem Kant has carried with him since he published the first *Critique* must be a relief for my interpretation, rather than hampering its credibility.

Am I taking a stance, however, on the need for alternate possibilities by going down this interpretative route? Two brief comments are in order before answering. First, I admittedly sympathize with readings that do not emphasize alternate possibilities in Kant's theory of freedom; as I have argued throughout the last chapters, Kant's doctrine is much more insistent on the need for free actions to come from the right sort of ground than about their contingency or the possibility that they be otherwise. Pereboom's (2007) characterization of Kant as a proponent of source freedom instead of leeway freedom may, then, be apt. Second, while alternate possibilities are defended by many as a condition of free will, Frankfurt's position is certainly not without allies[120], so a clear ruling against a Kantian source freedom will not come from the ranks of the philosophy of action. Having said that, alternate possibilities are not closed off by my proposed approach to Kant's theory of freedom. Hudson (1994: 12) argues that there is a modal fallacy at play in claiming that, were free action grounded on a sufficient reason, then whatever course of action is taken will *necessarily* have been so, and thus there could be no alternate possibilities[121]. He thus argues for a compatibilist reading of Kant that, nevertheless, observes the principle of alternate possibilities. I have already expressed my disagreement with Hudson's overall approach (§6), but, if he can argue for alternate possibilities and compatibilism, my grounds-based proposal should not be much more problematic in that regard. Hence, I will not take a stance on this question: alternate possibilities are not precluded in principle by my reading.

§20. The epistemic boundaries of critical philosophy

I have made a rather cavalier use, one could argue, of noumenal grounds in my reading of Kant's theory of freedom. Where do I get the authorization to estipulate any such things, if we are dealing with the critical Kant, who claims that "without [empirical intuition, concepts] have no objective validity at all, but are rather a mere play, whether it be with representations of the imagination or of the understanding" (KrV A 239/ B 298)? Have I illicitly trespassed into the sphere of the transcendent, instead of embracing the unknowability of freedom?

---

[119] See Insole (2013) for a thorough analysis of this matter.

[120] Like the aforementioned Sartorio (2016).

[121] The fallacious argument is: (i) P; (ii) $\Box(P \rightarrow Q)$; (iii) $\vdash \Box Q$. For a critique of this approach to alternate possibilities, see O'Connor (2000).

Freedom, as an idea (GMS 4: 448; see KrV A 419/ B 447), has a job to do in Kant's philosophy. As I argued in chapter 1, one of the responsibilities under this job title is giving an account of how certain phenomenal actions can be said to be freely caused by agents. This is no easy task, as I have attempted to show in the past chapters. There is a whole host of requirements that an acceptable resolution to the problem should abide by, and these include the key insight of transcendental idealism: we must posit noumena as the ground of appearances, and only there, though in a very particular relation with phenomena, is there room for freedom. We can compare the development of these rational requirements of freedom to Kant's approach to theology:

> "The idea of a moral ruler of the world is a task for our practical reason. Our concern is not so much to know what he is in himself (his nature) but what he is for us as moral beings; even though for the sake of this relation we must think the divine nature by assuming it to have the full perfection required for the execution of his will (e.g. as the will of an immutable, omniscient, all-powerful, etc. being)." (RGV 6: 140)

God is beyond the boundaries of our possible knowledge, but the idea of God has an important role to play for us, as the condition of possibility for the highest good (KpV 5: 125). We cannot simply do, however, with a vacuous concept of God; we need such a concept that will be able to account for the possibility of the highest good. That is how Kant argues for divine omniscience, omnipotence, and so forth.[122] Similarly, a vacuous concept of freedom will not suffice, we need one that is capable of achieving, without contradiction, all that it is set up to do, and we can explore such a concept. I stand with Insole regarding this point:

> "if the spatio-temporal causally determined series that we experience are features of our reception and cognition of the world, rather than being in the world in itself, then it is possible at least that the way things are in themselves is fundamentally different from the way in which we experience them. Epistemic humility brings both security within

---

[122] "In Kant's view, the rational idea of God has a twofold interest for us: on the one hand, it is the focus of a natural and inevitable (though necessarily fruitless) theoretical inquiry; on the other, it is an object of the highest practical (moral) concern. Of course, there can be no doubt that for Kant the latter interest takes precedence […] Yet even Kant's moral theism is in a way dependent on the theoretical idea of God." (Wood, 1978: 25).

> its parameters, and a conceptual freedom about that which we do not know” (2013: 103).

As with God, we cannot have theoretical knowledge about freedom[123]. Freedom is only granted assertoric status practically, not theoretically[124] (KpV 5: 104-5). Prior to the warrant that our awareness of the moral law gives to us, the edifice of freedom is purely hypothetical in form. *If* there are free beings, then they will have the properties we have argued they must have. But once we have a practical mandate to posit the reality of free beings, *modus ponens* kicks in, and we gain assertoric access to the entire structure of free action. Kant makes this point quite clearly throughout his essay *How to orient oneself in thinking:*

> “A pure rational faith is therefore the signpost or compass by means of which the speculative thinker orients himself in his rational excursions into the field of supersensible objects; but a human being who has common but (morally) healthy reason can mark out his path, in both a theoretical and a practical respect, in a way which is fully in accord with the whole end of his vocation […] The *concept* of God and even the conviction of his *existence* can be met with only in reason” (DO 8: 142)

But practical assent is not the same as knowledge. It is “not inferior in degree”, but it is “completely different from it in kind” (DO 8: 141). Where does it differ? As we have been pointing out, practical assent does not grant any access to the internal possibility, and thus real determinations of its objects, which are completely beyond our access (Chignell, 2010). It lets us “think of something supersensible in a way which is serviceable to the experiential use of our reason”, by determining the relation of the transcendental object we posit to the phenomena we can experience (DO 8: 136). This accounts for the opacity that shrouds Kant’s use of these objects, always prefaced by a “such that” clause. We cannot know the legal grounds of nature, we can only posit that they relate in a certain way to appearances; we cannot know the real character behind our actions, we can only posit that we have such a character, and that, acting through it, we are free. This is as far as transcendental idealism will allow us to go, but it is far enough for my model. We have a practical warrant to posit a model of freedom such that it works in regarding phenomenal actions to be free, but this warrant only authorizes us to posit that humans can be free, it gives us no privileged knowledge into the hearts of people – not even our own.

---

[123] For a thorough analysis of the modes of holding-as-true (*Fürwahrhalten*) in Kant’s epistemology, see Chignell (2007).

[124] From the *Critique of Practical Reason* onwards, at least. See footnote 3.

6

CONCLUSION

I have argued, in the past chapters, for a novel approach to understanding Kant's theory of freedom. Kant requires that freedom be an absolute causal beginning for actions, and one grounded on rational principles. Actions amid appearances can be called free if they obtain in virtue of an intelligible, noumenal ground, which pertains to their empirically unified authors. In order for us to legitimately posit these grounds, we must take into account the requirements of our power of judgment, which draws us into assuming that there is a purposive ground of the necessity of empirical laws, and one which must account for the possibility of organisms. As rational organisms, we are the right sort of empirical being that can have the ground needed for free agency. We can never properly access these grounds, but we are entitled to assume that the appropriate sort of substrate is there.

Kant's transcendental idealism opens the door to letting what appear to be two contradictory images of the world live together in our single model of reality. By doing this, however, we risk creating a disjointed universe, with two unconnected realms ruled by fundamentally incompatible laws. Bringing the domains of theoretical and practical philosophy back to their common ground was the explicit goal of the *Critique of Judgment*. It has been the aim of this dissertation to show that this seminal work does not exhaust itself at giving an answer to the possibility of completion of our ultimate moral goals. The *Critique of Judgment* shows us the way forward in making sense of how we judge particular empirical beings as expressive of an intelligible ground; it gives us the tools to understand how freedom can irrupt amid nature without breaching its unity. Insofar as we need to judge ourselves to be real members of the substrate of nature, we regard freedom as a latent ground of appearances in the world. Human freedom, however, being the freedom of imperfect beings, exhibits a dual character. In acting in a way imputable to us, we can either remain obstinate in the dominion of pathological impulses over nature or instantiate a new sort of order in the world: that in which the moral law determines, not only what ought to be, but also what actually is. Making sense of this possibility, I believe, and I have argued, is what lies at the heart of Kant's theory of freedom.

BIBLIOGRAPHY

Works by Kant are cited according to the volumes and page numbers of the *Akademie-Ausgabe*, except for the *Critique of Pure Reason*, which is cited following the A- and B-editions. Translations are from the *Cambridge Edition of the Works of Immanuel Kant*. Abbreviations are as follows:

Coll. *Collins' Lectures on Ethics.*

DO: *Was heißt, sich im Denken orientieren?* [*What Does it Mean to Orient Oneself in Thinking?*]

FM: *Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnitzens und Wolfs Zeiten in Deutschland gemacht hat?* [*What real progress has metaphysics made in Germany since the time of Leibniz and Wolff?*]

GMS: *Grundlegung zur Metaphysik der Sitten.* [*Groundwork for the Metaphysics of Morals*]

ID: *De mundi sensibilis atque intelligibilis forma et principiis.* [*On the form and principles of the sensible and the intelligible world (Inaugural Dissertation)*].

KpV: *Kritik der praktischen Vernunft*. [*Critique of Practical Reason*].

KrV: *Kritik der reinen Vernunft*. [*Critique of Pure Reason*].

KU: *Kritik der Urteilskraft*. [*Critique of the Power of Judgment*].

L: *Logik. Ein Handbuch zur Vorlesungen.* [*Jäsche Logic*].

MS: *Metaphysik der Sitten*. [*Metaphysics of Morals*].

NE: *Principiorum primorum cognitionis metaphysicae nova dilucidation.* [*New Elucidation of the first principles of metaphysical cognition*].

RGV: *Die Religion innerhalb der Grenzen der bloßen Vernunft*. [*Religion within the Boundaries of Mere Reason*].

References

Allison, Henry E. (1982). "Practical and Transcendental Freedom in the Critique of Pure Reason", *Kant Studien* 73 (1-4): 271-290.

Allison, Henry E. (1990). *Kant's Theory of Freedom*. CUP.

Allison, Henry E. (1996). *Idealism and freedom. Essays on Kant's theoretical and practical philosophy*. CUP.

Allison, Henry E. (2020). *Kant's Conception of Freedom*. *A Developmental and Critical Analysis*. CUP.

Armstrong, D. M. (2004). *Truth and Truthmakers.* CUP.

Asay, Jamin (2023). *Truthmaking.* Cambridge Elements, CUP.

Beck, Lewis (1963). *A Commentary on Kant's Critique of Practical Reason*. University of Chicago Press.

Beck, Lewis (1975). *The Actor and the Spectator*. Yale University Press.

Bennett, Jonathan (1984). "Commentary: Kant's theory of freedom", in *Self and Nature in Kant's Philosophy*, Allen W. Wood (ed.). Cornell University Press.

Berlin, Isaiah (1969). "Two Concepts of Liberty", in *Four Essays on Liberty*. OUP.

Breitenbach, Angela (2014). "Biological Purposiveness and Analogical Reflection", in *Kant's Theory of Biology*, Ina Goy and Eric Watkins (eds.). De Gruyter.

Chignell, Andrew (2007). "Kant on belief", *The Philosophical Review* 116(3): 323-360.

Chignell, Andrew (2010). "Real Repugnance and our Ignorance of Things-in-Themselves: A Lockean Problem in Kant and Hegel", in Stolzenberg, J. and Rush, F. (eds.), *Internationales Jahrbuch des Deutschen Idealismus: Glaube und Vernunft*. De Gruyter.

Cohen, Alix A. (2009a). "Kant's Concept of Freedom and the Human Sciences", *Canadian Journal of Philosophy,* 39(1): 113-35.

Cohen, Alix A. (2009b) *Kant and the Human Sciences: Biology, Anthropology and History.* New York: Palgrave-Macmillan.

Davidson, Donald (1970). "Mental Events", in *Essays on Actions and Events*. OUP, 1985.

Esser, Andrea (2008). "Kant on Solving Moral Conflicts", in Betzler, M. (ed.), *Kant's Ethics of Virtue*.

Fodor, Jerry (1990). *A Theory of Content and Other Essays*. MIT Press.

Frankfurt, Harry G. (1969). "Alternate Possibilities and Moral Responsibility", *The Journal of Philosophy* 66(23: 829-39.

Freyenhagen, Fabian (2008). "Reasoning Takes Time: On Allison and the Timelessness of the Intelligible Self", *Kantian Review* 13(2): 67-84.

Freyenhagen, Fabian (2020). "Acting Irrespective of Hope", *Kantian Review* 25(4): 605-30.

Friedman, Michael (2014). "Laws of Nature and Causal Necessity", *Kant-Studien* 105(4): 531-53.

Frierson, Patrick (2003). *Freedom and Anthropology in Kant's Moral Philosophy*. CUP.

Frierson, Patrick (2005). "Kant's Empirical Account of Human Action", *Philosophers' Imprint* 5(7).

Frierson, Patrick (2010). “Two Standpoints and the Problem of Moral Anthropology”, in *Kant’s Moral Metaphysics: God, Freedom, and Immortality*, Benjamin B. Lipscomb and James Krueger (eds.). De Gruyter.

Frierson, Patrick (2014). *Kant’s Empirical Psychology*. OUP.

Galvin, Richard (2019). “Freedom and the Fact of Reason”, *Kantian Review* 24 (1): 27-51.

Gambarotto, Andrea and Nahas, Auguste (2022). “Teleology and the organism: Kant’s controversial legacy for contemporary biology”, *Studies in History and Philosophy of Science* 93: 47-56.

Gardner, Sebastian (2015). “German Idealism, Classical Pragmatism, and Kant’s Third *Critique*”, in Gava, G. and Stern, R. (eds.), *Pragmatism, Kant and Transcendental Philosophy*. Routledge.

Gardner, Sebastian (2017). “The metaphysics of human freedom: from Kant’s transcendental idealism to Schelling’s *Freiheitsschrift*”, *British Journal for the History of Philosophy* 25(1): 133-56.

Geiger, Ido (2022). *Kant and the Claims of the Empirical World. A Transcendental Reading of the Critique of Pure Reason.* CUP.

Ginsborg, Hannah. “Two kinds of mechanical inexplicability in Kant and Aristotle”, *Journal of the History of Philosophy* 42(1): 33-65.

Grenberg, Jeanine (2010). “In Search of the Phenomenal Face of Freedom”, in *Kant’s Moral Metaphysics: God, Freedom, and Immortality*, Benjamin B. Lipscomb and James Krueger (eds.). De Gruyter.

Guyer, Paul (1990a). “Reason and Reflective Judgment: Kant on the Significance of Systematicity”, *Nous* 24: 17-43.

Guyer, Paul (1990b). “Kant’s Conception of Empirical Law”, *Proceedings of the Aristotelian Society,* suppl. (64): 221-42.

Guyer, Paul (2016). *Virtues of Freedom: Selected Essays on Kant*. OUP.

Hanna, Robert and Moore, A. W. (2007). “Reason, freedom and Kant: An exchange”, *Kantian Review* 12(1): 113-33.

Heidegger, Martin (1929). *Kant und das Problem der Metaphysik*. Frankfurt: Verlag Klostermann.

Hudson, Hud (1994). *Kant’s Compatibilism*. Cornell University Press.

Hutton, James (2021). “Kant, causation and laws of nature”, *Studies in History and Philosophy of Science Part A.* 86: 93-102.

Illetterati, Luca (2014). “Teleological Judgment: Between Technique and Nature”, in *Kant’s Theory of Biology,* Ina Goy and Eric Watkins (eds.). De Gruyter.

Indregard, Jonas Jervell (2018). “A gradual reformation: empirical character and causal powers in Kant”, *Canadian Journal of Philosophy* 48(5): 662-683.

Juárregui, Claudia (2020). “Kant and the regulative character of the principle of mechanism in the antinomy of the teleological power of judgment”, *Con-textos Kantianos* 11: 92-109.

Kitcher, Patricia (1982). “Kant on self-identity”, *Philosophical Review* 91(1): 41-72.

Kitcher, Patricia (1984). “Kant’s real self”, in *Self and Nature in Kant’s Philosophy,* Allen W. Wood (ed.). Cornell University Press.

Kohl, Markus (2014). “Transcendental and Practical Freedom in the *Critique of Pure Reason*”, *Kant Studien* 105 (3): 313-335.

Korsgaard, Christine (1996). *Creating the Kingdom of Ends*. CUP.

Korsgaard, Christine (2008). *The Constitution of Agency: Essays on Practical and Moral Psychology*. OUP.

Korsgaard, Christine (2009). *Self-Constitution. Agency, Identity and Integrity.* OUP.

Korsgaard, Christine (2018). *Fellow creatures. Our obligations to the other animals*. OUP.

Kreines, James (2012). "Kant and Hegel on Teleology and Life from the Perspective of Debates about Free Will", in Khurana, T. (ed.), *The Freedom of Life*: *Hegelian Perspectives*. Goethe-Universtiät Frankfurt.

Kreines, James (2017). "Kant on the Laws of Nature: Restrictive Inflationism and Its Philosophical Advantages", *The Monist* 100(3): 326-41.

Kryluk, Michael (2017). "Gallows Pole: Is Kant's Fact of Reason a Transcendental Argument?", *The Review of Metaphysics* 70 (4): 695-725.

Longuenesse, Béatrice (2007). "Kant on the Identity of Persons", *Proceedings of the Aristotelian Society* 107(1) pt 2: 149-67.

Marshall, Colin (2013). "Kant's Appearances and Things in Themselves as Qua-Objects", *The Philosophical Quarterly* 63(252): 520-45.

Massimi, Michela (2017). "Grounds, Modality, and Nomic Necessity in the Critical Kant", in Massima, M. and Breitenbach, A. (eds.), *Kant and the Laws of Nature*. CUP.

McCarty, Richard (2009). *Kant's Theory of Action*. OUP.

McLaughlin, Peter (2003). "Newtonian Biology and Kant's Mechanistic Concept of Causality", in Guyer (ed.), *Kant's Critique of the power of judgment. Critical Essays*. Oxford: Rowman and Littlefield Publishers.

McLaughlin, Peter (2014). “Mechanical Explanation in the ‘Critique of the Power of Judgment’”, in *Kant’s Theory of Biology*, Ina Goy and Eric Watkins (eds.). De Gruyter.

Meerbote, Ralf (1984). “Commentary: Kant on Freedom and the Rational and Morally Good Will”, in *Self and Nature in Kant’s Philosophy*, Allen W. Wood (ed.). Cornell University Press.

Messina, James (2017). “Kant’s Necessitation Account of Laws and the Nature of Natures”, in Massima, M. and Breitenbach, A. (eds.), *Kant and the Laws of Nature*. CUP.

Michalson, Gordon E. (1990). *Fallen freedom. Kant on radical evil and moral regeneration.* CUP.

Mill, John Stuart (2015 [1859]). “On Liberty”, in *On Liberty and Other Essays*, Mark Philip and Frederick Rosen (eds.). OUP.

Mumford, Stephen and Anjum, Rani Lill (2015). “Freedom and Control: On the Modality of Free Will”, *American Philosophical Quarterly* 52(1): 1-11.

Munzel, G. Felicitas (1999). *Kant’s Conception of Moral Character. The “Critical” Link of Morality, Anthropology and Reflective Judgment.* Chicago University Press.

Nelkin, Dana K. (2000). “Two Standpoints and the Belief in Freedom”, *The Journal of Philosophy* 97(10): 564-76.

O’Connor, Timothy (2000). *Persons and Causes. The Metaphysics of Free Will.* OUP.

O’Neill, Onora (1989). *Constructions of Reason. Explorations of Kant’s practical philosophy.* CUP.

O’Shea, James (1997). “The Needs of Understanding: Kant on Empirical Laws and Regulative Ideas”, *International Journal of Philosophical Studies* 5(2): 216-54.

O'Shea, James (2013). "Kant and Dennett on the epistemic status of teleological principles", in O'Rourke, F. (ed.), *Human Destinies: Philosophical Essays in Memory of Gerald Hanratty*. University of Notre Dame Press.

Palatnik, Nataliya (2022). "Kantian freedom at a distance", *British Journal for the History of Philosophy* 30(6): 1033-54.

Pereboom, Derk (2007). "Kant on Transcendental Freedom", *Philosophy and Phenomenological Research* 73(3): 537-67.

Quarfood, Marcel (2014). "The Antinomy of Teleological Judgment: What It Is and How It Is Solved", in *Kant's Theory of Biology,* Ina Goy and Eric Watkins (eds.). De Gruyter.

Rawls, John (2000). *Lectures on the History of Moral Philosophy*, Harvard University Press.

Rodríguez-González Barredo, Álvaro (2023). "Can Kant tell us what to do? Intentions, dilemmas, and the long journey towards moral perfection", *Revista de Estudios Kantianos* 8(1): 23-35.

Rychter, Pablo (2008). "Perdurance, Endurance, and 'Having a Property *Atemporally'*, *Metaphysica* 9: 159-171.

Sartorio, Carolina (2016). *Causation & Free Will.* OUP.

Saunders, Joe (2016). "Kant and the Problem of Recognition: Freedom, Transcendental Idealism, and the Third-Person", *International Journal of Philosophical Studies* 24(2): 164-82.

Saunders, Joe (2021). "Some Hope for Kant's *Groundwork III*", *Inquiry*.

Saunders, Joe (2022). "Timeless Freedom in Kant: Transcendental Freedom and Things-in-Themselves", *History of Philosophy Quarterly* 39(3): 275-292.

Saunders, Joe. “Korsgaard’s Two Standpoints”. Unpublished manuscript.

Sethi, Janum (2023). “Watkins on Kant’s Laws of Nature”, *Kantian Review* 26(2): 307-14.

Stern, Robert (2000). *Transcendental arguments and scepticism: answering the question of justification*. OUP.

Steward, Helen (2012). *A Metaphysics for Freedom*. OUP.

Strawson, Peter F. (1966). *The Bounds of Sense. An Essay on Kant’s Critique of Pure Reason*. Routledge.

Sticker, Martin and Saunders, Joe (2022). “Why we go wrong: beyond Kant’s dichotomy between duty and self-love”, *Inquiry*.

Stroppa, Lucca (2023). “A dilemma for welfare pluralists”. Conference given at *Understanding Value XI*, U. of Sheffield.

Thompson, Michael (2012). *Life and Action. Elementary Structures of Practice and Practical Thoughts*. Harvard University Press.

Timmermann, Jens (2022). *Kant’s Will at the Crossroads. An Essay on the Failings of Practical Rationality*. OUP.

Tugby, Matthew (2022). *Putting Properties First: A Platonic Metaphysics for Natural Modality*. OUP.

Ware, Owen (2014). “Rethinking Kant’s Fact of Reason”, *Philosophers’ Imprint* 14.

Ware, Owen (2023). *Kant on Freedom*. Cambridge Elements, CUP.

Watkins, Eric (2005). *Kant and the Metaphysics of Causality*. CUP.

Williams, Neil (2019). *The Powers Metaphysic*. OUP.

Wood, Allen W. (1978). *Kant's Rational Theology*. Cornell University Press.

Wood, Allen W. (1984). "Kant's Compatibilism", in *Self and Nature in Kant's Philosophy*, Allen W. Wood (ed.). Cornell University Press.

Zuckert, Rachel (2007). *Kant on Beauty and Biology. An Interpretation of the 'Critique of Judgment'*. CUP.