Reason, Freedom and Morality:

An Interpretation and Defence of Kant’s *Groundwork* III

Josiah Paul Saunders

Thesis submitted for the degree of: Doctor of Philosophy

March 2015

Department of Philosophy,

University of Sheffield

Abstract

Kant worries that if we are not free, morality will be nothing more than a phantasm for us. In the final section of the *Groundwork*, he attempts secure our freedom, and with it, morality. Here is a simplified version of his argument:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

In this thesis, I offer an interpretation and defence of this.

I begin by defending the first two premises. I follow Kant to argue that reason involves freedom, and offer an account of the relationship between freedom and morality.

I then turn to two prominent objections. Commentators often complain that Kant has not managed to establish that we are rational beings with willsin the first place, and that he equivocates in his use of ‘free’ between premise 1 and 2. I argue that both of these objections can be overcome, and thus seek to defend Kant’s approach in *Groundwork* III.

In doing so, I depart from Kant (and Kantians) at several points. Most significantly, I argue against a non-metaphysical account of freedom, an anti-realist meta-ethics, and transcendental idealism itself. I stay with the spirit of Kant’s project, but often depart from the letter. In this, I find my project to be post-Kantian; I begin with Kant, but end up alongside his successors. Through this, I hope to vindicate our conception of ourselves as free, such that morality is no phantasm.

Acknowledgements

I want to begin by thanking Philip Catton for introducing me to Kant, and to him, Paul Studtmann, Carolyn Mason, Sagar Sanyal and Celia Sheerin for suffering through years of me discussing Kant in New Zealand.

Speaking of suffering, I must thank Robert Stern for reading and listening to me struggle through *Groundwork* III. I have been lucky enough to have an on-going conversation with Bob about reason, freedom and morality for 4 years now, and this PhD – and my thought in general – would be *much* worse without him.

I also owe a big thanks to Chris Bennett. Chris has an incredible eye for philosophy, which is a shame, because it means he knows exactly what I am up to, and exactly where I go wrong. Again, this thesis would be *much* worse without him.

Sheffield is home to a great graduate community, who have all had to listen to me talk about *Groundwork* III. For this, I would like to apologise. I owe special apologies to Charlotte Alderwick, who patiently spent 3 years trying to convince me that transcendental idealism did not provide a satisfying account of freedom; I did not listen at all, but it turns out she was right. I owe her and Neil Williams a massive apology for asking them to read through a draft of this thesis. I also owe apologies to Carl Fox and Jack Wadham, for asking them to read through drafts of various chapters.

I owe big apologies to both Martin Sticker and Irina Schumski. They have both read countless papers of mine, and always provide me with excellent criticisms, sometimes even related to the papers I send them!

I owe some institutions apologies as well, primarily the University of St. Andrews (and the Kingdom of Jens) for putting up with me criticising Kant every year, and the University of Pittsburgh (and in particular Stephen Engstrom and Tyke Nunez), for listening to me criticise Kant’s formalism over and over again. I also presented parts of this thesis at King’s College London, the Central European University in Budapest, University and Trinity College Dublin, the University of Konstanz, Bogazici University in Istanbul, and the University of Groningen, and would like to thank them all for listening. I should also thank the *Arts and Humanities Research Council* for funding this PhD, and my visiting-studentship in Pittsburgh.

I would like to apologise to my friends and family. Anna, John, MC, Donna, Mum, Dad, Dan, Nicole, Pete, Tash, Charlie, Joe, Max, Jess, Fabien, Ciaran, Nick, Tom, Rob, Daniel, Steve, Josh, Rich, and Tjalling, you have all had to listen to me talk about Kant, reason, freedom and morality for quite some time now. I can only promise that it will continue.

And last, but no by no means least, I would like to thank Kant. I find it easy to forget that Kant was human. It’s actually more than that: I find it hard to remember that he was human. But he was. He was a single human being writing just over two hundred years ago, trying to secure our freedom, and with it, morality. And for that, I want to thank him.

Contents

[**Introduction** 9](#_Toc414625208)

[1. The Identification and Corroboration of the Supreme Principle of Morality 10](#_Toc414625209)

[1.1 Identifying the Supreme Principle of Morality 11](#_Toc414625210)

[1.2 Corroborating the Supreme Principle of Morality 12](#_Toc414625211)

[`.3 Kant’s Deduction of Freedom 14](#_Toc414625212)

[2. Two Types of Argument for Freedom in Kant 18](#_Toc414625213)

[3. What *Groundwork* III is not 22](#_Toc414625214)

[4. Some Methodological Remarks 23](#_Toc414625215)

[5. Overview 25](#_Toc414625216)

[**1 – Reason, Spontaneity and Agency** 29](#_Toc414625222)

[1. Freedom must be Presupposed as a Property of the Will of all Rational Beings 30](#_Toc414625223)

[1.1 We must act under the Idea of Freedom 31](#_Toc414625224)

[1.2 The Inescapability of Freedom 32](#_Toc414625225)

[2. Reason’s Capacity for Ideas and the Intelligible World 36](#_Toc414625226)

[3. Reason and Spontaneity 41](#_Toc414625227)

[3.1 The Spontaneity Thesis 42](#_Toc414625228)

[3.2 Practical Reason 52](#_Toc414625229)

[3.3 The Metaphysics of Spontaneity 55](#_Toc414625230)

[Conclusion 58](#_Toc414625231)

[**2 – Freedom, Autonomy and Morality** 61](#_Toc414625232)

[1. Freedom, Autonomy and Morality: The Ambitious Account 62](#_Toc414625233)

[1.1 Allison’s Reciprocity Thesis 65](#_Toc414625234)

[1.2 Korsgaard’s Regress 66](#_Toc414625235)

[1.3 Ambiguities of Autonomy 68](#_Toc414625236)

[1.4 Kant’s Meta-Ethical Formalism 72](#_Toc414625237)

[1.5 Against Kant’s Formalism and Kantian Constructivism 80](#_Toc414625238)

[2. Freedom and Morality 96](#_Toc414625243)

[2.1 A Free Will stands under the Moral Law 96](#_Toc414625244)

[2.2 Freedom as the *ratio essendi* of Morality 103](#_Toc414625245)

[Conclusion 111](#_Toc414625246)

[**3 – Practical Reason** 113](#_Toc414625247)

[1. Evidence for our Agency 115](#_Toc414625248)

[1.1 Kant’s Pre-Critical Argument for Freedom 116](#_Toc414625249)

[1.2 Grounds for Doubt 120](#_Toc414625250)

[1.3 The Phenomenology of Reason 125](#_Toc414625251)

[2. The Possibility of Freedom 129](#_Toc414625252)

[2.1 Transcendental Idealism 129](#_Toc414625253)

[2.2 Against Transcendental Idealism 134](#_Toc414625254)

[2.3 Nature and Freedom 151](#_Toc414625262)

[Conclusion 154](#_Toc414625263)

[**4 – Rational Agency and the Moral Law** 155](#_Toc414625264)

[1. A Gap 155](#_Toc414625265)

[2. Reason and Motivation 158](#_Toc414625266)

[**Conclusion** 163](#_Toc414625267)

[**Bibliography** 165](#_Toc414625268)

Introduction

Kant’s *Groundwork of the Metaphysics of Morals* is one of the most important works of philosophy. And its task is relatively clear. In his own words, it is:

[…] nothing more than the identification and corroboration *of the supreme principle of morality*. (IV: 392. 3-4)

In the first two sections of the *Groundwork*, Kant identifies and analyses the supreme principle of morality. This is part of the canon of western philosophy. His attempt to corroborate it is not. For this, we have to turn to the third and final section of the book – *Groundwork* III.

Paul Guyer sums up a common attitude towards the *Groundwork* when he writes the following:

[Kant’s] analysis of the moral law and its several formulations is one of the masterpieces of western philosophy, and his attempt to provide a transcendental deduction of it is one of its most spectacular train wrecks. (Guyer 2007: 445)

Guyer is not alone in this diagnosis of *Groundwork* III. Karl Ameriks thinks it suffers “shipwreck” (2000: 191), while Henry Allison claims that “there is virtual unanimity that the attempt fails” (1990: 214).[[1]](#footnote-1) Indeed, even Kant himself appears to abandon this approach only a few years later in the *Critique of Practical Reason*. So much for *Groundwork* III – the only question that seems to remain concerns which vehicular metaphor captures the catastrophe best.

I look to provide some hope for *Groundwork* III, and in particular Kant’s argument for freedom in it. Kant appears to argue from our reason to freedom, and then from our freedom to morality. In this thesis, I offer an interpretation and defence of this argument. Here is a simplified version of it:[[2]](#footnote-2)

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

Over the course of this thesis, I set out and defend this argument.

This falls into four chapters. The first two chapters set up the first two premises. I argue that reason involves spontaneity, and offer an account of the relationship between freedom and morality. The final two chapters address two major objections. The first concerns whether our reason is practical, that is, whether we have wills. The second considers whether the freedom involved in reason is the same as the freedom involved in morality. I argue that these objections can be overcome, and that *Groundwork* III is not such a wreck after all.

That is a brief sketch of things to come. As for this introduction, I want to outline what to expect – and what not to expect – from Kant’s argument for freedom, and *Groundwork* III in general. I begin by discussing what it means to *corroborate* the supreme principle of morality (§1). I then turn to look at two different types of argument that Kant gives for freedom, one moral and the other non-moral. I argue that *Groundwork* III primarily attempts a non-moral argument for our freedom (§2), before clarifying precisely what a non-moral argument for freedom amounts to (§3). Here, I distance myself from those who read *Groundwork* III as a proof or justification of the moral law. I then offer some remarks on my methodology (§4), before ending with an overview of the thesis itself (§5).

# The Identification and Corroboration of the Supreme Principle of Morality

I began by quoting Kant’s own description of the *Groundwork*. Here it is again, in a little more detail. In the preface, Kant tells us that:

The present groundwork, however, is nothing more than the identification [*Aufsuchung*] and corroboration [*Festsetzung*] of the supreme principle of morality. (IV: 392. 3-4)[[3]](#footnote-3)

Let me say something about the identification, and the supreme principle of morality itself.

## 1.1 Identifying the Supreme Principle of Morality

Kant begins the *Groundwork* by analysing common human reason, and its conception of duty. Through this, he uncovers the following principle:[[4]](#footnote-4)

act only according to that maxim through which you can at the same time will that it become a universal law. (IV: 421. 7-8)[[5]](#footnote-5)

This is Kant’s supreme principle of morality. There are several different formulations of it,[[6]](#footnote-6) and the difficult question of their supposed equivalence.[[7]](#footnote-7) But we can circumvent these complications for now.[[8]](#footnote-8) For present purposes, what matters is that the supreme principle of morality is a categorical imperative.

There are two basic features of a categorical imperative: it is categorical, and an imperative. Beginning with the first of these, Kant conceives of morality as categorical, or unconditional. This emerges from his analysis of common human reason in *Groundwork* I and II, and concerns the normative force of moral claims. They are normative independent of our desires.[[9]](#footnote-9) Unlike hypothetical imperatives that derive their normative force from desires (if you want x, you ought to φ), Kant sees morality as commanding unconditionally (regardless of whether or not you want x, you ought to φ).[[10]](#footnote-10)

The categorical imperative is also an imperative. And there are two elements to this. The first is relatively straightforward: an imperative is a normative claim – one *ought* to φ. The second concerns Kant’s specific use of the term ‘imperative’. Kant distinguishes between the moral law and the categorical imperative: the moral law is the supreme principle of morality, and the categorical *imperative* is the particular binding form that this principle takes for creatures like us.[[11]](#footnote-11)

Kant views us as sensibly affected rational beings. We are rational, and can thereby easily recognise the moral law.[[12]](#footnote-12) However, we are also sensibly affected, and thus feel the pull of desires and inclinations. In less Kantian terms, we are rational animals. If we were fully rational and not sensibly affected, the moral law would not appear as an imperative to us, as we would have no desires that stood in the way of following it. Kant details this as follows:

Thus a perfectly good will would just as much stand under objective laws (of the good), but it could not be represented as thereby necessitated to actions that conform with laws, because it can of itself, according to its subjective constitution, be determined only by the representation of the good. Therefore no imperatives hold for the *divine* will and generally for a *holy* will: here the *ought* is out of place, because *willing* already of itself necessarily agrees with the law. Therefore imperatives are only formulae to express the relation of objective laws of willing as such to the subjective imperfection of the will of this or that rational being, e.g. of the human will. (IV: 414. 1-11)

We are sensibly affected rational beings, and it is because of this that the moral law appears to us as an imperative.[[13]](#footnote-13)

These are two key features of the supreme principle of morality. It is a categorical imperative, both unconditionally normative and an imperative (for us).

## 1.2 Corroborating the Supreme Principle of Morality

We now face the difficult question of what it means to *corroborate* the supreme principle of morality. The first two sections of the *Groundwork* leave something open. Kant analyses morality and identifies its supreme principle, but always postpones some heavy lifting for the third and final section. Here is a characteristic passage:[[14]](#footnote-14)

[F]or, how such an absolute command is possible, even if we know how to express it, will still require particular and arduous effort, which we suspend, however, until the final section. (IV: 420. 21-3)

This “particular and arduous effort” (IV: 420. 22-3) is ours in this thesis.

Here is the task. In *Groundwork* III, Kant attempts a deduction of the supreme principle of morality. It is important to clarify Kant’s use of the term ‘deduction’. He begins the transcendental deduction in the first *Critique* as follows:

Jurists, when they speak of entitlements and claims, distinguish in a legal manner between the questions about what is lawful (*quid juris*) and that which concerns the fact (*quid facti*), and since they demand proof of both, they call the first, that which is to establish the entitlement or the legal claim, the **deduction**. (A 84/B 116)

Here, we see that a ‘deduction’ has a particular meaning for Kant. As Dieter Henrich has famously shown, it is a now out-dated legal term, which concerns the securing of an entitlement to something that we take ourselves to possess.[[15]](#footnote-15) In the case at hand, we take ourselves to be subject to the supreme principle of morality, and in *Groundwork* III, Kant attempts to entitle us to this.

The question remains as to what exactly it is about the supreme principle of morality that Kant is attempting a deduction of. He uses the term ‘deduction’ three times in *Groundwork* III. In the first subsection, he writes of a “deduction of the concept of freedom” (IV: 447. 22-3), which he has yet to “make comprehensible” (IV. 447. 22). In the fourth subsection, he shows how a categorical imperative, qua imperative, is possible, and then writes of “the correctness of this deduction” (IV: 454. 20-1). And finally, at the end of the book, he refers back to “our deduction of the supreme principle of morality” (IV: 463. 21-2). The first two references set out two distinct tasks.[[16]](#footnote-16) Kant seeks to establish: 1) that we are free, and in particular, sufficiently free to stand under the moral law; 2) how the categorical imperative, *qua imperative*, is possible. These are the two major tasks of *Groundwork* III.[[17]](#footnote-17)

The recent literature is divided on the respective importance of these tasks. For a while, freedom was seen as the main issue in *Groundwork* III, and accordingly the first two or three subsections received the most attention.[[18]](#footnote-18) Recently however, scholars have recognised the importance of the imperatival form of the categorical imperative, and attention has shifted to the fourth subsection of *Groundwork* III.[[19]](#footnote-19)

These are both important issues.[[20]](#footnote-20) However, in this thesis, I confine my attention to freedom. The imperatival form that the categorical imperative takes for us is clearly important, but it has been dealt with excellently in the recent literature.[[21]](#footnote-21) Kant’s deduction of freedom on the other hand is not in such good shape. As noted at the outset, even Kant himself seems to abandon this strategy only a few years later.

## `.3 Kant’s Deduction of Freedom

Getting Kant’s deduction of freedom into focus is not easy. In general, *Groundwork* III is notoriously obscure. It is: “one of the most enigmatic of the Kantian texts” (Allison 1990: 214); “often […] described as rather dark or elusive and hardly accessible to meaningful reconstruction” (Horn and Schӧnecker 2006: viii); “a riddle wrapped within an enigma – and all this covered by a shroud for good measure” (Ameriks 2003: 226). Making sense of it is not easy. Henrich captures the nature of the text well when he writes that:

It is characteristic of the peculiar difficulty of the *Groundwork*, however, that even with a complete clarification of the arguments actually given by Kant the text is still not fully transparent. (Henrich 1975: 322)

No one seems sure exactly what or how Kant is arguing. The only agreement seems to be that, whatever it is, it fails.

One thing is relatively clear: Kant intended a deduction of freedom.[[22]](#footnote-22) How this proceeds is less clear.[[23]](#footnote-23) I find four different arguments for freedom in *Groundwork* III. In the order they appear in the text, they are:

1. The claim that freedom must be presupposed as a property of the will of all rational beings (IV: 447. 26 – 448. 22).
2. An appeal to transcendental idealism and the activity involved in (discursive) cognition (IV: 451. 1-36).
3. The claim that reason’s capacity for ideas reveals a “spontaneity so pure” (IV: 452. 18) that it thereby takes us far beyond anything sensibility can afford us, and marks us out as members of an intelligible world (IV: 452. 6-22).
4. An appeal to the consciousness of the causality of our reason (IV: 457. 4-7, IV: 457. 22-4, IV 459. 9-14, IV: 461. 17-25).

I think that these are four different lines of argument (and I will expound them over the course of this thesis).[[24]](#footnote-24)

The lack of unity in Kant’s thoughts on freedom is often remarked.[[25]](#footnote-25) I want to add that this lack of unity is not only inter-textual – most famously, the shift from *Groundwork* III to the second *Critique* (the so-called ‘great reversal’[[26]](#footnote-26)) – but also intra-textual. And *Groundwork* III exhibits this intra-textual diversity more than any other of Kant’s works. Kant is desperate to establish that we are free (such that we stand under the moral law). The first *Critique* made freedom possible (through transcendental idealism), and the moral law (as analysed in *Groundwork* I and II) requires it. The stakes are high, and Kant throws everything he has got at this.

This makes *Groundwork* III an unbelievably rich text. In a mere 18 (Academy) pages, we encounter (at least) 4 different arguments for freedom. Kant appeals to the spontaneity of rational judgment, the activity involved in (discursive) cognition (and transcendental idealism), reason’s capacity for ideas, and our consciousness of the causality of our reason.

I will attempt to reconstruct a single coherent argument for freedom from *Groundwork* III. In doing so, I will draw upon various parts of the text and circumvent others. And I am happy to do this. I do not seek to provide a complete clarification of the text of *Groundwork* III, such that it provides a single coherent argument for freedom.[[27]](#footnote-27) I think that might be impossible. In this respect, *Groundwork* III is like an old piece of Tupperware: You can pin three corners down, but in attempting to pin down the fourth, one of the others inevitably pops up.

Allow me to begin to get things in view. For this, it helps to set out Kant’s target. As noted earlier, a deduction has a specific meaning for Kant: it is an attempt to entitle us to something that we take ourselves to possess, in this case, freedom. Now, a deduction is only needed when our claim to possession is put in doubt.[[28]](#footnote-28) In this case, we take ourselves to be free and something puts this into question.

There are two worries here. The first concerns the general threat of causal determinism. If our behaviour is completely causally determined, it will not be up to us how we act. Whether or not we follow the supreme principle of morality would not be up to us, and morality would be a “phantasm” (IV: 445. 9) for creatures like us. At one point, Kant expresses this in legal terminology of a deduction:[[29]](#footnote-29)

[…] it is not left to the philosopher’s discretion whether he wants to remove the seeming conflict [between freedom and natural necessity], or leave it untouched; for in the latter case the theory about this is a bonum vacans [or vacant good[[30]](#footnote-30)], of which the fatalist can with good reason seize possession and chase all moral science from its supposed property as possessing it without title. (IV: 456. 29-33)

Before I continue, I should note that Kant has little time for (what we would call) a compatibilist conception of freedom. He famously derides it as the “freedom of a turnspit” (V: 97. 19), and declares any such attempt to solve the free-will problem as a “wretched subterfuge” (V: 96. 1). I agree with him on this, but will not make this case here. Instead, I follow Kant in his attempt to show how a libertarian conception of freedom (that allows for agents to do otherwise) is possible.

The second issue concerns the type of freedom that we possess. We might have some form of libertarian freedom, but it could be limited. We might still be ultimately governed by instincts and desires – in total, the sensuous part of human nature. If this were the case, our limited freedom might allow us to follow hypothetical imperatives, but we would be incapable of following a categorical imperative, or acting for the sake of the supreme principle of morality.[[31]](#footnote-31)

These two worries about freedom threaten that morality (as Kant analyses it in *Groundwork* I and II) might be a phantasm for us. In *Groundwork* III, Kant seeks to overcome this. He hopes to secure our right to think of ourselves as (suitably) free, and thereby as standing under the moral law.

This is the topic of my thesis. I argue that Kant can entitle us to regard ourselves as free in such a way that we stand under the moral law. Allow me to briefly re-sketch my simplified version of this argument:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

I will work my way through this argument in the first two chapters, then turn to defend it against the above two worries in the third and fourth. The stakes are high. Morality – or at least Kant’s conception of it – hangs in the balance.

There is an important additional element to this. For Kant, morality itself depends upon freedom.[[32]](#footnote-32) As I have just argued, if we are not free, we would be unable to stand under the moral law, and would thereby not be moral agents. In addition however, we would also lack our distinctive moral status. Our (distinctive) moral status is due to our freedom; In *Groundwork* II, Kant tells us that rational nature is the ground of the categorical imperative (IV: 429. 2-3), and that autonomy is the ground of the dignity of rational nature (IV: 436. 6-7). I will return to consider the complicated relationship between freedom, autonomy and morality in the second chapter of this thesis. For now, it suffices to note that, if we were not free, we would lack the distinctive moral status we take ourselves to possess. Moreover, if there were no free beings in existence, then morality would be a “chimerical idea without any truth” (IV: 445. 6), “the mere phantasm of a human imagination overreaching itself through self-conceit” (IV: 407. 17-9).[[33]](#footnote-33) This serves to raise the stakes for our argument even further.

Before we turn to *Groundwork* III, and the argument itself, I would like to clarify some key issues that help situate it. I want to emphasise that this is a non-moral argument for freedom (§2), and not a proof or justification of the moral law (§3).

# Two Types of Argument for Freedom in Kant

Freedom lies at the heart of Kant’s philosophy.[[34]](#footnote-34) And over the years, he attempted many different arguments for it. In general, we can find two distinct types of argument: one moral, and the other non-moral. On the non-moral approach, it is usually something about reason in general that proves we are free. On the moral approach, it is something about morality specifically that proves we are free.

For a while, especially in the first half of the 1780s, Kant attempted a non-moral argument for freedom, based on the spontaneity involved in reason.[[35]](#footnote-35) We can find this approach in the first edition of the first *Critique*, his lectures on metaphysics, his *Review of Schulz*, and of course *Groundwork* III.[[36]](#footnote-36)

Immediately following the *Groundwork* however, Kant appears to abandon this approach, to instead present a moral argument for freedom.[[37]](#footnote-37) Very roughly, we recognise the moral law and this reveals our freedom to us; we judge that we ought to do something, and this reveals that we can. This moral approach seems to end up as Kant’s considered view. And we can find traces of this argument in his earlier texts, including *Groundwork* III itself.[[38]](#footnote-38) Some commentators claim that, even in *Groundwork* III, Kant is primarily offering a moral argument for freedom.[[39]](#footnote-39) I depart from this, opting for a non-moral reading of *Groundwork* III, on both textual and other grounds.

As far as the text is concerned, I think that a non-moral reading of *Groundwork* III is the most natural. Alongside the four non-moral arguments for freedom that Kant appears to give, the notorious circle passage also provides strong evidence in favour of this reading. In the third subsection of *Groundwork* III*,* Kant raises the objection that his approach might be circular. The worry is that:

[…] our inference from freedom to autonomy and from it to the moral law contained a covert circle, namely that perhaps we were presupposing the idea of freedom only for the sake of the moral law, in order afterwards in turn to infer it from freedom, and hence were unable to state any ground of it […] (IV: 453. 4-8)

I will say more about the circle passage itself in due course. But for now, I want to draw attention to Kant’s worry that he might be “presupposing the idea of freedom only for the sake of the moral law” (IV: 453. 5-7). Here, I side with Allison in taking Kant’s claim to have avoided a circle at face-value.[[40]](#footnote-40)

The non-moral reading also finds a potentially unique argument in *Groundwork* III, and not just a jumbled telling of the moral argument for freedom that Kant would later tell from the second *Critique* onwards.

Alongside its textual merits, there are other good reasons to opt for a non-moral reading. For one, a non-moral argument for freedom addresses a challenge that a moral argument does not. The worry at hand is whether we are sufficiently free to comply with moral demands. The moral argument begins with our consciousness of standing under the moral law, and infers our freedom from that. In this way, it runs afoul of Kant’s circle – “presupposing the idea of freedom only for the sake of the moral law” (IV: 453. 5-7). Of course, this might not be so problematic. But, for a while at least, Kant thought that it was. As we have just seen, the circle passage suggests that Kant was interested in finding a non-circular response to this challenge. And a non-moral argument for freedom attempts just this.

But more importantly, the non-moral approach allows for a non-moral conception of freedom. It is reason in general that establishes our freedom, which allows for us to be free in both moral and non-moral matters. This approach has two main benefits. It allows for a unity between theoretical and practical reason and also helps circumvent an age old worry about Kant’s theory of freedom – Reinhold’s objection. Let me say a little more about each of these points in turn.

In the preface to the *Groundwork*, Kant tells us that:

[…] I require that the critique of a pure practical reason, if it is to be complete, also be able to present its unity with speculative reason in a common principle; because in the end there can only be one and the same reason which must differ merely in its application. (IV: 391. 24-9)[[41]](#footnote-41)

In this thesis, I will argue that reason in general involves spontaneity, and then extend this insight to the practical case, encompassing both moral and non-moral matters. I thus offer a reading of Kant where theoretical and practical reason both involve a similar freedom.

A non-moral conception of freedom also helps us overcome Reinhold’s objection. Karl Reinhold famously worried that if we equate freedom with acting morally, we will be unable to make proper sense of immoral action.[[42]](#footnote-42) Very briefly, if freedom consists in acting morally, then immoral action is unfree, and thereby something that can neither be imputed to us nor something that we can be held responsible for. Michelle Kosch suggests that Kant’s reversal from the *Groundwork* to the second *Critique* lays the path for this objection:

This epistemic subordination of freedom to the moral law will have interesting consequences for what Kant is able to say about moral evil (Kosch 2006: 37).[[43]](#footnote-43)

Allison argues that Reinhold’s objection – and various other classic worries with Kant’s theory of freedom – can be overcome with a proper understanding of Kant’s view on rational agency in general, that is, with a developed account of non-moral freedom.[[44]](#footnote-44) I agree with him on this.

An exclusively moral conception of freedom is unsatisfying. Not only does it open Kant up to Reinhold’s objection, but it also robs us of choice. The moral approach reveals that we have a capacity to act morally, but says nothing about whether this extends to non-moral choices. Timmermann has convincingly argued that, on Kant’s later moral approach, freedom “has nothing to do with choice of different options” (2007: 177). It consists in the ability “to do otherwise *only when we do wrong*” (Timmermann 2007: 177), and we end up having to “bite the Humean bullet with regard to non-moral projects” (Timmermann 2006: 91n39). I hope to avoid this bullet. Freedom is important in non-moral matters – relationships, projects and so on.

Timmermann insists that Kant had such an exclusively moral conception of freedom (2006: 84-5; 2007: 123n7, 177) and argues that Reinhold (and Sidgwick) went astray in opting for a non-moral conception of freedom (2007: 164). However, he also recognises that Kant’s exclusively moral conception of freedom is problematic, in that it does run afoul of a variant of Reinhold’s objection, and concedes that: “Maybe Reinhold and Sidgwick have a point after all” (Timmermann 2007: 166-7).

I agree with Timmermann that from the second *Critique* onwards, Kant primarily operates with an exclusively moral conception of freedom.[[45]](#footnote-45) I also share the worry that such a conception of freedom is problematic. Through turning back to the *Groundwork*, I hope to find a non-moral conception of freedom that offers a way around this.

Of course, much more needs to be said here.[[46]](#footnote-46) Reinhold’s objection is both complicated and significant, as it gets to the heart of the relationship between freedom and morality in Kant. One issue that stands out is how we are to understand Kant’s claim (in our second premise) that a free will is *under* the moral law. I will return to address this in detail in chapter 2.

I do not take the preceding considerations to count decisively against a moral argument for (or moral conception of) freedom in Kant. That would be far too hasty. My point is rather that there are good reasons (both textual and otherwise) to consider a non-moral argument for, and non-moral conception of, freedom in Kant. In the end though, this will depend on whether these arguments work, and that is something that concerns this thesis as a whole, for which the verdict will have to wait.

# What *Groundwork* III is not

It is important to emphasise one thing that *Groundwork* III is *not*. It is not an attempted proof or justification of the moral law.[[47]](#footnote-47) *Groundwork* III does deal with sceptical challenges, but not this one.

In the third subsection of *Groundwork* III, Kant poses the following question:

But why, then, ought I to subject myself to this principle, and do so as a rational being as such[?] (IV: 449. 11-3).

This is not however, the voice of the moral sceptic, who requires non-moral reasons for acting morally. Moreover, this is for the best.[[48]](#footnote-48) If Kant were to provide non-moral reasons for acting morally, he would renounce the idea of duty for duty’s sake, and reduce morality to hypothetical imperatives[[49]](#footnote-49) – *if* you want certain non-moral things, *then* you ought to act morally. A non-moral proof of the moral law is hopeless, and Kant does not attempt one in *Groundwork* III.

This point has gained traction in the recent literature: Timmermann has convincingly argued that Kant is not addressing a (radical) moral sceptic (2007: 129-30); Stern has noted that Kant is not addressing a sceptic who stands outside morality, but instead one who is inside the moral life, but finds it problematic or puzzling from within (2010: 463); Allison sees Kant as addressing a “curious and sympathetic meta-ethicist” (2011: 309).

I agree with these accounts. The *Groundwork* takes common rational cognition as its starting point, and analyses that. In *Groundwork* III, Kant is addressing someone who has followed him through the book, but is unsure of whether *we* are capable of standing under the moral law. And this worry stems from our constitution. We are rational, but we are also animals (sensibly affected creatures). Unless we are free, the moral law will be a phantasm for us. And here I think that Kant is offering a non-moral argument, to the effect that we are free and thereby entitled to think of ourselves as standing under the moral law. This is the sceptical challenge that *Groundwork* III addresses (alongside the issue of the imperatival force of the moral law).

In setting things up this way, I depart from a constructivist reading of *Groundwork* III, popularised by Christine Korsgaard and Onora O’Neill, where Kant (supposedly) attempts something like a proof of morality through thinner non-moral premises about the nature of rational agency. I detail the relevant aspects of this departure in chapter 2 (§§1.4-1.5).

# Some Methodological Remarks

My key reference is the *Groundwork* itself. After that, I draw heavily upon the work of Henry Allison. Perhaps more than anyone else, Allison has extensively defended Kant’s claims that reason involves freedom, and that to be free is to stand under the moral law. As noted earlier, Allison finds a non-moral conception of freedom in Kant, which he defends at length. This forms the basis of his Incorporation Thesis. He also famously interprets and defends Kant’s claim that freedom and the moral law are reciprocal concepts, which he calls the ‘Reciprocity Thesis’. Allison thus defends both premises of my argument. However, he also thinks that the argument of *Groundwork* III fails. Returning to my brief sketch:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

Allison extensively defends versions of premises 1 and 2. However, alongside many other commentators, he thinks that Kant fails to establish that we have wills, such that premise 1 fails. He also insists that the conception of freedom involved in premise 1 is weaker than the conception of freedom required by premise 2, such that the argument itself is invalid (the conclusion supposedly does not follow due to a crucial equivocation in the use of the term ‘free’). I think that Allison is mistaken in this diagnosis, and will attempt to show so.

More broadly, Allison has offered the most extensive philosophical engagement with Kant’s deduction of freedom to date. Dieter Schӧnecker has recently offered a comprehensive account of *Groundwork* III, but his approach is mostly textual. We can see this clearly in the introduction to one of his papers, where he tells us that:

The purpose of this paper is not to criticize or make philosophical use of Kant’s deduction. As I have argued elsewhere, one of the major obstacles of serious (i.e. historical) Kant-research is the inability (and unwillingness) to distinguish between the question of what a text means and the question of whether what it manifests is true. (Schӧnecker 2006: 301)

Of course, what a text means and whether what it manifests is true are separate questions. However, the purpose of this thesis is precisely to make philosophical use of Kant’s deduction of freedom. And here, I want to resist Schӧnecker’s equating serious and historical Kant-research. Careful textual analysis and careful philosophical reconstruction are both serious tasks. Schӧnecker has extensively engaged in the former, but the latter is my focus.

One final methodological remark: While I defend a version of *Groundwork* III, I also distance myself from Kant at certain key points. Some of these departures are minor, and I will spare you the details for now. Others are major – I end up rejecting a non-metaphysical account of freedom (ch.1), a formalist or constructivist meta-ethics (ch. 2), and transcendental idealism itself (ch.3).

Certain departures are inevitable – the text often pulls in two (or more) directions, and one is forced to choose. I think that other departures are required to preserve Kant’s key insights. I accept (what I take to be) Kant’s key insights about reason, freedom and morality, but often worry that his system cannot sustain them. In general, I stay with the spirit of Kant’s philosophy, but in places, depart from the letter. In this, I locate my project in the post-Kantian tradition. I begin with Kant, but often end up alongside his successors (and some of their present-day followers).

# Overview

I want to end this introduction by providing an overview of things to come. This thesis defends Kant’s argument that we are sufficiently free to stand under the moral law. Here it is again, at its most basic:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

I present this argument in my first two chapters, and then go on to defend it against two major objections in the final two chapters.

Allow me to spell this out in a little more detail.

## Chapter 1: Reason, Spontaneity and Agency

I begin by considering the spontaneity involved in reason. I proceed in the same way that Kant does in the second subsection of *Groundwork* III. I start with the spontaneity involved in rational judgment, before extending this to the practical case, establishing the spontaneity involved in practical judgements, action, maxims, and more broadly, rational agency. I conclude, with Kant, that a rational will is a free will.

## Chapter 2: Freedom, Autonomy and Morality

In my second chapter, I turn to consider the relationship between freedom and morality in Kant. Famously, Kant holds that freedom and morality are reciprocal concepts.[[50]](#footnote-50) I explore precisely what to make of this claim.

I consider both an ambitious and modest account of the relationship between freedom and morality. On the ambitious account, freedom provides something like a justification of morality. I argue against this approach, along with Kant’s formalism and Kantian constructivism, which I associate with it. In its place, I offer my own modest account of the relationship between freedom and morality, where freedom allows us to stand under the moral law.

With the premises of my argument defended, I then turn to consider the objections. Many commentators also defend the premises of Kant’s argument, but reject the argument itself. This is the main debate my thesis plugs into, a debate between people who are sympathetic to the Kantian project (broadly conceived). Thus, while I do offer a defence of premises 1 and 2, I realise that elements of this will be unsatisfying to those unsympathetic to Kant’s project in general. My target lies elsewhere. I am interested in those who are sympathetic to Kant, and his claims about reason and morality, but think that *Groundwork* III fails.[[51]](#footnote-51) And here, two objections stand out amongst the others.

## Chapter 3: Practical Reason

The first objection is that Kant has failed to establish that we have wills, or that our reason is practical. The starting point of *Groundwork* III is a rational will, where the first premise states that ‘A rational will is a free will’. From here, we attempt to move through freedom to the moral law, but the crucial issue of whether or not we are rational beings *with wills* still remains.

I argue that this objection can be overcome. In doing so, I draw upon Kant’s appeal to the phenomenology of agency in the often overlooked fifth subsection of *Groundwork* III. I argue that this provides excellent evidence for our being rational beings with wills.

I also offer some general epistemological reflections about doubt. I consider what drives Kant to doubt that we have wills. His worry is that natural necessity might make freedom impossible. Kant wants to show how freedom is possible in the face of natural necessity. Transcendental idealism is his solution, which locates freedom outside of nature. I accept that this makes freedom possible, but object that (amongst other things) it precludes the recognition of other rational agents.

I conclude that transcendental idealism fails to provide an adequate account of freedom. In its place, I sketch an alternative picture of how freedom is possible, one that locates freedom within, rather than outside of nature. Through this, I hope to vindicate our conception of ourselves as free.

## Chapter 4: Rational Agency and the Moral Law

The second key objection to *Groundwork* III is that the freedom involved in rational agency is weaker than the freedom required by the moral law. Kant’s argument then apparently contains an equivocation in the use of the term ‘free’ between premises 1 and 2, and is thereby invalid. I argue that, with the ground laid in the previous chapters, this objection drops away.

## Conclusion

With these two objections overcome, I conclude that we can follow Kant from reason to freedom to the moral law. We are entitled to think of ourselves as free. The moral law is no phantasm for us.

1 – Reason, Spontaneity and Agency

The first premise of my argument is that:

1. A rational will is a free will

In this chapter, I unpack this claim as I find it in *Groundwork* III. For this, I turn to the text. As noted in the Introduction (§1.3), I find four different arguments for freedom in *Groundwork* III. They are:

1. The claim that freedom must be presupposed as a property of the will of all rational beings (IV: 447. 26 – 448. 22).
2. An appeal to transcendental idealism and the activity involved in (discursive) cognition (IV: 451. 1-36).
3. The claim that reason’s capacity for ideas reveals a “spontaneity so pure” (IV: 452. 18) that it thereby takes us far beyond anything sensibility can afford us, and marks us out as members of an intelligible world (IV: 452. 6-22).
4. An appeal to the consciousness of the causality of our reason (IV: 457. 4-7, IV: 457. 22-4, IV 459. 9-14, IV: 461. 17-25).

We now face the question of which – if any – of these arguments establish a link between reason and freedom. The appeal to the consciousness of the causality of our reason (D) and the activity involved in (discursive) cognition (B) do not. They draw upon an awareness of our own activity to establish that we are rational beings, and specifically, rational beings with wills. This is the topic of chapter 3, and I will return to consider these claims in detail there.

That leaves us with A and C. Kant’s claim that freedom must be presupposed as a property of the will of all rational beings (A) has received a lot of attention in the literature, whereas his claim that reason’s capacity for ideas marks us out as members of an intelligible world (C) has not. I will introduce and consider both arguments, before defending a version of A.

This chapter takes the following structure. In the first section, I introduce (A) Kant’s claim that freedom must be presupposed as a property of the will of all rational beings(§1). I begin by looking at this claim at its most basic, namely, that we must act under the idea of freedom. I consider a powerful objection to this: Inescapability does not entail justification, and thus the basic claim that we must act under the idea of freedom does not thereby justify our employing the concept of freedom. I argue that inescapability is not enough, and that Kant needs to establish a link between reason and spontaneity.

I consider two attempts to make this link. The first is (C) Kant’s claim that reason’s capacity for ideas reveals our freedom (§2). I criticise this argument, and move on to consider one key sentence in the original passage (A), where Kant posits a connection between rational judgement and spontaneity. The heart of this chapter concerns my attempt to unpack this in a way that establishes a link between reason and spontaneity (§3). Once I have done this, I extend the result to the practical case, to argue that a rational will is a free will (§3.2), and consider some of the metaphysical implications of this (§3.3).

# Freedom must be Presupposed as a Property of the Will of all Rational Beings

The second section of *Groundwork* III is titled: *Freedom must be presupposed as a property of the will of all rational beings*. It consists of a single long paragraph, which is worth quoting in full. Here it is:

It is not enough that, on whatever grounds, we attribute freedom to our will if we do not have grounds sufficient to attribute it to all rational beings as well. For since morality serves as a law for us only as for *rational beings*, it must hold for all rational beings as well, and as it must be derived solely from the property of freedom, freedom must also be proved as a property of the will of all rational beings; and it is not enough to establish it from certain supposed experiences of human nature (though this is actually absolutely impossible and can solely be established a priori), but one must prove it as belonging to the activity of rational beings endowed with a will as such. Now I say: every being that cannot act otherwise than *under the idea of freedom* is actually free, in a practical respect, precisely because of that; i.e. all laws that are inseparably bound up with freedom hold for it just as if its will had also been declared free in itself, and in a way that is valid in theoretical philosophy. Now I assert: that we must necessarily lend to every rational being that has a will also the idea of freedom, under which alone it acts. For in such a being we conceive a reason that is practical, i.e. has causality with regard to its objects. Now, one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse. Reason must view herself as the authoress of her principles, independently of alien influences, and must consequently, as practical reason, or as the will of a rational being, by herself be viewed as free; i.e. its will can be a will of its own only under the idea of freedom, and must thus for practical purposes be ascribed to all rational beings. (IV: 447. 28 – 448. 22)

Kant begins by asserting the importance of ascribing freedom to all rational beings with wills (or rational agents), something that is crucial given his project in the *Groundwork* – he insists that morality, if it is to be absolutely necessary, must hold for all rational agents.[[52]](#footnote-52) He then claims that all rational agents must act under the idea of freedom, before going on to argue for this.

The argument occurs over the last four sentences of the passage. This is an incredibly dense piece of text, and I will leave its finer details aside for the moment. I will instead begin by exploring Kant’s claim at its most basic, namely that rational agents must act under the idea of freedom (§1.1). I will then consider a powerful objection to this (§1.2), before eventually returning to the text to offer a defence of this claim, and more broadly, Kant’s move from reason to freedom (§3).

## 1.1 We must act under the Idea of Freedom

It is instructive to begin with the basic claim that we must act under the idea of freedom. Kant makes this claim in various places prior to the *Groundwork* [1785]. Consider the following passage from Kant’s lectures on metaphysics [1782-3]:

Freedom is a mere idea and to act in conformity with it is what it means to be free in the practical sense […] freedom is practically necessary – man must therefore act according to an idea of freedom and he cannot act otherwise (XXIX: 898).[[53]](#footnote-53)

He makes a similar claim in his *Review of Schulz* [1783]:

[…] the practical concept of freedom has nothing to do with the speculative concept, which is abandoned entirely to metaphysicians. For I can be quite indifferent as to the origin of my state in which I am now to act; I ask only what I now have to do, and then freedom is a necessary practical presupposition and an idea under which alone I can regard commands of reason as valid. Even the most obstinate skeptic grants that, when it comes to acting, all sophistical scruples about a universally deceptive illusion must come to nothing.

In the same way, the most confirmed fatalist, who is a fatalist as long as he gives himself up to mere speculation, must still, as soon as he has to do with wisdom and duty, always act *as if he were free*, and this idea also actually produces the deed that accords with it and can alone produce it. It is hard to cease altogether to be human. (VIII: 13)

In these passages, Kant explicitly states that freedom is practically necessary – we cannot act other than under the idea of freedom.

Before we continue, it is worth noting that in the passage from the *Review of Schulz*, we can find both a non-moral and a moral argument for freedom, which correspond to two different ways we might read the claim that freedom is practical necessity; the first concerns the possibility of action, the second specifically moral action.[[54]](#footnote-54) As noted in the Introduction (§2), it is the first type of argument that interests me in this thesis.

Leaving aside the moral reading, the basic idea is that, insofar as one acts, one must regard oneself as free. It is what it means to be an agent that one takes things as – in some sense – up to oneself. This general point seems plausible, and it is relatively well accepted in the Kant literature,[[55]](#footnote-55) and some places beyond.[[56]](#footnote-56) Daniel Dennett, for one, devotes a chapter of *Elbow Room* to acting under the idea of freedom, and claims that:

We cannot help acting under the idea of freedom, it seems; we are *stuck* deliberating as if our futures were open. (Dennett 1984: 108)

There is some agreement that we must act under the idea of freedom. The real question is what exactly this amounts to. To bring this out, let us turn to consider an objection.

## 1.2 The Inescapability of Freedom

Our basic claim meets with a basic objection. The fact that we must act under the idea of freedom does not make us actually free. And given this, a question arises as to whether we are justified in acting under the idea of freedom. In general, inescapability (or non-optionality) does not entail justification.[[57]](#footnote-57) To say that we cannot act otherwise than under the idea of freedom seems to offer an excuse for this, rather than a justification.[[58]](#footnote-58)

Robert Stern confronts Kant’s argument with this very issue:[[59]](#footnote-59)

The philosophical worry is that this is just not a very satisfactory argument. For, it might be asked, how does proving that we must believe we are free, or assume we are free, or act as if we were free, or regard ourselves as free, even though we might not be, help settle our worries on this issue? In fact, might it not seem to make them worse, as we now seem to have to believe, assume, etc things about ourselves that (for all the argument shows) could well be false? And isn’t this the worst kind of sceptical nightmare, locking us into views of ourselves that we cannot escape even if they are in error? (Stern: forthcoming).

So much for the inescapability of freedom. Our acting under the idea freedom, if it only admits of exculpation (as we cannot do otherwise) might turn out to be “the worst kind of sceptical nightmare”.

As we discussed in the Introduction (§1.2), Kant tells us that he is offering a *deduction* of freedom,[[60]](#footnote-60) and as such, his distinction between *quid facti* and *quid juris* is useful.[[61]](#footnote-61) *Quid facti* concerns the fact of our possession, whereas *quid juris* concerns our entitlement to a possession. In the case at hand, the *quid facti* concerns our employing the concept of freedom and the *quid juris* concerns our being entitled to do so.[[62]](#footnote-62) More precisely, the *quid facti* concerns whether we act under the idea of freedom and the *quid juris* concerns whether we are entitled to act as such.

This brings us back to our basic claim that we must act under the idea of freedom. The practical necessity answers the *quid facti*, but leaves open the crucial issue of the *quid juris*, namely, whether we are justified in acting under the idea of freedom.[[63]](#footnote-63)

For now, it will be instructive to clarify what exactly is meant by ‘must’ in the claim that we must act under the idea of freedom.[[64]](#footnote-64)

Perhaps the simplest form of non-optionality is something like brute necessity. At times, Korsgaard (echoing Sartre) suggests such an approach. She opens her book *Self-Constitution* with the claim that: “Human beings are *condemned* to choice and action” (Korsgaard 2009: 1). She notes concerning the nature of this necessity that:

The necessity of choosing and acting is not causal, logical or rational necessity. It is our *plight*: the simple and inexorable fact of the human condition. (Korsgaard 2009: 2)

Such necessity however – no matter how profound – does not entail justification. As we saw above, non-optionality may serve to explain our acting under the idea of freedom, but this is a matter of exculpation rather than justification.[[65]](#footnote-65)

Let us turn to briefly look at another form of necessity, the psychological. There might be some peculiarity of human psychology that means that *we* must act under the idea of freedom. This strategy however suffers a similar fate to the last.[[66]](#footnote-66) We can draw upon Wizenmann’s challenge to Kant to illustrate this.[[67]](#footnote-67)

Perhaps it is a necessary feature of being in love that one assumes the existence of the object of one’s love. If it were, this supposed quirk would neither establish the existence of the object of one’s love, nor rationally justify belief in its existence.[[68]](#footnote-68) In general, psychological necessitation does not entail justification.[[69]](#footnote-69)

Imagine a trip to the psychoanalyst: “Doc, I’m in love with Elvis Presley. I can’t get him out of my head, even though he’s dead. We’re going to be married and live together forever”. The doctor examines me and determines that I’m stuck with this belief, due to some incurable form of compulsion.[[70]](#footnote-70) That might help me in coming to terms with my belief, but it certainly would not help justify it. Elvis Presley is *not* going to marry me, and it would be rational for me to be responsive to this. We might then not blame me for holding this belief, but *if I could* get rid of it, I would be more rational for doing so.[[71]](#footnote-71) The necessity in question merely offers an excuse for the erroneous belief associated with infatuation.

These forms of necessity fail to justify our acting under the idea of freedom. Another strategy is to claim that any attempt to argue against freedom is self-stultifying.[[72]](#footnote-72) This thought runs as follows: If we must act under the idea of freedom, then the very act of arguing against freedom occurs under the idea of freedom, and thereby undermines itself.

At its most basic, this strategy does not help. It is certainly tempting to put the denial of freedom into the voice of a hypothetical opponent, and then just sit back and watch them performatively undermine their own claim. In doing this however, we fail to take the claim itself seriously.[[73]](#footnote-73) And to see this we only need to assume that, in fact, we are not free. If we were not free, but stuck acting under the idea of freedom, this would not count against the truth of the claim that we are not free, but rather, it would count against us, as being unable to respond to the truth of the situation. And as Stern recognises, this again, is the “worst kind of sceptical nightmare, locking us into views of ourselves that we cannot escape even if they are in error” (forthcoming).

There is another line of thought sometimes at work here, which is that freedom is necessary for reason. If we accept this connection, then there would be something incoherent about rationally arguing against our freedom, but the fact that we must act under the idea of freedom is not doing the work. The work is instead done by the connection between reason and freedom itself. And on this note, let us move on to the next section of this chapter.

In this section, we have looked for a justification for our acting under the idea of freedom. We considered some basic attempts to make sense of the necessity of acting under the idea of freedom, and a self-stultification argument. All of these attempts failed. However, as I have just suggested, there is another avenue, namely a connection between reason and freedom.[[74]](#footnote-74) In what follows, I turn to the details of the third (C) and second subsections (A) of *Groundwork* III to consider two attempts to make this connection. This will require some effort, but will allow us to eventually provide a response to the objection at hand (§3.3).

# Reason’s Capacity for Ideas and the Intelligible World

At one point in the third subsection of *Groundwork* III, Kant appears to argue from the possession of (theoretical) reason to our membership in an intelligible world and hence our freedom. Here is the relevant passage:

Now, a human being actually does find in himself a capacity by which he is distinguished from all other things, even from himself, in so far as he is affected by objects, and that is *reason*. As pure self-activity, it is elevated even above the *understanding* in this: that though the latter is also self-activity and does not, like sense, contain merely representations that arise when one is affected by things (and thus passive), still it can produce from its activity no other concepts than those which serve merely *to bring sensuous representations under rules* and thereby to unite them in one consciousness, and without this use of sensibility it would think nothing at all; whereas reason under the name of the ideas shows a spontaneity so pure that thereby he goes far beyond anything that sensibility can ever afford him, and provides proof of its foremost occupation by distinguishing the world of sense and the world of understanding from each other, and thereby marking out limits for the understanding itself.

On account of this a rational being must view itself, *as an intelligence* (thus not from the side of its lower powers), as belonging not to the world of sense, but to that of understanding; […]

As a rational being, hence as one that belongs to the intelligible world, a human being can never think of the causality of his own will otherwise than under the idea of freedom; for independence from the determining causes of the world of sense (such as reason must always ascribe to itself) is freedom. (IV: 452. 7-35).

This argument draws upon a supposed absolute spontaneity involved in Reason’s capacity to form ideas.[[75]](#footnote-75) As we have just seen:

[…] reason under the name of the ideas shows a spontaneity so pure that thereby he goes far beyond anything that sensibility can ever afford him (IV: 452. 17-9)

It will be helpful here to clarify what Kant means by ideas, and to spell out their connection to Reason and freedom.

For this, we need to turn to the first *Critique*. In the first book of the Dialectic, *On the concepts of pure reason*, in the section, *On the ideas in general*, Kant tells us that:

A concept is either an **empirical** or a **pure concept**, and the pure concept, insofar as it has its origin solely in the understanding (not in a pure image of sensibility), is called *notio*. A concept made up of notions, which goes beyond the possibility of experience, is an **idea** or a concept of reason. (A 320/B 377)

In an earlier discussion of Plato, he notes how ideas take us beyond the sensible world:[[76]](#footnote-76)

Plato made use of the expression **idea** in such a way that we can readily see that he understood by it something that not only could never be borrowed from the senses, but that even goes far beyond the concepts of the understanding[[77]](#footnote-77) (with which Aristotle occupied himself), since nothing encountered in experience can ever be congruent to it. (A 313/B370)

Ideas go beyond concepts of the understanding, and the possibility of experience. Kant continues to note that:

Plato found his ideas pre-eminently in everything that is practical, i.e. in what rests on freedom (A 314-5/B 371).

If we abstract from its exaggerated expression, then the philosopher’s spiritual flight, which considers the physical copies in the world order, and then ascends to their architectonic connection according to ends, i.e., ideas, is an endeavour that deserves respect and imitation; but in respect of that which pertains to principles of morality, legislation and religion where the ideas first make the experience (of the good) itself possible, even if they can never be fully expressed in experience, perform a wholly unique service, which goes unrecognized precisely because it is judged according to empirical rules, whose validity as principles should be cancelled by those very ideas. For when we consider nature, experience provides us with the rule and is the source of truth, but with respect to moral laws, experience is (alas!) the mother of illusion, and it is most reprehensible to derive the laws concerning what I **ought to do** from what **is done**, or to want to limit it to that. (A 318-9/B 375)

It is interesting that Plato’s “spiritual flight […] deserves respect and imitation; but in respect of that which pertains to principles of morality, legislation and religion” (A 318/B 375). Kant appears to claim that the ideas of Reason have practical, but not theoretical merit.

If morality is the only appropriate domain of the ideas of Reason, then it is not so clear that Kant’s appeal to Reason’s capacity to form ideas counts as a non-moral argument for freedom. However, ideas themselves are still formed by theoretical Reason, and in the *Critique*, Kant does proceed to discuss the important theoretical role that ideas of Reason have.[[78]](#footnote-78)

Moving on from this, we need to clarify what exactly it is about Reason and ideas that reveals an absolute spontaneity. As we saw earlier, Kant claims that, “reason under the name of the ideas shows a spontaneity so pure that thereby he goes far beyond anything that sensibility can ever afford him” (IV: 452. 17-9). Here, it seems that it is Reason’s capacity to form ideas that entitles us to view ourselves as members of an intelligible world.

Timmermann, for one, seems to think that this is Kant’s argument:

For now, the only thing that matters is the fact that human beings are aware of a capacity within themselves to think beyond the boundaries of experience. When we conceive of these ideas we make what might be called a leap of reason. Reason as pure spontaneity thus reaches far beyond the mere comparative spontaneity of the understanding, which is tied to empirical conditions (Timmermann 2007: 136)

This accurately reflects the passage at hand (along with A 313/ B370), but it does not strike me as a promising avenue.[[79]](#footnote-79) For one, an empiricist account of these ideas in terms of abstraction might be available. And even if abstraction is unable to adequately explain the origin of ideas, this might just point towards a form of nativism, rather than the “spontaneity so pure” (IV: 452. 18) that Kant suggests. Of course, the plausibility of empiricist and nativist accounts of ideas is beyond the scope of our discussion.[[80]](#footnote-80)

The more important issue however, is that it is unclear how Reason’s capacity for ideas could translate to the practical sphere. We do still need to establish that we have wills, or that our reason is practical, but that is the topic of chapter 3. The worry here is that, even if we have wills, if the only spontaneity involved in reason concerns its ability to form ideas, then it is not clear how this is of any use in the practical case. We might possess this spontaneity in our use of theoretical reason, but in the practical case, reason could still be entirely subservient to the sensible world.[[81]](#footnote-81) What we are after is something about reason in general that involves spontaneity. If we found this, it would apply equally in the theoretical and the practical case – both theoretical and practical reason would involve spontaneity. And Reason’s capacity to form ideas does not serve this role.

Paton latches on to a similar point:

It is not quite clear what Kant gains by his concentration on this special function of pure reason: indeed he himself goes on immediately to speak as if his argument were based on reason in its most general sense […]

But in spite of all this the argument seems to hold, if it holds at all, for theoretical reason in all its forms; for all reason must assume itself to be a power of spontaneous activity in accordance with its own principles, whatever be the objects to which it is directed. (Paton 1947: 239)

Paton is onto something. Throughout the rest of *Groundwork* III, Kant proceeds to talk of reason in general as involving freedom. Indeed, he does so at the end of the very passage at hand: “independence from the determining causes of the world of sense (such as reason must always ascribe to itself) is freedom” (IV: 452. 33-5).

Not only does the appeal to reason in general have textual support, it also seems more suited for the task at hand. We are looking for something about reason in general that involves spontaneity, such that it can be extended to the practical case. And we can find just this in the previous section of *Groundwork* III, in Kant’s argument for the claim that freedom must be presupposed as a property of the will of all rational beings (A). In this passage, Kant claims that rational judgement (in both the theoretical and practical cases) involves spontaneity. He even says as much in the very passage at hand. Recall the following:

As pure self-activity, it is elevated even above the *understanding* in this: that though the latter is also self-activity and does not, like sense, contain merely representations that arise when one is affected by things (and thus passive), still it can produce from its activity no other concepts than those which serve merely *to bring sensuous representations under rules* and thereby to unite them in one consciousness, and without this use of sensibility it would think nothing at all; whereas reason under the name of the ideas shows a spontaneity so pure […] (IV: 452. 9-18)

Here we see that the understanding involves self-activity, but that reason goes further than it to offer a “spontaneity so pure” (IV: 452. 18).[[82]](#footnote-82) I take it that Kant is concerned with our second objection to the approach of *Groundwork* III, namely whether the freedom involved in rational agency (in general) is sufficient for morality. This is an interesting question, but one I will return to in the fourth chapter of this thesis (where I will argue against a supposed gap between rational agency and the moral law). For now, I want to turn to the details of the second subsection of *Groundwork* III (A).

# Reason and Spontaneity

In this section, I argue that reason involves spontaneity. In doing so, I return to Kant’s argument that freedom must be presupposed as a property of the will of all rational beings (A). Let us take a closer look at the second half of this highly dense paragraph.

Now I say: every being that cannot act otherwise than *under the idea of freedom* is actually free, in a practical respect, precisely because of that; i.e. all laws that are inseparably bound up with freedom hold for it just as if its will had also been declared free in itself, and in a way that is valid in theoretical philosophy. Now I assert: that we must necessarily lend to every rational being that has a will also the idea of freedom, under which alone it acts. For in such a being we conceive a reason that is practical, i.e. has causality with regard to its objects. Now, one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse. Reason must view herself as the authoress of her principles, independently of alien influences, and must consequently, as practical reason, or as the will of a rational being, by herself be viewed as free; i.e. its will can be a will of its own only under the idea of freedom, and must thus for practical purposes be ascribed to all rational beings. (IV: 448. 4-22)

There are four sentences here, and it will help to go through these one by one.

Kant begins with an assertion:

Now I say: every being that cannot act otherwise than *under the idea of freedom* is actually free, in a practical respect, precisely because of that (IV: 448. 4-6)

As we saw earlier, the *necessity* of acting under the idea of freedom seems to be doing the work. It is “precisely because” we cannot but act under the idea of freedom that, in a practical respect, we are actually free. Kant attaches a famous footnote to this (IV: 448n), which appears to downplay the metaphysical aspects of this argument. I will return to consider this shortly (§3.3). For now, let us continue with the passage at hand.

Kant goes on:

Now I assert: that we must necessarily lend to every rational being that has a will also the idea of freedom, under which alone it acts. (IV: 448. 9-11)

This is the main claim of the second section. Indeed, it echoes the title: *Freedom must be presupposed as a property of the will of all rational beings*. Kant’s argument for this begins with a claim about the nature of reason and judgement:[[83]](#footnote-83)

Now, one cannot possibly think of a reason that would self-consciously receive guidance [*Lenkung empfienge[[84]](#footnote-84)*] from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse [*Antriebe*]. (IV: 448. 13-7)

The basic thought here seems to be that reason cannot self-consciously receive guidance with regard to its judgements, because if it did, it would attribute the determination of its judgements to this guidance and not to reason itself.

This sentence is crucial for my reading of *Groundwork* III. And so, I would like to name it. I will refer to it as the ‘Spontaneity Thesis’.[[85]](#footnote-85) I read it to signal an essential connection between reason, judgement, self-consciousness, and spontaneity. In what follows, I hope to illuminate these connections. This is a tough puzzle to piece together. I will begin by looking at some of these ideas as they figure in the first *Critique*. I will argue that reason involves spontaneity, and then attempt to clarify the essential insight with the help of an example (§3.1). I then extend this to the practical case (§3.2), before considering its metaphysical significance (§3.3).

## 3.1 The Spontaneity Thesis

At the start of the Transcendental Logic, Kant characterises human cognition as follows:

Our cognition arises from two fundamental sources in the mind, the first of which is the reception of representations (the receptivity of impressions), the second the faculty of cognizing an object by means of these representations (spontaneity of concepts); through the former an object is **given** to us, through the latter it is **thought** in relation to that representation (as a mere determination of the mind) […]

If we will call the **receptivity** of our mind to receive representations insofar as it is affected in some way **sensibility**, then on the contrary the faculty for bringing forth representations itself, or the **spontaneity** of cognition is the **understanding**. (A 50-1/B 74-5)

Kant views human cognition as essentially discursive, that is to say, it requires both intuitions and concepts. The former concerns sensibility and the latter spontaneity, which Kant identifies with the understanding; “the faculty for bringing forth representations itself, or the **spontaneity** of cognition is the **understanding**” (A 51/B 75).

Later on, towards the start of the Transcendental Analytic, Kant explains what the understanding does – it judges. He writes:

We can, however, trace all actions of the understanding back to judgments, so that the **understanding** in general can be represented as a **faculty for judging**. (A 69/ B 94)

The understanding subsumes intuitions under concepts, that is to say, it judges. We receive intuitions, and we spontaneously apply concepts to them. This gives us a very basic picture of the relationship between the understanding, judgement and spontaneity.

One major obstacle here is that it is never exactly clear what Kant means by spontaneity. Pippin notes that:

What is typical of such passages is that Kant leaves unexplained the meaning of such phrases as ‘producing representations from itself’ and concepts being ‘based on’ spontaneity, apparently confident that the contrast with receptivity clarifies what he is trying to get at. (Pippin 1987: 452-3)

Pippin is right. Kant invokes spontaneity as a catch all contrast to the passivity of reception, but this can be unhelpful, as there is more than one thing that we need to contrast with the passivity of reception. One such thing concerns the synthesis of the manifold of experience. Put crudely, Kant sees sensibility as a mess. It needs to be unified, and this requires us. This is not receptive, but active, or as Kant puts it, spontaneous. This is one way in which Kant uses the term ‘spontaneity’.[[86]](#footnote-86)

I am interested in another aspect of spontaneity, one that is intimately related to judgement and self-consciousness.[[87]](#footnote-87) In what follows, I explore some of the passages that suggest this conception of spontaneity. For this, we need to introduce a key player into the mix – the subject.

Kant famously begins §16 of the B-deduction as follows:[[88]](#footnote-88)

The **I think** must **be able** to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible or else would be nothing for me. (B 131-2)

We now have a subject in the picture. The ‘I think’ is a necessary condition of the very possibility of representation. It allows me to combine representations in one consciousness and thus makes thought possible for me; without such a unity, representations would amount to nothing *for me*. Kant calls this the transcendental unity of apperception. Some simple examples will help illuminate the claim.

I look around me, and judge that this coffee is cold, this computer slow, and this sentence cumbersome. In order for these judgements to amount to anything *for me*, I have to be able to grasp the elements involved – this coffee, the concept ‘cold’ – and be able to hold them together – this coffee *is* cold. It is no good being able to think about coffee, and the concept ‘cold’ independently, if I cannot hold them together in a unified consciousness. In Kant’s terms, the ‘I think’ must be capable of accompanying the representation ‘A is B’ insofar as this amounts to anything *for me*.

Kant then continues to link this to spontaneity. Immediately after introducing the ‘I think’, he writes that “this representation is an act of **spontaneity**, i.e., it cannot be regarded as belonging to sensibility” (B 132). He also claims that:

[…] the manifold representations that are given in a certain intuition would not all together be **my** representations if they did not all together belong to a self-consciousness, i.e., as my representations (even if I am not conscious of them as such) they must yet necessarily be in accord with the condition under which alone they **can** stand together in a universal self-consciousness, because otherwise they would not belong to me. From this original combination much may be inferred. (B 132-3)

As we can see in the last sentence of this passage, Kant does not shy away from the significance of the transcendental unity of apperception. He goes on to call it “the understanding itself” (B 134n), the “highest point” (ibid) in transcendental philosophy, and claims that this “principle is the supreme one in the whole of human cognition” (B 135).

With this, we find ourselves near the heart of the Critical system. Once again, there are two elements at play in Kant’s discussion of spontaneity in these passages. The first concerns the importance of our synthesis of the manifold, and the second the importance of self-ascription. We can see these two elements come apart in another key passage. Later on in the B-deduction, Kant writes that:

The synthetic unity of consciousness is therefore an objective condition of all cognition, not merely something I myself need in order to cognize an object but rather something under which every intuition must stand **in order to become an object for me**, since in any other way and without this synthesis, the manifold would **not** be united in one consciousness (B 138).

Here, Kant emphasises the importance of our synthesis of the manifold.[[89]](#footnote-89) He then immediately goes on to claim that:

This last proposition is, as we said, itself analytic, although, to be sure, it makes synthetic unity into the condition of all thinking for it says nothing more than that all **my** representations in any given intuition must stand under the condition under which alone I can ascribe them to the identical self as **my** representations, and thus can grasp them together, as synthetically combined in an apperception, through the general expression **I think** (B 138)

In this sentence, we see the importance of the possibility of self-ascription. Let me return to my simple examples to begin to explore the significance of this.

We have just seen that the ‘I think’ must be capable of accompanying the thought ‘A is B’ insofar as this thought amounts to anything *for me*. There is an additional element to this. For Kant, the fundamental act of human cognition is judgement. I do not just passively hold that ‘A is B’, I *judge* that ‘A is B’. Returning to my cold cup of coffee, insofar as ‘this coffee is cold’ amounts to a rational judgement *for me*, I must be able to take the attribution of ‘cold’ to the coffee as up to me.[[90]](#footnote-90)

This suggests that there are two senses in which the ‘I think’ must be able to accompany all of our thoughts. Firstly, insofar as the thought that ‘A is B’ amounts to anything for me, I have to be able to hold together A and B in a unified consciousness. Secondly, insofar as ‘A is B’ amounts to a rational judgement for me, I have to be able to take the attribution of B to A as up to me; *I* think A is B.

This second point is important. I want to return to the examples at hand, to address a potential counterexample here. And so, I look around again, and judge that this coffee is still cold, this computer still slow, and this sentence as cumbersome as ever. Now, there might be other ways to reliably track these things: my mug changes colour with the temperature of the coffee; my computer measures its own speed; and Microsoft Word puts squiggly green lines under my sentences. These look like potential counterexamples to Kant’s claim that rational judgement requires spontaneity. We can see this by returning to the Spontaneity Thesis, where Kant tells us:

[…] one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse. (IV: 448. 13-7)

It might seem that Kant is considering two basic ways in which our judgements might be determined: either spontaneously, or by an impulse. The latter seems inadequate, pushing us towards the former. The examples at hand (my heat responsive mug, speed monitoring computer, and disapproving Microsoft Word) point towards a possible neglected alternative: judgements might not be determined spontaneously, but might still reliably track the relevant features of the world. My coffee is cold, and my mug reliably tracks this without the type of spontaneity that Kant thinks rational judgement requires.

Self-consciousness is crucial here. My mug, computer and Microsoft Word all (in some sense) detect that A is B, but these are not rational judgements *for them*.[[91]](#footnote-91) A rational judgment is only a rational judgement *for me*, when I can take it as up to me. We can see this clearly in the Spontaneity Thesis:

[…] one cannot possibly think of a reason that would *self-consciously* receive from any other quarter with regard to *its* judgements, since the subject would not then attribute the determination of judgement to *his* reason, but to an impulse. (IV: 448. 13-7) [Emphases mine]

As self-conscious, I must be able to regard my judgements as up to me. This is what it means for them to be *my* judgements. And Kant himself goes on to make this very claim in the next sentence. He summarises the Spontaneity Thesis by saying that “[r]eason must view herself as the authoress of her principles” (IV: 448. 17), then extends this to the practical case, before claiming of a rational being that “its will can be a will of its own only under the idea of freedom” (IV: 448. 20-21).

The I is doing a lot of work here.[[92]](#footnote-92) Indeed, in Kant’s pre-critical lectures on metaphysics, he attempted a proof of the transcendental freedom of the soul on the basis of the I. He argues:

[…] the I proves that I myself act […] When I say: I think, I act, etc., then either the word I is applied falsely , or I am free. Were I not free, then I could not say: I do it, but rather I would have to say: I feel in me a desire to do, which someone has aroused in me. But when I say: I do it, that means spontaneity in the transcendental sense. (XXVIII: 268-9)[[93]](#footnote-93)

With the critical turn, Kant can no longer use this argument in quite the same way. The restrictions placed upon our knowledge in the first *Critique* stop Kant from inferring anything about a soul.[[94]](#footnote-94) However, the general point still holds: Self-consciousness involves spontaneity, “[w]hen I say: I think, I act, etc., then either the word I is applied falsely, or I am free” (XXVIII: 269).

One might object that, in helping himself to the I, Kant is taking too much for granted here. A related objection is that perhaps one’s judgements do receive guidance, but not self-consciously. Ameriks pushes this point against similar readings of *Groundwork* III:

This reading, however, does not explain why a rational will could not *allow* that in each case in which a justified judgment is made, there might *also* be, at the level of causes, a real determination of our mind by natural factors such as hidden impulses. (Ameriks 2003: 243):

I agree with Ameriks that this *might* be the case, but I think that this possibility lacks reasonable grounds. But I will not address this here. In chapter 3, I will offer a discussion about doubt, reason and agency, where I attempt to defuse such sceptical challenges.

In the meantime, I find myself in agreement with Matthew Boyle, when he writes (of a similar but related task) that:

My aim here is not to make a case that would persuade such skeptics, but to address those philosophers who take the sorts of familiar facts described here at face value. (Boyle 2011: 2n1)

For now, let us return once more to the Spontaneity Thesis:

[…] one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse. (IV: 448. 13-7)

Kant’s claim is that one cannot *self-consciously* receive guidance with regard to one’s judgements, because if one did, one would attribute the determination of one’s judgements to this guidance and not to one’s reason.

What makes the Spontaneity Thesis so difficult to unpack (but also so rewarding) is that there are two elements to it. It is for this reason that I have resisted substantially rephrasing it myself. And here, Beck is spot on:

It is often remarked by Kant-scholars that when, after dogged effort, they come to the root of some perplexing sentence or paragraph, they find that the complexities of the writing cannot always be attributed to the ineptitude of the writer, but more often to the demands of the thought expressed. They often realize that it would be very difficult, if not impossible, to say better exactly what Kant was saying and that any simplification of style would almost inevitably involve oversimplification of the thought. (Beck 1960: 4)[[95]](#footnote-95)

But enough about the complexity of the sentence at hand. We can now begin to unpack the two elements involved. The first concerns necessity of the possibility of self-attribution. This is the ‘I take’ bit, which involves self-consciousness and the possibility of imputation. The second element concerns the responsiveness to reasons involved in rational judgement. Kant is talking about the importance of the possibility of self-attribution specifically in connection to reason – “one cannot possibly think of a reason that […]” (IV: 448. 13). We have spent some time looking at the first of these elements, noting the important connection between self-consciousness and the ‘I think’, which we found in the B-deduction.

We have to be careful not to over-emphasise this element. There are two elements at play in the Spontaneity Thesis, and they can come apart. If we over-emphasise the importance of self-ascription, we run the risk of overplaying the ‘I take’ element, and losing the connection to reason. We risk ending up with a voluntaristic conception of spontaneity, where things are radically up to us, but we no longer have any grip on reasons (or the world).[[96]](#footnote-96) My coffee *is* cold, this computer *is* slow, this sentence *is* cumbersome, and I need to be responsive to all of this. I will return to consider this in more depth in the next chapter (§§1.4-1.5), where this dialectic (between freedom and reason-responsiveness) becomes crucial in ethics. For now however, let us turn to the second element of the Spontaneity Thesis, in order to illuminate the connection to reason.

At one point, Allison offers the following gloss on (the sentence that I have called) the Spontaneity Thesis:

Otherwise expressed, to make a cognitive judgment is to place oneself in the logical space of reasons and to hold oneself responsible to its norms. Consequently, a cognizer cannot regard itself as determined to judge in a certain way by factors external to the self-imposed norms (“from elsewhere”), since that would require locating the act of judgment in the logical space of causes and therefore as not governed by epistemic norms. (Allison 2011: 308)

The talk of norms is important. As noted in the Introduction (§2), in his later works, Kant ends up defending a moral argument for freedom, where (very roughly) our consciousness of the moral law reveals our freedom to us. We recognise that we ought to do something, and this reveals that we can. The general principle of ‘ought implies can’ is doing the work here. And this extends beyond the domain of morality. Prudential (and epistemic oughts) also involve spontaneity. We can see this in the first *Critique*. Here are two relevant passages:

Now that this reason has causality, or that we can at least represent something of the sort in it, is clear from the imperatives that we propose as rules to our powers of execution in everything practical. (A 547/B 575)

The reference to “everything practical” suggests that Kant is considering both moral and non-moral imperatives. This is made explicit shortly thereafter:

Whether it is an object of mere sensibility (the agreeable) or even of pure reason (the good), reason does not give in to those grounds which are empirically given, and it does not follow the order of things as they are presented in intuition, but with complete spontaneity it makes its own order according to ideas [...] (A 548/B 576)

There are practical norms, and we stand under them. We recognise that we ought to follow them, and from this, we infer that we can. Of course, this could quickly collapse into Kant’s moral argument for freedom, in a way that would run afoul of his circle – “presupposing the idea of freedom only for the sake of the moral law” (IV: 453. 5-7) – but we are considering more than just moral norms. We are also considering non-moral practical norms. And this gives us non-moral grounds for spontaneity. Something similar applies to epistemic norms.

Allison’s invocation of the space of reasons also brings the connection to reason into sight, and is worth exploring. For this, allow me to share a couple of key passages from Sellars’ seminal *Empiricism and the Philosophy of Mind*:[[97]](#footnote-97)

[…] to be the expression of knowledge, a report must not only *have* authority, this authority must *in some sense* be recognized by the person whose report it is. And this is a steep hurdle indeed. (Sellars 1956: 74; §34)

The essential point is that in characterising an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says. (Sellars 1956: 76; §36)

These passages are hard to unpack, but the basic thought is simple enough: to know something is to be able to be endorse it, to be able to justify it, or in Sellars’ iconic phrase, to be able to place it in the space of reasons.

As with Kant, there are lots of divergent readings of Sellars.[[98]](#footnote-98) In what follows, I focus on John McDowell’s reading, where spontaneity plays a central role.[[99]](#footnote-99) McDowell claims that:

[…] judging, making up our minds what to think, is something for which we are in principle responsible – something we freely do, as opposed to something that merely happens in our lives. (McDowell 2009: 6)

This emphasis on responsibility differs in letter from Kant’s Spontaneity Thesis, but the spirit of the passage is similar. McDowell notes the connection himself:

Sellars describes the logical space of reasons as the space of “justifying and being able to justify what one says”. We can see this as a distinctly twentieth-century elaboration of a Kantian conception, the conception of the capacity to exercise, paradigmatically in judgement, a freedom that is essentially a matter of responsiveness to reasons. (McDowell 2009: 6)

We need to clarify what exactly McDowell means when he says that “freedom is essentially a matter of responsiveness to reasons” (2009: 6). We can agree that to know something, or to make a rational judgement, requires being able to locate it within the space of reasons. But a further question remains, namely, *who* is doing this? Who is endorsing these claims, or locating them within the space of reasons? The answer is simple: a (self-conscious) subject. And McDowell makes this connection clear elsewhere:

It is incumbent on thought to be responsive to reasons recognized as such, and nothing can count as a reason for a thinking subject unless its authority as a reason can be freely acknowledged by the subject. (McDowell 2009: 90).

Alongside the importance of responsiveness to reasons, we also need to recognise the importance of a (self-conscious) subject.

Kant’s Spontaneity Thesis captures these two essential elements. Rational judgement requires a subject capable of locating their judgements within the space of reasons. Returning to my current predicament, we can see why it is not enough to reliably detect that this coffee is cold, this computer slow, and this sentence cumbersome. These claims must be able to be located within a space of reasons, and that requires a subject. In order for *me* to *know* these things, I must be able to locate them within the space of reasons; otherwise, these claims would amount to nothing for me.

As I noted at the outset, this is a tough puzzle to piece together. Through looking at one crucial sentence in *Groundwork* III – Kant’s Spontaneity Thesis (IV: 448. 13-7) – I have attempted to unpack some of the connections between reason, judgement, self-consciousness and spontaneity. I hope to have shown that rational judgement requires a subject capable of locating their judgments within the space of reasons. Or more simply, that reason involves spontaneity.

## 3.2 Practical Reason

We left the *Groundwork* at Kant’s claim about rational judgement, what I have called the Spontaneity Thesis:

Now, one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgements, since the subject would not then attribute the determination of judgement to his reason, but to an impulse. (IV: 448. 13-7)

For Kant, rational judgement involves spontaneity. And likewise, as rational, practical judgements will also involve spontaneity. Of course, we still face the question of whether our reason is practical, but that is the topic of chapter 3. Leaving this aside for now, we find Kant continue to claim, immediately after the preceding argument, that:[[100]](#footnote-100)

Reason must view herself as the authoress of her principles, independently of alien influences,[[101]](#footnote-101) and must consequently, as practical reason, or as the will of a rational being, by herself be viewed as free; i.e. its will can be a will of its own only under the idea of freedom (IV: 448. 17-21)

Rational judgement involves spontaneity. The same applies to practical judgement as well. If I judge that I ought to φ, I must be able to locate this within the space of reasons. Allison makes essentially the same point:

The point is that just as a cognizer must view itself as a self-determiner with respect to its judgments, so a rational agent must regard itself as determining itself to act on the basis of self-imposed principles and therefore as standing in the logical space of *practical* reasons. (Allison 2011: 308)

This also extends to maxims, which we can see by turning to Kant’s *Religion within the Boundaries of Mere Reason*.

For Kant, rational agents do not just act, they act on maxims, where a maxim is “the subjective principle for acting” (IV: 420n) or a “subjective principle of willing” (IV: 400n).[[102]](#footnote-102) In a famous passage in the *Religion*, Kant contends that:

[…] freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot bedetermined to action through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule forhimself, according to which he wills to conduct himself); only in this waycan an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom). (VI: 23. 3 – 24. 5)

Allison has termed this the Incorporation Thesis, and contends that it “underlies virtually everything that Kant has to say about rational agency” (1990: 40).[[103]](#footnote-103) The main idea is that, for a rational agent, an inclination only provides an incentive to act insofar as it is taken up on behalf of the agent, and incorporated into a maxim. If I incorporate an incentive into a maxim, I must be able to take the formation of this maxim as up to me, and be able to locate it within the space of reasons. Only in this way, can my acting upon an incentive be consistent with my rational agency.

The Incorporation Thesis is the application of the Spontaneity Thesis to the case of maxims. The Spontaneity Thesis shows us that rational judgement involves spontaneity, as it requires a subject capable of locating their judgments within the space of reasons. This applies to reason across the board: theoretical judgements involve this spontaneity, as do practical judgments, and rational agency in general. For Kant, rational agency involves the formation of maxims, and the Incorporation Thesis concerns the spontaneity involved in this.

Here is an important passage from Allison, illuminating the Incorporation Thesis in terms of the space of reasons:[[104]](#footnote-104)

[…] “justification goes all the way down.” In other words, one cannot appeal to a desire, no matter how strong, in justification of a course of action, and, indeed, for essentially the same reason as in the cognitive case: to do so is to remove oneself from the space of reasons; it is to take one’s behaviour as caused (and therefore as perhaps excused), but not as rationally justified. In that case, one is no longer acting under the idea of freedom, because one is no longer *acting*. (Allison 2012: 93)[[105]](#footnote-105)

To be rational is to be able to take my judgements, maxims and actions as up to me, and to be able to locate them within the space of reasons and provide justification for them.

In summary then, Kant sees rational judgements as involving spontaneity. By the same token, practical rational judgements will also involve spontaneity, or more generally, rational agency involves freedom.

## 3.3 The Metaphysics of Spontaneity

In this chapter, I hope to have outlined and offered a defence of my first premise:

1. A rational will is a free will

It is worth commenting on my wording here. The second section of *Groundwork* III explicitly states that: Freedom *must be presupposed* as a property of the will of all rational beings. And, as we saw earlier (§3.0), Kant claims that:

Now I say: every being that cannot act otherwise than *under the idea of freedom* is actually free, in a practical respect, precisely because of that; i.e. all laws that are inseparably bound up with freedom hold for it just as if its will had also been declared free in itself, and in a way that is valid in theoretical philosophy. (IV: 448. 4-9)

This is accompanied by a famous footnote:

I follow this route – of assuming freedom only, sufficient for our purpose, as made the foundation by human beings in their actions merely *in the idea* – so that I may not incur the obligation of proving freedom in its theoretical respect as well. For even if this latter point is left unsettled, the same laws that would bind a being that was actually free yet hold for a being that cannot act otherwise than under the idea of its own freedom. Here we can thus liberate ourselves from the burden that weights upon theory. (IV: 448n)

These passages are often read as downplaying the metaphysical significance of the connection between reason and freedom in Kant.[[106]](#footnote-106) Korsgaard for instance also offers a two premise summary of *Groundwork* III, but offers the following as the first premise:

A rational will must be regarded as a free will (Korsgaard 1996: 24a)

Allison provides another good example of this approach. He ends a paper on autonomy and spontaneity in the self as follows:

[…] Kant’s insistence on the connection between rational agency and spontaneity is to be understood as a conceptual claim rather than a putative metaphysical explanation. In other words, freedom in the sense of spontaneity is not something that we must add to our conception of ourselves as rational agents in order to make some metaphysical sense out of it; it is rather the defining feature of this very conception. This is the force of the claim that we must act under the idea of freedom. […] for Kant this remained a mere conceptual claim regarding the manner in which we must conceive ourselves insofar as we take ourselves as rational agents, whereas in his idealistic successors it became inflated into an ontological truth. Whether this conception is necessary and, if so, whether its inflation is warranted are, of course, among the big questions posed for us by German idealism. (Allison 1996: 142)

In the previous two sections, I have agreed with much of Allison’s analysis. Here, I want to depart from him and Korsgaard to defend a stronger connection between reason and freedom.

Before I do so, it is worth briefly noting that parts of the second subsection of *Groundwork* III itself suggest a strong connection. Most importantly, Kant talks of a *proof* [*Beweis*] of freedom. He says that “freedom must also be proved as a property of the will of all rational beings” (IV: 447. 33-4), and that “one must prove it as belonging to the activity of rational beings endowed with a will as such” (IV: 448. 2-4).[[107]](#footnote-107)

In what follows, I argue that the connection we have found (in §§3.1-3.2) amounts to a proof rather than a presupposition, and consider the metaphysical significance of this. We have just seen Allison insist against there being any metaphysical implications of this connection. This looks like a limited conception of metaphysics. If spontaneity is the defining feature of reason, that seems like a metaphysical truth. Perhaps Allison is equating metaphysics with claims about the noumenal – bad metaphysics, if you will. But rather than engage in a debate about the term ‘metaphysics’[[108]](#footnote-108), I want to claim that freedom is more than just a presupposition. Allison ends the above passage by asking whether the inflation of the necessity of spontaneity into an ontological truth is warranted. I think that it is. I also think that Allison (and Kant) need it to be, but I will return to press this point in chapter 3 (§2.2).

Allow me to share some passages that set the scene. Allison’s treatment of the spontaneity involved in rational judgement is in line with ours. He argues that in order for ‘A is B’ to amount to a rational judgement for me, I have to be able to take A as B. Moreover, he claims that:

[...] the activity of “taking as” is constitutive of judgment. (Allison 1990: 37)

In another paper, when considering the spontaneity involved in judgment, he claims that:[[109]](#footnote-109)

[…] the “must” here is conceptual, for unless I am aware of taking x as F […], I have not in fact taken it as such. In short, conceptual recognition […] is an inherently reflective and, therefore, self-conscious act. (Allison 1996: 62)

This aligns nicely with our treatment of the link between reason and spontaneity. Allison goes on however, to insist that:

It is just at this point, however, that the Kantian position is most open to misunderstanding. Thus, contrary to what one might suppose, the “critical” Kant expressly refrains from drawing any ontological conclusions regarding the absolute spontaneity of the self from his epistemological analysis. The spontaneity-claim concerns rather the way in which the thinking subject must be conceived (or conceive itself) *qua* engaged in cognition. In other words, spontaneity functions in the technical Kantian sense as an idea in light of which the act of thinking must be conceived in order to retain its normative status, but this does not license a metaphysical inference to the absolute spontaneity of the “thing which thinks”. (Allison 1996: 64)

Allison insists that I must conceive myself as spontaneous insofar as I take myself to be rational, but that is not quite right. The connection is much simpler: I am rational, and thereby spontaneous. Our previous analysis established a link between reason and spontaneity, and we should not shy away from that.[[110]](#footnote-110) As rational, I am spontaneous.

Allison and Korsgaard sometimes contrast Kant’s claim that we must act under the idea of freedom with the claim that we must *believe* that we are free, and deny this latter claim.[[111]](#footnote-111) Allison argues that this latter claim would be refuted by the mere existence of a single (hard) determinist or fatalist (2012: 113). This is an unsatisfyingly narrow conception of belief, but their point seems to be that you do not have to explicitly think that you are free, and that seems fair. However, the main issue is whether Kant’s discussion provides grounds for believing that we are spontaneous. And it does. Spontaneity is constitutive of rational judgment. We are justified in believing that we make rational judgments, and are thereby also justified in believing that we are spontaneous. Of course, there are the further questions of whether we are justified in extending this to the practical case, and whether this conception of freedom is sufficient for Kant’s conception of morality, but those are the topics of chapters 3 and 4.

Now, we can finally return to our objection. Earlier on in this chapter, we considered the objection that acting under the idea of freedom, although necessary, might not be justified (§1.2). We considered two varieties of inescapability (brute and psychological), and concluded that neither of them offered justifications, but only exculpations of our acting under the idea of freedom.

Since then, we have seen that reason involves spontaneity. We can now clarify the nature of the ‘must’ in Kant’s claim that we must act under the idea of freedom. Here is Allison, one more time:

[...] the activity of “taking as” is constitutive of judgment. (Allison 1990: 37)

[…] the “must” here is conceptual, for unless I am aware of taking x as F […], I have not in fact taken it as such. In short, conceptual recognition […] is an inherently reflective and, therefore, self-conscious act. (Allison 1996: 62)

Allison is right. The ‘must’ is conceptual, but it can also be distracting (which we saw in the failure of the self-stultification argument (§1.2)).What matters is that reason involves spontaneity, not how we must conceive ourselves. Spontaneity is constitutive of rational judgment. We are rational, and we are thereby also spontaneous.

Before I continue, I want to pre-empt a possible objection. I have just claimed that we have grounds to believe that we are spontaneous, and that we should not shy away from this. One worry is that this might conflict with Kant’s transcendental idealism, and the strict limits it places on making any claims that go beyond the bounds of possible experience. I have not said much about transcendental idealism in this thesis so far, but I will do so in chapter 3, where I distance myself from it.

# Conclusion

In this chapter, I have considered two parts of *Groundwork* III that attempt to link reason and freedom. Moreover, I have argued that one of these works. I locate Kant’s key insight in a sentence in the second subsection of *Groundwork* III (IV: 448. 13-7), which I have called the ‘Spontaneity Thesis’. The basic idea is that to be rational is to be able to locate one’s own judgements within the space of reasons. Put succinctly, reason involves spontaneity. Likewise, rational agency also involves spontaneity, or freedom – a rational will is a free will. Of course, I have yet to establish that we are rational beings *with wills*, but that is the topic of chapter 3. Before we get to that, I want to explore the connection between freedom and morality in Kant. This is the topic of the next chapter.

2 – Freedom, Autonomy and Morality

The second premise of my argument is that:

1. A free will stands under the moral law

In this chapter I attempt to unpack this claim, exploring the relationship between freedom, autonomy and morality in Kant.

I will consider an ambitious and a modest account of this relationship. On the modest account, freedom enables us to act morally. The ambitious account goes beyond this, proposing that freedom provides something like a proof of morality. The exact nature of the difference between these two accounts will emerge over the course of this chapter.

I split the chapter up into two parts. In the first (§1), I consider the ambitious account, which I associate with Kant’s (meta-ethical) formalism and Kantian constructivism. I wish to distance myself from all of this. In doing so, I depart from the argument in the opening section of *Groundwork* III. I reject Kant’s argument, but accept its conclusion, namely that a free will stands under the moral law. In the second half of this chapter (§2), I lay out my modest account of this claim.

# Freedom, Autonomy and Morality: The Ambitious Account

As with the previous chapter, it helps to start with the text itself. *Groundwork* III begins with the section: *The concept of freedom is the key to the explanation of the autonomy of the will*. In this, Kant details the connection between freedom, autonomy and morality as follows:[[112]](#footnote-112)

A *will* is a kind of causality of living beings in so far as they are rational, and *freedom* would be that property of such a causality, as it can be efficient independently of alien causes *determining* it; just as *natural necessity* is the property of the causality of all nonrational beings to be determined to activity by the influence of alien causes.

The explication of freedom stated above is *negative* and therefore unfruitful for gaining insight into its essence; but there flows from it a *positive* concept of freedom, which is so much the richer and more fruitful. Since the concept of causality carries with it that of *laws* according to which, by something that we call a cause, something else, namely the consequence, must be posited: freedom, though it is not a property of the will according to natural laws, is not lawless because of that at all, but must rather be a causality according to immutable laws, but of a special kind; for otherwise a free will would be an absurdity. Natural necessity was a heteronomy of efficient causes; for every effect was possible only according to the law that something else determines the efficient cause to causality; what else, then, can freedom of the will be, but autonomy, i.e. the property of the will of being a law to itself? But the proposition: the will is in all actions a law to itself, designates only the principle of acting on no maxim other than that which can also have itself as its object as universal law. But this is just the formula of the categorical imperative and the principle of morality: thus a free will and a will under moral laws are one and the same.

Thus if freedom of the will is presupposed, morality along with its principle follows from it, by mere analysis of its concept. (IV: 446. 7 – 447. 9)

There are four main steps to this. Kant argues that:

1. A free will is independent of determination from alien causes (IV: 446. 7-10).
2. Nevertheless, a free will must have some law determining it (IV: 446. 15-21).
3. As independent from determination from alien causes (1), the law that determines the free will (2) must therefore be the will’s own law, that is, the will must be autonomous (IV: 446. 24 – 447. 2).
4. The claim that the will is a law to itself is just the same as the claim that we are to act on universalisable maxims, which is the supreme principle of morality (IV: 447. 2-7).

As is often the case in *Groundwork* III, Kant moves fast. We need to unpack these four steps.

The first two steps go together, and lay out Kant’s conception of a rational free will.

1. A free will is independent of determination from alien causes (IV: 446. 7-10).
2. Nevertheless, a free will must have some law determining it (IV: 446. 15-21).

In the preceding chapter we saw that there are two elements to rational judgement: spontaneity, and responsiveness to reasons. And we can read Kant as making a related claim here: It is not enough for a free will to be independent from alien causes, it must also have some law determining it.

One immediate complication though involves what Kant means by a ‘law’ in the claim that a free will must have some law determining it. In the passage, he claims that a “*will* is a kind of causality of living beings in so far as they are rational” (IV: 446. 7-8), and that “the concept of causality carries with it that of *laws*” (IV: 446. 15-6). He then draws an analogy with “*natural necessity* […] the property of the causality of all nonrational beings to be determined to activity by the influence of alien causes” (IV. 446. 10-2), but notes that a free will, while “not lawless […] must rather be a causality according to immutable laws, but of a special kind” (IV. 446. 19-21).

We can read Kant’s talk of ‘laws’ in either a broad or a narrow sense (and the text pulls us in both directions).[[113]](#footnote-113) In the narrow sense, the law is an actual practical *law* that determines the will. Read this way, the analogy with natural necessity is strong – nonrational beings are causally determined by the laws of nature, whereas rational beings are determined by the moral law. Timmermann opts for this reading, claiming that:[[114]](#footnote-114)

Kant conceives of ‘laws of freedom’ and ‘laws of nature’ as parallel descriptive laws that govern separate realms. They concern two kinds of metaphysical doctrine. Freedom and natural necessity are qualities of types of causality that operate in different causal spheres. (Timermann 2007: 122)

I find this reading problematic.

My main worry is that it conceives of the relationship between a rational will and the moral law in the wrong kind of way, as it views the rational will as (in some sense) causally determined by the moral law. This opens Kant up to Reinhold’s objection, and misses what is distinctive about the moral law, namely its normativity. This is a complicated issue, and one I will return to in the second half of this chapter (§2.1), where I outline what it means to be *under* the moral law; there I will argue that the moral law is not merely descriptive, but also essentially normative and prescriptive (both for us, and a perfectly rational will).

Alongside this worry, the narrow reading of ‘law’ does not quite fit with the argumentative structure of the passage. In the first two steps of the passage, Kant briefly lays out his conception of rational agency. He then hopes to move from this, through autonomy (step 3), to the moral law (step 4). If we read ‘law’ in a narrow sense, the argument from rational agency to autonomy to the moral law is over and done with before it has begun, and more importantly, unconvincing – Kant has given us no reason why a rational will must act on a *law* (in the strong sense).

We can however, read Kant’s talk of ‘laws’ in a broader sense, as maxims or principles of action. I opt for this reading. In doing so, I lose the tightness of the analogy with causality and natural necessity, but the gains make up for this. Alongside the issue of the normativity of the law, it also allows us to make sense of Kant’s claim that a rational will must act on a law. As we saw in the previous chapter (§3.2), Kant conceives of rational agency as involving maxims; rational agents do not just act, they act on principles.[[115]](#footnote-115) This follows from Kant’s conception of rational agency, but the claim that a free being must act on *a law* (in the strong sense) does not.

Before we continue, I should note that we have run up against a larger interpretive issue, and one that we have already touched upon. As we saw in the introduction (§2), there are moral and non-moral conceptions of freedom in Kant, and these roughly match up with the strong and weak readings of ‘law’ in the claim that a free will must have a law determining it:[[116]](#footnote-116) if our freedom is exclusively moral, then it consists in the ability to follow the moral law – the moral law is the *law* of freedom; if our freedom is also non-moral, then it also applies to maxims in general. In this thesis, I am interested in the latter option. Let us accordingly leave aside the exclusively moral reading of freedom to move on to the next step.

Things begin to heat up with the third step, where Kant claims that a free will must be autonomous:

1. As independent from determination from alien causes (1), the law that determines the free will (2) must therefore be the will’s own law, that is, the will must be autonomous (IV: 446. 24 – 447. 2).

This needs considerable unpacking, and in what follows, I do just this. I look at Allison’s (§1.1) and Korsgaard’s (§1.2) treatments of Kant’s claim, before introducing a crucial ambiguity in Kant’s conception of autonomy (§1.3). With this ambiguity in plain sight, I distance myself from Kant’s (meta-ethical) formalism (§1.4) and Kantian constructivism (§1.5). I conclude that the attempt to move from freedom to autonomy to the moral law in the first section of *Groundwork* III fails. I then lay aside this ambitious account of the relationship between freedom, autonomy and morality to offer a modest alternative (§2).

## 1.1 Allison’s Reciprocity Thesis

In several places, Kant contends that freedom and the moral law are reciprocal concepts.[[117]](#footnote-117) Allison has offered perhaps the canonical treatment of this relationship, which he calls the ‘Reciprocity Thesis’.[[118]](#footnote-118) He primarily focuses on our third step. He claims that the move from freedom to autonomy works, but only because Kant is operating with a radical conception of freedom – complete independence from determination by external conditions.[[119]](#footnote-119) This is *transcendental* freedom, and it involves “an independence of […] reason itself […] from all determining causes of the world of sense” (A803/B831). The thought is that insofar as a free will is completely independent from determination by external conditions, it cannot be determined by anything external to it, and thus must provide its own law.[[120]](#footnote-120)

Allison insists that transcendental freedom is the key to this move. He argues that:

[…] without the assumption of such freedom […] a free choice based on self-interest, happiness, or some such putatively ultimate yet non-moral end or motivational ground could be justified by an appeal to ‘human nature’ or some given determinant of behaviour. (The details are irrelevant to the argument.) (Allison 1990: 208)

The assumption of transcendental freedom blocks such moves:

Put simply, if self-preservation, self-interest, or happiness is the principle of my behaviour, if it dictates my maxims, it is I (not nature in me) that gives it this authority. At least this is the case under the presupposition that I am free in the transcendental sense. (Allison 1990: 208)

The Reciprocity Thesis posits a reciprocal relationship between *transcendental* freedom and the moral law.

We can plug this into our argument:

1\*) A *transcendentally* free will is *totally* independent of determination from alien causes

2\*) Nevertheless, a *transcendentally* free will must have some law determining it

3\*) As *totally* independent from determination from alien causes (1\*), the law that determines the *transcendentally* free will (2\*) must therefore be the will’s own law, that is, the will must be autonomous.

I will return to criticise this shortly (§§1.3-1.5). Before that, it is worth taking a look at Korsgaard’s treatment of the relationship between freedom and autonomy.

## 1.2 Korsgaard’s Regress

Kant’s move from freedom to autonomy serves a central role in Korsgaard’s work. In what follows, I detail this move as it occurs in *The Sources of Normativity* (one of her longer treatments of it).[[121]](#footnote-121)

Korsgaard sets out the issue in terms of how we can achieve reflective success, or act on reasons. Say we decide to act on a certain desire. We then face the question of the normative status of this desire. In appealing to another (perhaps more fundamental) desire or principle of action in order to justify the initial desire, we in turn face the question of the normative status of this new desire or principle – and thus, Korsgaard claims, “the usual regress threatens” (1996b: 97).[[122]](#footnote-122) We face the challenge as to what could bring such a course of reflection to a successful end.

Korsgaard then employs variants of Kant’s argument in the first section of *Groundwork* III, running through steps (2) and (3):

[W]e may say that since the will is practical reason, it cannot be conceived as acting and choosing for no reason. Since reasons are derived from principles, the free will must have a principle. But because the will is free, no law or principle can be imposed on it from outside. Kant concludes that the will must be autonomous: that is, it must have its own law or principle. (Korsgaard 1996b: 97-98)

She then asks what such a principle could be. Her answer is Kant’s, and it parallels steps (3) and (4):[[123]](#footnote-123)

The problem faced by the free will is this: the will must have a law, but because the will is free, it must be its own law. And nothing determines what that law must be. *All that it has to be is a law*. Now consider the content of the categorical imperative, as represented by the Formula of Universal Law. The categorical imperative merely tells us to choose a law. Its only constraint on our choice is that it has the form of a law. And nothing determines what the law must be. *All that it has to be is a law* … Therefore the categorical imperative is the law of a free will. (Korsgaard 1996b: 98)

For Korsgaard, a free will is thus autonomous.[[124]](#footnote-124)

I should mention that Korsgaard thinks this argument establishes that we are autonomous, but not that we stand under the moral law. Korsgaard thinks that Kant gets us to something like step (3), but not to step (4). In order to complete this final step, she invokes a notion of practical identity, and employs a transcendental argument, which she sees in Kant’s argument for the formula of humanity (Korsgaard 1996b: 100-102). Korsgaard’s novel move from autonomy to morality has (deservedly) received much attention, but it is not our concern here (I offer an alternative account of Kant’s argument later in this chapter – §2.2).[[125]](#footnote-125) We shall continue to investigate the relationship between freedom and autonomy.

## 1.3 Ambiguities of Autonomy

We have just seen three accounts of the claim that a free will is autonomous – Kant’s, Allison’s and Korsgaard’s. Now is an appropriate time to introduce a crucial ambiguity. There are several different meanings of the term ‘autonomy’ in Kant, and I want to draw attention to three of them.[[126]](#footnote-126)

The first is a meta-ethical position. Autonomy in this sense is contrasted with heteronomy and concerns what gives normative force to our principles, whether it is something internal or external to our will.[[127]](#footnote-127) (I will return to expand upon this shortly).

The second type of autonomy is a property of the will. And here, we can understand this in a weak or strong sense. In the weak sense, ‘autonomy’ has a familiar meaning, concerning rational agency in general – insofar as one is able to act on the basis of maxims, and choose and deliberate for oneself, one is autonomous. The strong sense of ‘autonomy’ points to the capacity for one’s maxims to conform with the moral law – insofar as one’s maxims can conform with the moral law, one is autonomous. This distinction corresponds to the two ways of reading Kant’s talk of ‘laws’, which should come as no surprise.[[128]](#footnote-128) ‘Autonomy’ means self-legislation, and the two different ways of thinking about laws leads to two different ways of conceiving self-*legislation* and autonomy.

Finally, the third type of autonomy is a principle of morality, namely, the formula of autonomy itself:

[…] to choose only in such a way that the maxims of your choice are also included as universal law in the same volition (IV: 440. 18-20).

We now have three different types of autonomy. The first is meta-ethical, the second concerns agency, and the third is an ethical principle.

Kant is not just clumsy here; he sees these three types of autonomy as connected. In *Groundwork* III, he tells us that “freedom and the will's own legislation are both autonomy, and hence reciprocal concepts” (IV: 450. 23-4), and also that “[w]ith the idea of freedom the concept of *autonomy* is now inseparably combined and with the concept of autonomy the universal principle of morality” (IV: 452. 35-7). In a key subsection in *Groundwork* II – *The autonomy of the will as the supreme principle of morality* – he connects them as follows:

Autonomy of the will is the characteristic of the will by which it is a law to itself (independently of any characteristic of the objects of willing). The principle of autonomy is thus: not to choose in any other way than that the maxims of one’s choice are also comprised as universal law in the same willing. […] that the envisaged principle of autonomy of autonomy is the sole principle of moral science can very well be established by mere analysis of the concepts of morality. For thereby we find that its principle must be a categorical imperative, and that it commands neither more nor less than just this autonomy. (IV: 440. 16-32)

Allison’s Reciprocity Thesis latches on to these connections, and attempts to link agency with the principle of morality, through the meta-ethical conception of autonomy. Allison contends that, because we are (transcendentally) free, no object of the will could serve to justify our principles of action, that is, (meta-ethical) heteronomy is insufficient. And given this, we are “to choose only in such a way that the maxims of your choice are also included as universal law in the same volition” (IV: 440. 18-20). There are thus three different ways in which Kant’s concept of autonomy comes apart, but these also connect back together.

I want to challenge this reunion. The claim under investigation is that a (transcendentally) free will is autonomous. It is crucial to clarify what is meant by ‘autonomy’ in this claim. Kant appears to be talking about autonomy as a property of the will. And so, we have two options: it is either a property of the will in the weak or strong sense. I will consider these options in turn.

If we understand autonomy in the third step as a property of the will in the weak sense, then the claim that a (transcendentally) free will is autonomous holds, but loses its significance. Understood in this sense, it would just amount to a restatement of the result of the last chapter – that rational agency involves maxims and freedom. This is not enough, as the type of autonomy that (3) establishes needs to match up to the fourth step of Kant’s argument:[[129]](#footnote-129)

1. The claim that the will is a law to itself is just the same as the claim that we are to act on universalisable maxims, which is the supreme principle of morality (IV: 447. 2-7).

It is worth setting out all four steps of the argument to see how this plays out.

1. A free will is independent of determination from alien causes (IV: 446. 7-10).
2. Nevertheless, a free will must have some law determining it (IV: 446. 15-21).
3. As independent from determination from alien causes (1), the law that determines the free will (2) must therefore be the will’s own law, that is, the will must be autonomous (IV: 446. 24 – 447. 2).
4. The claim that the will is a law to itself is just the same as the claim that we are to act on universalisable maxims, which is the supreme principle of morality (IV: 447. 2-7).

Kant looks to move from our agency (1 and 2) to autonomy (3) to morality (4). If autonomy in (3) is understood in a weak sense – as being a rational agent who is able to act on maxims – it follows from Kant’s conception of agency (1 and 2), but does not look like it will get us all the way to the moral law (4).

The strong sense of autonomy fares no better. It allows us to comfortably move from autonomy (3) to the moral law (4), but then it is not clear how this strong sense of autonomy follows from the considerations of agency (1 and 2).

Autonomy stands between rational agency and the moral law, and Kant struggles to connect these up.[[130]](#footnote-130) I should note that, in the final chapter of this thesis, I will defend a version of Kant’s position, attempting to disarm the threat of a gap between rational agency and the moral law. However, I will not defend the ambitious account of the relationship between reason, freedom and the moral law, but instead my own modest account. For that, we will have to wait until the second half of this chapter, and so, for now, I continue to discuss the shortcomings of the ambitious account.

Kant needs to connect rational agency and the moral law. One route, which we have just seen Allison take, is through the meta-ethical conception of autonomy. I am going to argue that this route fails. And I take this to have important consequences. I think it counts against both Kant’s (meta-ethical) formalism and Kantian constructivism. But I am getting ahead of myself again. We need to work our way through Kant’s attempt to move from freedom to autonomy. In doing so, the failure of this move and its significance will emerge.

Here is Allison’s reconstruction of Kant’s move from freedom to the moral law:

1) As a “kind of causality” the will must, in some sense, be law governed or, in the language of the second Critique, “determinable” according to some law (a lawless will is an absurdity); (2) as free, it cannot be governed by laws of nature; (3) it must therefore be governed by laws of a different sort, namely, self-imposed ones; and (4) the moral law is the required self-imposed law. (Allison 1990: 203)

Once more, the third step is crucial. And again, everything hangs on how we read ‘self-imposed’ in the claim at hand. If we understand it in a weak sense, then we are talking about maxims; as we saw in the last chapter (§3.2), they are up to us, and in this sense self-imposed. That is fine, but then it is no longer clear why the moral law (4) follows from this. The other option is to understand ‘self-imposed’ in a stronger sense, as independent of all objects of the will. In this case, the moral law (4) follows from self-imposition (3), but this strong sense of self-imposition (3) no longer clearly follows from considerations of agency (1 and 2). Let me elaborate on this last point.

Allison claims that, because a (transcendentally) free will cannot be governed by laws of nature, it must therefore be governed by self-imposed laws (in a strong sense). But there is a neglected alternative here, namely, laws that are neither of nature nor self-imposed. Now perhaps there are no such laws. But if there were, they seem suitable candidates for the principles of a transcendentally free will. Indeed, at one juncture, Allison himself suggests something similar:

[T]he idea that the existence of something that is an end in itself is a sufficient condition of a categorical imperative seems relatively unproblematic, since, arguably, entities with this status (if they exist) could be the source of unconditioned commands (Allison 2011: 205).

This seems unproblematic, and yet it conflicts with what Allison takes to be the central claim of the *Groundwork*, namely that only a self-imposed law could be unconditionally normative (and thereby suitable for a transcendentally free will). This claim serves as the centrepiece of his recent commentary.[[131]](#footnote-131) Unfortunately, it appears to neglect an alternative, namely an unconditionally normative claim that is not self-imposed.

It is worth noting that Kant also seems guilty of this. In step (3), he asks:[[132]](#footnote-132)

[…] what, then, can freedom of the will be other than autonomy, that is, the will’s property of being a law to itself? (IV: 446. 25 – 446. 2)

And later on in *Groundwork* III, he notes that: “With the idea of freedom the concept of *autonomy* is now inseparably combined and with the concept of autonomy the universal principle of morality” (IV: 452. 35-7), and also talks of “our inference from freedom to autonomy and from the latter to the moral law” (IV: 453. 4-5).

I have moved quickly over this point, which touches upon both Kant’s formalism and Kantian constructivism. I want to now say something about each of these. In doing so, the failure of Kant’s move from freedom to autonomy – and with it, the ambitious account of the relationship between freedom and the moral law – should become clearer.

## 1.4 Kant’s Meta-Ethical Formalism

Kant often claims that the moral law must be merely formal, that it must abstract from any matter, end, or object of the will.[[133]](#footnote-133) Unless it does so, he argues, only hypothetical imperatives will result.

This *meta-ethical* formalism is my target. There is another version of formalism that I leave aside, namely a first-order ethical or political formalism, where the Formula of Universal Law (for example) is seen as an excellent decision procedure. The main focus of this section is the meta-ethical position that there is no moral content independently of (or prior to) a merely formal principle. For this, we turn once more to the text.

In an important passage in *Groundwork* II, Kant tells us that:

Wherever an object of the will has to be made the foundation for prescribing the rule that determines it, there the rule is nothing other than heteronomy; the imperative is conditional, namely: *if* or *because* one wills this object, one ought to act in such a way; hence it can never command morally, i.e. categorically. (IV: 444. 1-5)

He concludes that:

An absolutely good will, whose principle must be a categorical imperative, will therefore, indeterminate with regard to all objects, contain merely *the form of willing* as such, and indeed as autonomy (IV: 444. 28-30)

Here, we see Kant explicitly state that the moral law must be indeterminate with regard to all objects, and instead contain merely the form of willing.[[134]](#footnote-134) He contends that otherwise, only hypothetical imperatives will result. This is especially clear in the section *The heteronomy of the will as the source of all spurious principle of morality:*

If it is *in anything other* than the fitness of its maxims for its own universal legislation, hence if – as it goes beyond itself – it is a characteristic of any of its objects that the will seeks the law that is to determine it, the outcome is always *heteronomy*. Then the will does not give itself the law, but the object by its relation to the will gives the law to it. This relation, whether it rests on inclination, or on representations of reason, makes possible hypothetical imperatives only: I ought to do something *because I want something else*. By contrast the moral and hence categorical imperative says: I ought to act in such or such a way, even if I did not want anything else. (IV: 441. 1-13)

Kant runs together two distinct lines of thought here, which correspond to two different ways of reading the distinction between hypothetical and categorical imperatives.[[135]](#footnote-135) The first concerns whether the normative force of a principle is conditional or not. On this reading, the normative force of hypothetical imperatives depends upon whether one wants the end in question, whereas the normative force of categorical imperatives does not. Again:

[…] hypothetical imperatives […say]: I ought to do something *because I want something else*. By contrast the moral and hence categorical imperative says: I ought to act in such or such a way, even if I did not want anything else. (IV: 441. 10-13)

This is an excellent distinction, and I am happy to follow Kant in accepting it (as I did in the Introduction – §1.1).

Unfortunately, another distinction accompanies it. At times, Kant claims that only *merely formal* practical principles can command unconditionally. This is his (meta-ethical) formalism. Otherwise expressed, the first distinction concerns whether the normative force of a practical principle depends upon *any object of desire*, and the second *any object whatsoever*. I follow Kant in the former distinction, but reject the latter.

Kant neglects an alternative. He writes as if there are only two options – a merely formal principle, or a conditionally normative principle – but this is mistaken.[[136]](#footnote-136) Morality does not have to be formal to be unconditionally normative. Quite simply, if there were any objects of unconditional worth, then they could ground unconditionally normative claims. Not only is this alternative plausible, but it also finds expression in parts of Kant’s moral philosophy.

Perhaps the most obvious example occurs in his introduction of the formula of humanity:[[137]](#footnote-137)

But suppose there were something *the existence of which in itself* has an absolute worth, that, as an *end in itself*, could be a ground of determinate laws, then the ground of a possible categorical imperative, i.e. of a practical law, would lie in it, and only in it alone.

Now I say: a human being and generally every rational being exists as an end in itself, *not merely as a means* for the discretionary use for this or that will, but must in all its actions, whether directed towards itself or also to other rational beings, always be considered *at the same time as an end*. (IV: 428. 3-11)

In this important passage, Kant explicitly makes reference to an end of absolute worth, which serves as the ground of the moral law.[[138]](#footnote-138)

Of course, others have offered formalist readings of this passage, and perhaps these can be made to fit the text, but I want to circumvent this exegetical debate here (I will return to say something about the plausibility of these positions later in this chapter – §2.2 ). For my purposes, what matters is that, taken at face-value, Kant’s discussion of the formula of humanity counts against his (meta-ethical) formalism – an end of absolute worth could (and maybe even does) serve as the ground of an unconditionally normative claim.

At times, Kant seems attuned to this. In the run up to his discussion of the formula of humanity, he introduces a subset of ends, purposes and objects of the will, before claiming that:

Practical principles are *formal* if they abstract from all subjective ends; they are *material* if they have these, and hence certain incentives, at their foundation. (IV: 427. 30-2)

If subjective ends amount to things that we want, then this takes us back to our first way of making out the distinction between hypothetical and categorical imperatives: a practical principle that derives its normative force from some subjective end – something that I might or might not want – is a hypothetical imperative.[[139]](#footnote-139) But this does not count in favour of Kant’s claim elsewhere that no object could serve as the determining ground of the will, or that the moral law must be merely formal. If there were objective ends, then these could be the ground of unconditionally normative claims.

This is a crucial point against Kant’s (meta-ethical) formalism. Allow me to run through it again, this time with reference to Kant’s treatment of the relationship between freedom and morality in the second *Critique*.

In the *Critique of Practical Reason,* Kant discusses this in terms of the following problem: “Supposing that a will is free: to find the law that alone is suitable for determining it necessarily” (V: 29. 12-3). His answer is that:

Since the matter of a practical law, i.e., the object of a maxim, can never be given except empirically, but a free will – as independent of empirical conditions (i.e., conditions belonging to the world of sense) – must nonetheless be determinable, a free will must, independently of the *matter* of the law, nonetheless find a determining basis in the law. But the law, apart from its matter, contains nothing more than the legislative form. Therefore solely the legislative form, insofar as it is contained in the maxim, can amount to a determining basis of the will. (V: 29. 14-22)

He concludes: “Thus freedom and unconditional practical law reciprocally refer to each other” (V: 29. 24-5).

As with *Groundwork* III, Kant begins with the assumption that a free will must have a law determining it. He then claims that the matter of a practical law (which here, he equates with a maxim) can only be given empirically. He continues to argue that a free will is independent of empirical conditions, and thereby all matter of the law, which leaves nothing left to determine it, but the mere legislative form.

This argument suffers the same faults as the last one. Once more, we can bring this into view by considering an end that appears to be of absolute worth – rational nature. There is no good reason why a free will must abstract from such an end (or matter). Of course, by ‘matter’, Kant could just mean things that we happen to desire, but then the claim that once we eliminate this option only the mere legislative form of the will remains, does not follow. Again, Kant neglects an alternative – objects of unconditional worth.

I will return to say more about objects of unconditional worth in the next section, when we debate the objections to (Kantian) moral realism. I have made an initial case that Kant neglects an alternative. In the next section (§1.5), we will consider the various ways in which one might defend Kant’s (meta-ethical) formalism against this charge. For now though, we can return to Kant’s claim that a free will must be autonomous. Allison’s gloss on this is that, if the will is transcendentally free – independent from everything external to it – it cannot be determined by anything external to it, and thus must provide its own law. But much like our treatment of spontaneity in chapter 1 (§3.1), we need to be careful to not attribute a radically voluntaristic conception of freedom to Kant.

McDowell is excellent on this point. He repeatedly emphasises that recognising reasons does not involve surrendering one’s freedom or autonomy. Here is McDowell himself:

Emancipating our intellectual activity from the pseudo-authority of the merely given cannot be freeing ourselves from the authority of reasons. (McDowell 2009: 104)

In a lengthier passage, he elaborates on this:[[140]](#footnote-140)

Cashing out the idea of rational self-determination in terms of an image of legislating for oneself goes back to Kant’s appropriation of Rousseau. The image is fine, but we need to be careful about what it can amount to. When we say that rational norms are self-legislated by rational subjects, the point is that acknowledging the authority of those norms is not abandoning control of one’s life to an alien power. But that is consistent with – and indeed requires – that we not pretend to make sense of the idea of a legislative act that confers authority on the norms of reason. If the legislative act is not already subject to the norms of reason, how can it be anything but arbitrary? But nothing instituted by an act that is arbitrary could be intelligible as the authority of reason. If self-legislation of rational norms is not to be a random leap in the dark, it must be seen as an acknowledgment of an authority that the norms have anyway. Submitting to that authority is not handing over control of the relevant areas of one’s life to a foreign power. What controls one’s life is still in oneself, in whatever it is about one that enables one to recognize that the norms are authoritative. But their authority is not a creature of one’s recognition. Seen in this light, the self-determination idea is a version of the basic commitment to rationalism. (McDowell 2009: 105)

McDowell draws inspiration from Hegel here, who makes a similar point against Kant.

In the introduction to the *Philosophy of Right*, Hegel discusses two elements of the will: the universal and the particular.[[141]](#footnote-141) The universal element involves the ability to abstract from any and all determinate content, whereas the particular element of the will involves its determinacy. For Hegel, each of these elements presents one side of the will, and they only find their truth together. What exactly this unity consists in is difficult – a difficulty Hegel acknowledges.[[142]](#footnote-142) However, his criticism of Kant is fairly clear. For this, we need only to look at the first element of the will – its universality.

The universality of the will concerns the ability to abstract from everything particular. Hegel characterises this as follows:

The will contains (α) the element of pure indeterminacy of of the ‘I’s pure reflection into itself, in which every limitation, every content, whether present immediately through nature, through needs, desires, and drives, or given and determined in some other way, is dissolved; this is the limitless infinity of absolute abstraction or universality, the pure thinking of oneself. (PR: §5, p.37)

It is inherent in this element of the will that I am able to free myself from everything, to renounce all ends, and to abstract from everything. The human being alone is able to abandon all things, even his own life: he can commit suicide. The animal cannot do this; it always remains only negative, in a determination which is alien to it and to which it merely grows accustomed. The human being is pure thinking of himself, and only in thinking is he this power to give himself universality, that is to extinguish all particularity, all determinacy. (PR: §5Z, p.38)

Hegel recognises the importance of this element of the will, but insists that it should not be over-emphasised:[[143]](#footnote-143)

Only one aspect of the will is defined here – namely this absolute possibility of abstracting from every determination in which I find myself or which I have posited in myself, the flight from every content as limitation. If the will determines itself in this way, or if representational thought considers this aspect in itself as freedom and holds fast to it, this is negative freedom or the freedom of the understanding. – This is the freedom of the void, which is raised to the status of an actual shape and passion. (PR: §5, p.38)

Hegel’s complaint is that this freedom is empty. It is, by its very nature, too universal, too abstract, indifferent to any content, particularisation or determinacy. It is important that Hegel does accept that there is this element to our will, he just insists that it is only one side of it.[[144]](#footnote-144) As such, it needs to be brought together with the particular element of the will. This is arguably Hegel’s key insight – that universality and particularity need to be united in an individual.[[145]](#footnote-145) But we do not have time for Hegel’s key insight here. What matters for our purposes is that the purely universal or abstract element of the will is one-sided.

It thereby makes for an incomplete conception of freedom. Here is Hegel again, later on in the *Philosophy of Right*:[[146]](#footnote-146)

A will that resolves on nothing is not an actual will; the characterless man can never resolve on anything. The reason for such indecision may also lie in an over-refined sensibility which knows that, in determining something, it enters the realm of finitude, imposing a limit on itself and relinquishing infinity; yet it does not wish to renounce the totality which it intends. Such a disposition is dead, even if its aspiration is to be beautiful. ‘Whoever aspires to great things’, says Goethe, ‘must be able to limit himself’. Only by making resolutions can the human being enter actuality, however painful the process may be; for inertia would rather not emerge from that inward brooding in which it reserves a universal possibility for itself. But possibility is not yet actuality. The will which is sure of itself does not therefore lose itself in what it determines. (PR: §13Z, p. 47)

The universal element of the will abstracts from all particularity and determinacy. And consequently, it “cannot achieve anything positive” (PS: §588, p.358), as this would involve “making itself into an *object* *and enduring being*” (ibid).[[147]](#footnote-147) Hegel concludes that:[[148]](#footnote-148)

Universal freedom, therefore, can produce neither a positive work nor a deed; there is left for it only *negative* action; it is merely the *fury* of destruction. (PS: §589, p.359)[[149]](#footnote-149)

The question we face is whether Kant’s conception of freedom is one-sided or incomplete. And insofar as Kant claims that, because we are (transcendentally) free, no object may determine the will, it appears to be just that.

Timmermann presumably has something like this objection in mind, when he claims that, for Kant:

Freedom is not opposed to determination as such, but to the wrong kind of determination. (Timmermann 2007: 122)

That is fine, but we need to ask what the wrong kind of determination is. If, as we have seen with Kant’s (meta-ethical) formalism, the wrong kind of determination involves *any* object, matter or end, anything but the mere legislative form of the will, then Hegel is exactly right to complain that Kant has an overly formal abstract conception of freedom.[[150]](#footnote-150)

We can draw this out further though considering a related point that Hegel makes against Kant. In the next section, I employ a variant of Hegel’s famous empty-formalism objection, against both Kant’s formalism and Kantian constructivism. In doing so, I hope to further unpack the relationship between freedom, autonomy and morality in Kant.

## 1.5 Against Kant’s Formalism and Kantian Constructivism

In this section, I want to continue to distance myself from Kant’s (and Kantian) anti-realism. Beginning with the basics, I think of the realism/anti-realism debate in Euthyphro-style terms:[[151]](#footnote-151) are things valuable just because they emerge from a certain procedure; or do they emerge from the procedure because they are valuable? I align realism with the latter, and anti-realist positions such as Kant’s (meta-ethical) formalism and Kantian constructivism with the former.

There has been a lot of debate recently over whether Kant is best understood as a constructivist or a realist.[[152]](#footnote-152) I want to circumvent this exegetical debate. For what it is worth, I think that neither of these (contemporary) labels tidily captures the entirety of Kant’s moral philosophy. A lot can be said in favour of reading him as either a constructivist or a realist, which I take to show that both these lines of thought figure in his moral philosophy, but I will not make this case here. Instead, I want to continue to focus on one prominent anti-realist line of thought, which I locate in Kant’s (meta-ethical) formalism.

As we have just seen, in analysing the notion of an unconditionally normative principle, Kant contends that no object could lay at its basis. This counts against the moral realist, who hopes that there are objects of unconditional worth that serve as the basis for unconditionally normative claims. Kant’s formalism is thus a type of anti-realism.

However, it is important to emphasise that for both realist and anti-realist readings of Kant, the moral law is not up to us in any voluntaristic sense. It is a formal principle of practical reason that holds independently of any particular act of willing.[[153]](#footnote-153) There is thus an element of realism to Kant’s formalism, but it is not realism in the substantial sense.[[154]](#footnote-154) And it is important not to gloss over this difference. As we have just seen, Kant’s formalism pulls against moral realism through claiming that only a merely formal principle could generate a categorical imperative. Kant’s formalism resists the idea of any moral content independently of, or prior to, a merely formal principle of practical reason.

In what follows, I continue to distance myself from this (meta-ethical) formalism, along with its contemporary constructivist counterparts.[[155]](#footnote-155) In the previous section, I objected to Kant’s argument for his formalism (claiming that it neglected an alternative), whereas this next objection is addressed to formalism as such. For this, we turn once more to Hegel.

Hegel’s empty formalism objection, at its most basic, is the thought that Kant’s formula of universal law lacks content.[[156]](#footnote-156) It is important to distinguish between two issues here. The first concerns whether the formula of universal law can generate any content, or place any restrictions on our conduct. I do not want to get involved in this debate.[[157]](#footnote-157) Many Kantians have offered an excellent account of how the formula of universal law can be formal and generate content, and I am happy to grant them this. Instead, I want to pose what I take to be a more fundamental Hegelian challenge to Kant, namely whether a purely formal principle can generate any *moral* content.[[158]](#footnote-158)

In his commentary on the *Reason as testing laws* section of the *Phenomenology*, Robert Stern poses a dilemma to the Kantian.[[159]](#footnote-159) Here is the first horn:[[160]](#footnote-160)

[…] on the one hand, if the test of non-contradiction is purely formal, it is not clear that failing the test reveals anything of moral relevance: why, if a maxim fails the test, does this show that acting on the maxim would be wrong? (Stern 2013a: 151)

The other horn is as follows:

If, on the other hand, the test is seen as a way in which the agent can discover whether or not by acting in a certain manner she would be free-riding, […] it then presupposes some moral content as part of that test (namely, the wrongness of free-riding, or of manipulating others), rather than determining what is right and wrong through the test, and so is no longer purely formal in this sense. (Stern 2013a: 151)

One could opt for the second horn, which would move us away from the purely formal nature of the moral law, and introduce some moral content. And, this is precisely what some readers of the *Groundwork* think Kant does with the introduction of the formula of humanity,[[161]](#footnote-161) where, as we have seen, every rational being has absolute worth, as an end in itself, which serves as the ground of the moral law (IV: 428. 3-11).

One could also opt for the first horn, where value consists solely in maxims conforming to the standard of universality. But then we lose any explanation of why *this* merely formal requirement is morally significant.

We can bring this into view by comparing the violations of two different requirements of practical reason – those that govern hypothetical imperatives, and those that govern categorical imperatives.[[162]](#footnote-162) An example might help.

If I want a clear head, I should not drink too much coffee. I want a clear head, but I end up drinking too much coffee. I am irrational here, as my behaviour conflicts with a formal principle of practical reason – if you will the end, then you ought to will the means. But this is of no moral significance. In this case, I am irrational, but not immoral.[[163]](#footnote-163)

Changing the example somewhat, imagine that I make a habit of throwing coffee at people who ask me tough questions. Let us grant that this is immoral, and that it (and the associated maxim) fails the universalisability test – another supposedly merely formal principle of practical reason. In this case however, unlike the last, the violation of a formal principle of practical reason is of moral significance. But why? Both cases involve the violation of formal principles of practical reason. In both cases I am being irrational, but in only one of these cases is this irrationality immoral. In treating the universalizability requirement as merely formal, the formalist is unable to account for this.

The realist on the other hand, has an easy answer at this point. The universalizability requirement is important because it calls upon us to not make exceptions for ourselves, not to manipulate others, or to treat them as mere means to the fulfilment of our own ends, and so on.[[164]](#footnote-164)

One commentator who grapples with this question is Andrews Reath. He asks:

If the Hypothetical Imperative is a formal principle that the will is to itself, it also applies unconditionally – to agents qua rational, and not in virtue of any particular ends or interests. What then gives the Categorical Imperative a different normative status? (Reath 2013: 46)

He locates the difference in the claim that, unlike hypothetical imperatives, categorical imperatives are synthetic and thereby yield ends. Here is Reath himself:

The special authority of the Categorical Imperative is due to its synthetic character – that it yields some substantive ends and principles of conduct that are independent of and take deliberative priority over an agent’s contingent ends. (Reath 2013: 46)

In contrast [to the Hypothetical Imperative], pure practical reason will address questions of evaluation that are beyond the scope of empirical practical reason. It will introduce standards for evaluating actions and ends that are non-instrumental, and apply independently of given desires and ends – principles which ground judgements of intrinsic goodness or acceptability to anyone, which for Kant are the basis of justification to others. There are grounds for thinking, in addition, that pure practical reason will be concerned with the evaluation and choice of ends for their own sake (in contrast to the choice of actions as means to ends). This will include the capacity to elect aims and goals viewed as intrinsically good or worthy of choice, which can initiate actions and structure larger practical pursuits. (Reath 2006: 69)

This seems right, but it will not help the formalist or the constructivist. Once more, a crucial question remains, namely, what is the normative status of these ends? Is failing to pursue them merely a rational failure, and if so, what makes the rational failure involved in failing to pursue these ends normatively distinctive from the rational failure involved in failing to pursue instrumental ends? Reath notes that the ends that the categorical imperative yields “take deliberative priority over an agent’s contingent ends” (2013: 46). I agree, but contend that the formalist cannot vindicate this claim; once more, there are many requirements of reason, and in treating the moral law as merely formal, one is unable to give it distinctive normative status. Of course, one could say that the ends yielded by the moral law are of independent moral worth, but that would move us away from formalism and constructivism to introduce some moral content.

Stephen Engstrom has offered another avenue to the formalist.[[165]](#footnote-165) He contends that unless our judgements conform to the form of practical knowledge, chiefly universalisability, they will not be proper judgements:

[…] by being in conflict with the understanding of practical knowledge that they involve merely in virtue of being practical judgments, they will in that sense be in conflict [with] their own form and hence with themselves and so will lack the self-sustaining character essential to cognition, a character that they themselves, as exercises of the capacity for such cognition, purport to have. (Engstrom 2009: 128)

I am happy to grant Engstrom all of this, but again, the question remains as to why it is that judgements being proper judgements – having the self-sustaining nature essential to cognition – is of *moral* significance? I make improper judgements all the time. Now sure, I do so at the pain of irrationality, but irrationality *per se* is not morally significant. It is worth pointing out that Engstrom is aware of the dilemma at hand, and the difficulty it poses.[[166]](#footnote-166) In the introduction to *The Form of Practical Knowledge*, he sketches his response. He sets out a view of universalization as a logical operation externally imposed upon a maxim, and then mirrors our objection:[[167]](#footnote-167)

Why should it matter to us whether a maxim we follow results in a contradiction if universalized?[[168]](#footnote-168) Why should we suppose that an externally imposed logical operation is invested with the authority to constrain the will? (Engstrom 2009: 18)

His view of universalization is supposed to overcome this. On his account, the universalization test makes manifest the form of practical knowledge, and so is not an *externally* imposed logical operation (Engstrom 2009: 18-9). So conceived, the contradiction arises between the form and the content of a maxim.

However, our objection can still be pressed. We can ask:

Why should it matter to us whether a maxim we follow results in a contradiction if universalized? Why should we suppose that an *internally* imposed logical operation is invested with the authority to constrain the will?

There are at least two issues in play here.

Firstly, if it is just a logical operation, why does it matter whether it is internally or externally imposed? The principle of non-contradiction, for instance, is a norm of reason and accordingly has normative authority over both our thoughts and actions, regardless of whether it is internally imposed or not.[[169]](#footnote-169)

Secondly, let me re-iterate my main objection: if this is merely a logical operation, why is it of moral significance? Hypothetical imperatives seem to involve certain internally imposed logical operations (if you will the end, you also ought – on pain of contradiction – to will the means), but failure to comply with these is not necessarily a moral failure. As we saw with the two coffee examples, the formalist is unable to explain why the addition of a merely formal requirement of universalizability is of moral significance.

It might seem like I have stacked the deck here, in asking for an account of the moral significance of the (supposedly) merely formal requirement of universalisability. The formalist (and constructivist) might reply that this both misses the point and begs the question against them.[[170]](#footnote-170) The whole point of (meta-ethical) formalism (and constructivism all the way down) is that there is no moral content prior to, or independently of, a formal procedure. But that does not get them off the hook. By committing to such an account, they are unable to explain the moral significance of the universalizability requirement, or differentiate it (normatively) from other formal principles of practical reason.

Before I continue, I want to make clear that nothing I have said counts towards jettisoning Kant’s formula of universal law. For all I have said, it might still be the case that it is an excellent heuristic or test, and accordingly has great epistemic significance.[[171]](#footnote-171) It might also work well in navigating competing conceptions of what is right in ethics or politics, but none of this directly bears upon the meta-ethical issues surrounding the possibility of moral content independent of this formal principle. Both the realist and the formalist can allow that this formal principle has epistemic and practical worth.

Part of the appeal of Kant’s formalism seems to be that, in treating the moral law as merely formal, it attempts to account for the unconditional nature of morality’s demands without having to introduce any metaphysics of value.[[172]](#footnote-172) And in linking morality to a formal principle of reason, it does link morality to something that is unconditionally normative. Unfortunately, being unconditionally normative is a necessary, rather than a sufficient condition of being moral. And this ties into the formalism objection. Any formal principle of reason is unconditionally normative in one sense – you ought not to contradict yourself, regardless of what you desire – but moral claims also override (or silence) what we desire. And for this, we need more than a merely formal principle of practical reason.

I take this to be a major weakness of Kant’s formalism (and Kantian constructivism). And, alongside this weakness (as we saw in the last section – §1.4), Kant’s argument for formalism neglects an alternative. Now the formalist might accept that these are serious problems, but still insist that, despite this, their position is still in a better state than the alternative, namely, the introduction of moral content. For this we need to turn to the Kantian objections to moral realism. These can be thought of in four camps, and they run as follows:

* Metaphysical: There are no independently valuable ends.
* Epistemological: Even if there were such ends, we could not know about them.
* Normative: Even if there were such ends, and we could know about them, they would not be unconditionally normative.
* Motivational. Even if there were such ends, and we could know about them, and they were unconditionally normative, they would not be able to motivate us.

In what follows, I will address each of these objections in turn.

In the previous section of this chapter (§1.4), I argued that Kant is mistaken to think that the moral law must be merely formal. I accused him of neglecting an alternative – objects of unconditional worth which ground unconditionally normative claims. Over the next few pages, I will consider various ways in which one might defend Kant’s formalism (and Kantian Constructivism) against this charge.

### Metaphysical: There are no independently valuable ends.

One simple option for the Kantian is to deny that there are any objects of unconditional worth. Of course, this does not count in favour of the strong claim that the moral law *must* be merely formal, but in the absence of any objects of unconditional worth, perhaps a formal principle is the only option left.

A revealing discussion occurs towards the end of *Groundwork* III. There, Kant considers the world of understanding, and its relationship to moral motivation. He begins by noting the limits of knowledge that transcendental idealism poses:

[…] if it [practical reason] were to fetch in addition an object of the will, i.e. a motive, from the world of understanding, then it would overstep its bounds, and presume acquaintance with something of which it knows nothing. (IV: 458. 16-8)

Kant continues to claim that the concept of a world of understanding is necessary:

[…] but without the least presumption to think of it further than merely according to its formal condition, i.e. to the universality of the maxim of the will, as a law, and hence to the autonomy of the latter, which alone is consistent with its freedom; whereas all laws that are determined with reference to an object yield heteronomy, which can be found only in laws of nature and also apply only to the world of sense. (IV: 458. 29-35)

There is an epistemic point in here, which I will return to shortly. Here though, I am interested in the metaphysical objection.

I have argued that Kant is mistaken in thinking that the moral law needs to be merely formal. But this might take us beyond transcendental idealism. I am fine with that, but Kant would not be. Either way, now is not the time to discuss the merits of transcendental idealism (I will return to this in chapter 3). If we conceive nature in the way Kant does, we will find no value in it.

Engstrom directs a similar complaint against moral realism:

[…] the familiar difficult questions of how these directions could be understood to have their footing in any such external setting, given the scientific knowledge we now have of nature and the physical universe […] the view that there is a teleological order in nature […] that provides us with directions for living – a view we today can look upon only with suspicion. (Engstrom 2009: 11)

Engstrom references Mackie here (1977: 38n), which is somewhat peculiar.[[173]](#footnote-173) The Kantian can accept that nature lacks a teleological order that provides us with directions for living,[[174]](#footnote-174) but there is more to the Kantian picture than nature. Crucially, there is also the transcendental. And however one conceives that, the Kantian picture moves beyond Mackie’s sparse ontology as it contains a distinctive type of being – the Kantian rational agent.[[175]](#footnote-175) These beings, in light of their reason and freedom, stand out from the rest of nature. Both the formalist and the (Kantian) realist admit this unique being into their ontology – the Kantian rational agent, who either possesses value, or the capacity to recognise and follow merely formal principles of practical reason.

Of course, the formalist might insist that rational agents do not possess independent value, but that is exactly the issue under debate. The appeal to nature, Mackie and metaphysics does not settle this. We can now turn to the epistemological challenge.

Epistemological: Even if there were such ends, we could not know about them.

Here is a fuller version of the above passage from Engstrom:

[…] the familiar difficult questions of how these directions could be understood to have their footing in any such external setting, given the scientific knowledge we now have of nature and the physical universe, *and of how we could ever know them if they were somehow there outside us, written into things* […] the view that there is a teleological order in nature that we somehow rationally apprehend and that provides us with directions for living – a view we today can look upon only with suspicion. (Engstrom 2009: 11) [Emphasis mine]

In this, we see the epistemic prong of the attack on moral realism. Let us continue to focus on the value of rational nature.

Perhaps we do only come to know that rational nature has value through the activity of practical reason, or though formal procedures, but that is fine. That would just show that these procedures have an important, or even essential, epistemic role.

One might still worry that attributing moral realism to Kant inevitably turns him into to some sort of rational intuitionist, as it might seem that we need to posit some distinct faculty to know that rational nature in general has value.[[176]](#footnote-176) We can find an answer to this in Kant’s discussion of the formula of humanity.

Kant tells us that the ground of a possible categorical imperative lies in the existence of something with absolute worth (IV: 428. 3-6). He then proceeds to argue from elimination as to what this is. He rules out the objects of our inclinations (IV: 428. 11-4), the inclinations themselves (IV: 428. 14-17), and non-rational beings (IV: 428. 19-21), to end up with rational beings as the only option standing (IV: 428. 21-33). He then claims that, if there is to be a categorical imperative, the ground of this principle will be that rational nature exists as an end in itself (IV: 428. 34 – 429. 3).

I will say more about the formula of humanity towards the end of the chapter (§2.2), but will stop here for now. For present purposes, what matters is that this reading outlines how it is possible to *know* that rational nature in general has value, without turning Kant into a rational intuitionist.[[177]](#footnote-177)

### Normative: Even if there were such ends, and we could know about them, they would not be unconditionally normative.

Here we encounter a familiar Kantian claim – that only a formal law could be unconditionally normative.[[178]](#footnote-178) Henry Allison has recently put forth a similar claim – that only a self-imposed law could be unconditionally normative – which serves as the centrepiece of his recent commentary on the *Groundwork*.[[179]](#footnote-179) Unfortunately, as we saw earlier (§§1.3-1.4), the claim does not hold, in either of these varieties. It neglects an alternative, namely an unconditionally normative claim that is neither merely formal nor self-imposed.

Allison’s main argument for his claim ties into the Reciprocity Thesis. He has repeatedly argued that, as transcendentally free, the will cannot be determined by anything external to it, and thus must provide its own law.[[180]](#footnote-180) However, this does not follow. Part of the problem stems from his (and Kant’s) conception of freedom.

As we saw earlier, another part of Hegel’s objection to Kant is that his conception of freedom is too universal, too formal, too abstract, too indifferent to determination. We can bring this to light by, once more, considering rational nature. This is an end that appears to be of unconditional value. Let us assume that it is, and that we can know this. Then why could this unconditionally valuable end not serve as the ground of unconditionally normative claims? We have seen Kant suggest that it could:

But suppose there were something *the existence of which in itself* has an absolute worth, that, as an *end in itself*, could be a ground of determinate laws, then the ground of a possible categorical imperative, i.e. of a practical law, would lie in it, and only in it alone. (IV: 428. 3-6)

We have also seen Allison himself accept something similar:

[T]he idea that the existence of something that is an end in itself is a sufficient condition of a categorical imperative seems relatively unproblematic, since, arguably, entities with this status (if they exist) could be the source of unconditioned commands. (Allison 2011: 205)

Allison is right on this, but it counts against Kant’s claim elsewhere that the moral has to be merely formal, and Allison’s own claim that only a self-imposed law could be unconditionally normative.

Timmermann offers a related objection to the realist, invoking the autonomy/heteronomy contrast. He claims that:

Acting *for the sake of the good* would be a mark of heteronomy (Timmermann 2007: 125)

I do not think this is a mark of heteronomy; and if it is, then heteronomy no longer looks so bad.[[181]](#footnote-181) Acting for the good just because one happens to like the good is problematic. But if we act *for the sake of the good*, that involves recognising something as good and acting accordingly. And there is nothing problematic about that.

To summarise, if there were any ends with absolute worth and we could know them (the metaphysical and epistemological issues), then they could be unconditionally normative for us.

### Motivational: Even if there were such ends, and we could know about them, and they were unconditionally normative, they wouldn’t be able to motivate us.

Engstrom gives a clear account of the motivational worry:

If we suppose that our capacity to observe moral requirements is based in a rational appreciation of them, then we are unable to account for how that appreciation can move us to act accordingly. (Engstrom 2013: 138-9)

He also notes that:

It was from a picture of this sort that Hume was understandably recoiling (Engstrom 2009: 12)

I find this picture Humean, all too Humean. That is not to say it is completely without merit, but rather that it should not be a defining worry for the Kantian. The Kantian can (and should) move beyond the Humean picture of requiring desires to motivate us – pure reason can be practical. And here, I see no significant difference between claiming that we can be motivated by merely formal practical principles rather than the value of rational nature.

In the previous subsection, I introduced Timmermann’s claim that acting for the sake of the good would be problematic, but did so without any context. In considering Timmermann’s views on motivation, we can see why he thinks this, and why he opts for Kant’s formalism. He offers the following account of moral action:[[182]](#footnote-182)

[…] all action involves both form and matter, i.e., an end as an object of volition and a law. As regards moral value, the decisive question is whether the formal or the material element determines the ‘dutiful’ act, i.e., the act that, on the face of it, coincides with what duty commands (*pflichtm*[*äß*](http://www.dict.cc/deutsch-englisch/pflichtgem%C3%A4%C3%9F.html)*ig*). An agent is thus *either* interested in an end because he is interested in the action, on moral grounds, because it is commanded by a purely formal law; *or* because he is interested in a certain manner for the sake of bringing about the end that appeals to him. (Timmermann 2013: 219-220)[[183]](#footnote-183)

These are not the only two options. Timmermann claims that we either act for the sake of an end that just happens to appeal to us, or we act “on moral grounds, because it is commanded by a purely formal law” (2013: 220), but we could also act on moral grounds in recognising an object of unconditional value.

Timmermann claims that – apart from the special exception of the moral law – we always need a desire to motivate us. This opens up an internalism/externalism debate about motivation, which I will return to consider in chapter 4. For now however, I want to note that it is important to pull apart the (internalist) claim that we need a (pre-existing) desire to motivate us.[[184]](#footnote-184) It can be understood in two senses: If it is understood in a weak sense (something like action requires an affective state), then it is true, but no longer counts against the externalist, whereas if it is understood in a strong sense (action requires an actual pre-existing desire), then it is no longer plausible.[[185]](#footnote-185) Once we do this, we can then disarm the motivational threat to moral realism.

We do not need to sign up to either Kant’s formalism or Kantian constructivism. For one, Kant’s argument for his formalism neglects an alternative. And secondly, viewing the moral law as a merely formal principle of practical reason leaves us unable to explain the moral significance of the – supposedly merely formal – universalisability requirement, or differentiate it (normatively) from other formal requirements of reason. This was the initial result of the previous two sections. We have just turned to consider the plausibility of introducing some moral content. We have seen a variety of (Kantian) arguments against this approach, but none of them are convincing.[[186]](#footnote-186)

# Conclusion

We can now finally tie this back into Kant’s claim that a free will must be autonomous. Kant’s claim appears to be that a free will cannot be determined by anything extrinsic to it, and thus must provide its own law. As we discussed at the start of this chapter, if this is just taken to mean that a free will must be autonomous in a weak sense – that it must act on maxims, which it takes as up to itself (the result of our previous chapter - §3.2) – that is fine. However, in order to move from freedom to morality, Kant needs more than this. He needs to link a conception of autonomy as agency with a meta-ethical conception of autonomy, and ultimately, the moral law itself. As we have seen though, these connections fall short. Just because we are free – even transcendentally free – does not mean that we have to abstract from all possible objects of the will. The move from autonomy as agency to this meta-ethical conception of autonomy fails. And we see this in McDowell and Hegel – we should not think of freedom as freedom from reasons.

Let us then return to the first section of *Groundwork* III. I broke this down into four steps:

1. A free will is independent of determination from alien causes (IV: 446. 7-10).
2. Nevertheless, a free will must have some law determining it (IV: 446. 15-21).
3. As independent from determination from alien causes (1), the law that determines the free will (2) must therefore be the will’s own law, that is the will must be autonomous (IV: 446. 24 – 447. 2).
4. The claim that the will is a law to itself is just the same as the claim that we are to act on universalisable maxims, which is the supreme principle of morality (IV: 447. 2-7).

I am happy to accept the first two steps, but cannot follow Kant through the third and fourth. And one last time, this all hinges upon how we understand ‘autonomy’ in the third step. The only way to make that step plausible is to read it in the weak sense of agency (namely that a free will must act on its own maxims or principles). But then, the fourth step no longer follows. In the end then, I view the argument of the opening section of *Groundwork* III, along with Kant’s formalism and Kantian constructivism, as unsuccessful.

# Freedom and Morality

In this chapter, we have considered the opening argument of *Groundwork* III, where Kant appears to move from freedom to autonomy to the moral law. I have argued that this ambitious attempt to connect freedom, autonomy and the moral law fails. I would now like to leave Kant’s argument aside, to present a modest account of its conclusion – that a free will stands under the moral law.

## 2.1 A Free Will stands under the Moral Law

Kant concludes his four-step argument in the first section of *Groundwork* III thus:

[…] a free will and a will under moral laws are one and the same (IV: 447. 6-7)

In this section, I will consider what it means to stand *under* the moral law.

I propose that to be under the moral law means that the moral law is normative and prescriptive for you. In this, I depart from Allison and Schӧnecker, who have both recently argued that the moral law is not prescriptive for a holy will. I dispute this. Over the course of this section, I make the case that the moral law is both normative and prescriptive (although not obligatory) for *all* (suitably) free wills. I also consider the significance of this claim, suggesting that my alternative puts Kant in a better place to respond to Reinhold’s objection that he cannot adequately account for evil actions.

What does Kant mean when he says that a free will is *under* the moral law? The claim could be read as saying that a free will is *bound* by the categorical imperative. But, as Schӧnecker has insisted, this is not quite right.[[187]](#footnote-187) Kant is talking about the moral law here (not the categorical *imperative*), and at places he mentions that a holy will also stands under the moral law.[[188]](#footnote-188) So ‘under’ cannot mean bound (or obligated), where as we saw in chapter 1 (§1.1), Kant conceives of obligation as the imperatival form that the moral law takes for finite rational beings.

Things are further complicated by the fact that, for a holy will, the moral law describes its actions. This can lead to the thought that standing under the moral law involves nothing normative, but I think that would be mistaken.[[189]](#footnote-189) The moral law is normative, but also descriptive of a holy will and obligatory for us.

We can see this in a passage from *Groundwork* II, where Kant writes that:

If reason determines the will without fail, then the actions of such a being that are recognized as objectively necessary are also subjectively necessary; i.e. the will is a capacity to choose *only that* which reason, independently of inclination, recognizes as practically necessary, i.e. as good. (IV: 412. 30-4)

Here, Kant talks of a holy will *recognising* that which is practically necessary [*practisch nothwendig*], or good.[[190]](#footnote-190) The good is normative – one should pursue the good. The holy will recognises what is practically necessary and acts accordingly without thereby feeling necessitated; or again, it responds to the normativity of the moral law without being obligated or bound by it.

I want to break with other commentators to contend that the moral law is prescriptive for both us and the holy will.[[191]](#footnote-191) It is uncontroversial that the moral law is obligatory for us, but not so for the holy will; the holy will is rational, but has no sensible inclinations that could resist the moral law, and so follows the moral law without being *bound* by it.[[192]](#footnote-192) The moral law thus describes the actions of a holy will. But it also prescribes how the holy will *ought* to act – more on this tricky word in a minute. Of course, this prescription is not experienced as an obligation, as the holy will has no sensible inclinations that could resist it.

To clarify this point, imagine some aliens pay a visit to earth. They are rational, and have no sensible inclinations that conflict with their reason. They can recognise the moral law, and so also recognise that they ought to “*act only according to that maxim through which you can at the same time will that it become a universal law”* (IV: 421. 7-8), and “act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means” (IV: 429. 10-2), and so on. They have no sensible inclinations that pull against these principles, and so follow them without reluctance or hesitation. But they still should follow these principles. And they do so because they recognise their normative force. The moral law prescribes how they – and indeed any rational agent – should act.

Here we run up against a limitation of language though, in that ‘ought’ and ‘should’ can be used in two ways: 1) to convey prescriptivity; 2) to convey an obligation. When I say that the moral law prescribes how any rational agent ought to act, I am using ‘ought’ and ‘should’ in the first sense. Kant himself runs up against this limitation. In *Groundwork* II, he claims that:

[…] no imperatives hold for the *divine* will and generally for a *holy* will: here the *ought* [*das Sollen*] is out of place, because *willing* [*das Wollen*] already of itself necessarily agrees with the law. (IV: 414. 5-8)

Once again though, there are two senses of ‘ought’: the prescriptive and the obligatory. The latter is out of place for the holy will, but not the former. In running these two senses of ‘ought’ together, we can miss this point.

Consider, for example, Kant’s treatment of one’s own happiness in the *Metaphysics of Morals*:

For *his own happiness* is an end that every human being has (by virtue of the impulses of his nature), but this end can never without self-contradiction be regarded as a duty. What everyone already wants unavoidably, of his own accord, does not come under the concept of *duty*, which is *constraint* to an end adopted reluctantly. Hence it is self-contradictory to say that he is *under obligation* to promote his own happiness with all his powers. (VI: 386. 1-7)

Kant makes the point that all human beings have their own happiness as their end. He then argues that we can therefore have no duty to promote our own happiness because “the concept of *duty*, […] is *constraint* to an end adopted reluctantly” (VI: 386. 4-6). With this set up, we miss out on the interesting issue of whether there are reasons to promote one’s happiness. Does one’s own happiness have any (direct[[193]](#footnote-193)) normative status?[[194]](#footnote-194) Is it a prescription, something that we should do, albeit one that – due to our constitution – we happen to follow without reluctance? By running together prescription and obligation, it becomes difficult to adequately address these issues.

The moral law is normative. It is thus both prescriptive for us and the holy will. Of course, it is also binding for us, but we should not conflate its prescriptivity and bindingness (even though, for us, they both come together). In setting things up this way, I depart from both Schӧnecker’s and Allison’s recent accounts of this issue. In what follows, I outline their positions, and then detail my departure from them.

Schönecker focuses on the following claim, which he calls Kant’s ‘thesis of analyticity’ [*Analytizitätsthese*]:

[…] a free will and a will under moral laws are one and the same.

Thus if freedom of the will is presupposed, morality along with its principle follows from it, by mere analysis of its concept. (IV: 447. 6-9)

Schönecker insists upon three points here: 1) ‘under’ in this claim does not mean bound (2006: 303; 2013: 230, 238-9);[[195]](#footnote-195) 2) the moral law is analytic for a perfectly rational being (2006: 303-4); 3) the moral law is descriptive of – but not prescriptive for – a perfectly rational being (2013: 231). I agree with the first claim, am unsure about the second,[[196]](#footnote-196) and (as we have seen) disagree with the third. I take this final disagreement to be the interesting issue.

Allison also departs from Schönecker in claiming that the moral law is normative for the holy will, but still insists that it is not prescriptive:

[… Kant’s] conception of a perfectly good or holy will is clearly normative, inasmuch as it concerns what is good in the highest respect, but it is also descriptive rather than prescriptive, since it describes how such a will necessarily acts (according to the moral law) rather than prescribing how it ought to act in order to be perfectly good or holy. Accordingly, for Kant there can be normativity without prescriptivity, though not the converse. (Allison 2011: 162)

He later claims that:

[…] the idea of conformity to universal law is not itself a prescription (in which case it would be vacuous), but it is still a description of the principle governing the volitions of a will considered as good without limitation. Prescriptivity, it was argued, only enters the story when this idea is connected with the will of a finite rational agent equipped with maxims embodying its own non-moral interests (Allison 2011: 170)

But it is not at all clear why the notion of prescriptivity here is vacuous, which we can see by looking at another passage in Allison:

[…] the actions of an absolutely good will must not only conform to universal law as such, they must also be performed, or, better, their underlying maxims adopted, *because they conformed*, which means that such conformity must provide the reason for adopting, or at least a sufficient reason for rejecting, maxims that fail to conform to universal law as such. Otherwise, the conformity would be a purely contingent matter, having no moral import. (Allison 2011: 170)

In the case of the holy will, conformity with the moral law is not a contingent matter. The holy will always follows the moral law, but it does so *because* it is the moral law. It is not enough to always conform with the moral law. The moral law must be *followed*, and this involves recognising its normativity and responding to it – one ought to treat humanity as an end-in-itself, and so on.

Another example might help illustrate this. Imagine that I make two robots: Robo-Leibniz and Robo-Kant. Robo-Leibniz is programmed to conform with the moral law, whereas Robo-Kant can recognise the moral law. Neither Robo-Leibniz nor Robo-Kant have any sensible inclinations. Robo-Kant thereby always follows the moral law, where Robo-Leibniz always conforms to it (without recognising it, nor acting for the sake of it).[[197]](#footnote-197) For Kant (the regular, non-robo Kant), there is a huge difference between these two beings. One responds to the normativity of the moral law, while the other does not.

Leaving the robots aside, my departure from Allison might just be a point of language, in that there are two senses of ‘prescriptivity’. One is close to obligation, and involves being told that you should do something. And this, like obligation, does seem out of place for the holy will – it does not need to be told what to do, and any such telling would be vacuous. There is another sense of ‘prescriptivity’ though, which goes hand in hand with normativity. To recognise the moral law is to recognise it as normative, and to recognise that one should act according to it. If we lose this, we lose precisely what is distinctive about the moral law. Consider what it would mean if the moral law was not prescriptive (in this latter sense): It would mean that the holy will would recognise the moral law and its normativity, but this would provide no reasons whatsoever for how the holy will should act. Pace Allison, it is not vacuous to think of the moral law as prescriptive for the holy will; what is vacuous is to think of the moral law as normative, but not at all prescriptive.

And we can find some basis for this reading in the text. Earlier on, we saw Kant claim that:

If reason determines the will without fail, then the actions of such a being that are recognized as objectively necessary are also subjectively necessary; i.e. the will is a capacity to choose *only that* which reason, independently of inclination, recognizes as practically necessary, i.e. as good. (IV: 412. 30-4)

Kant talks of a holy will *recognising* that which is practically necessary [*practisch nothwendig*], or good.[[198]](#footnote-198) The good is normative *and* prescriptive – one should pursue the good. The holy will recognises what is practically necessary and acts accordingly without thereby feeling necessitated; or again, it responds to the prescriptivity of the moral law without being obligated or bound by it.

The question of whether the moral law is just descriptive of, or also prescriptive for, a holy will might appear to be nothing more than a scholars’ quibble (alongside the question how many angels can dance on the head of a pin?) However, it is crucial to understanding the claim that a free will is *under* the moral law, especially if we are to overcome Reinhold’s objection.[[199]](#footnote-199)

And here, once again, I depart from Schönecker. As we have seen, he emphasises the parts of *Groundwork* III that count in favour of treating the moral law as descriptive of – but not prescriptive for – a holy will. He is aware of the difficulties that this reading generates, but contends that:

[…] if an interpretation *a* is available that is better than interpretation *b* – roughly, an interpretation is better if it coherently (i.e., consistently and comprehensively) accounts for more textual observations than another – then *a* must be preferred over *b* even if *b* ascribes *a* theory to the text that seems more attractive than the theory based on *a*. (Schönecker 2013: 226)

He then applies this directly to the issue at hand:

For instance, the theory of analyticity brings about the question of how evil actions are possible: if to act freely is to act morally, and vice versa, then there appears to be no room for evil actions that are free. If the best interpretation show that bisubjunction to be Kant’s position, it should not be given up because it has an unwelcome implication (from ‘our’ point of view); maybe the implication is a reason to give up that bisubjunction but it is not a reason to give up the interpretation if the latter is the best available. (Schönecker 2013: 226-7)

Schönecker’s commitment to the text is admirable, but I am happy to opt for the most philosophically plausible reading. After all, I view Kant’s moral philosophy as alive, and am accordingly interested in the parts of it that work. So if option *b* has less textual support than *a*, but works better, I am happy to opt for *b*.[[200]](#footnote-200) Schönecker acknowledges that his reading of Kant’s claim that we are ‘under’ moral laws renders Kant unable to account for evil actions (2013: 230), but claims so much the worse for Kant. I seek to avoid this unwelcome implication, and hope to have provided a reading of Kant’s claim that does so. If we treat the moral law as merely descriptive of a free will, then insofar as we are free, the moral law will just describes our actions. This opens Kant up to Reinhold’s objection that immoral actions are not free and thereby not imputable. If however, we conceive of a free will as responding to the normativity (and prescriptivity) of the moral law, this opens up a way of circumventing this problem.

In the first half of this chapter, we saw Allison and Korsgaard argue that, if we are (suitably) free, then we are thereby committed to the moral law, in virtue of our freedom (§§1.1-1.3).[[201]](#footnote-201) On this ambitious account, the opening section(s) of *Groundwork* III establish the moral law (for us) through our freedom. I have attempted to bypass this approach. And once we do so, we can then read the move from freedom to the moral law as relatively unproblematic. This is my modest account.

To stand under the moral law is to be subject to it.[[202]](#footnote-202) It is for the moral law to prescribe how one should act. This principle prescribes the behaviour of all (suitably) free beings. It is also descriptive of a holy will, and obligatory for us. However, we are all free and are thereby subject to the moral law.

I want to emphasise one final point here, namely that I follow Allison in claiming that, if we are to make this move work, *transcendental* freedom is required. The moral law is unique in that, unlike a hypothetical imperative, it does not derive its normative force from any antecedent desire. By contrast, it commands unconditionally. Insofar as we are capable of standing under such a principle, we will need to be able to abstract from our desires. What exactly this involves, and whether we get such freedom from rational agency is the issue of our fourth and final chapter. There, I will attempt to disarm the threat of a gap between rational agency and the moral law. For now though, we can continue to explore one final aspect of the relationship between freedom and morality in Kant.

## 2.2 Freedom as the *ratio essendi* of Morality

In this chapter, we have considered the freedom involved in moral agency. In the introduction to this thesis, I noted that freedom is also required for our status as moral patients. Kant’s discussion of the formula of humanity is crucial here. In what follows, I consider this aspect of the relationship between freedom and morality. In doing so, I sketch a realist alternative to the formalism and constructivism we have just looked at.

As usual, let us begin with the relevant text. In the build up to the introduction of the formula of humanity, Kant tells us that:

[…] what serves the will as the objective ground of its self-determination is the *end*, and this, if it is given by mere reason, must hold equally for all rational beings. (IV: 427. 21-4)

He proceeds to introduce this end, and its role (which we saw earlier):

But suppose there were something *the existence of which in itself* has an absolute worth, that, as an *end in itself*, could be a ground of determinate laws, then the ground of a possible categorical imperative, i.e. of a practical law, would lie in it, and only in it alone.

Now I say: a human being and generally every rational being exists as an end in itself, *not merely as a means* for the discretionary use for this or that will, but must in all its actions, whether directed towards itself or also to other rational beings, always be considered *at the same time as an end*. (IV: 428. 3-11)

He then offers a short argument from elimination to support the claim that rational beings are ends of absolute worth. He considers what objects could have unconditional worth, and rules out the objects of our inclinations (IV: 428. 11-4), the inclinations themselves (IV: 428. 14-17), and non-rational beings (IV: 428. 19-21). Rational beings end up as the only option standing (IV: 428. 21-33).[[203]](#footnote-203)

Kant concludes:

If, then, there is to be a supreme practical principle and, with regard to the human will, a categorical imperative, it must be such that, from the representation of what is necessarily an end for everyone, because it is an *end in itself*, it constitutes an *objective* principle of the will, and hence can serve as a universal practical law. The ground of this principle is: *a rational nature exists as an end in itself*. That is how a human being by necessity represents his own existence; to that extent it is thus a *subjective* principle of human actions. But every other rational being also represents its existence in this way, as a consequence of just the same rational ground that also holds for me; thus it is at the same time an *objective* principle from which, as a supreme practical ground, it must be possible to derive all laws of the will. (IV: 428. 34 – 429. 9)

Aside from Kant’s hasty overlooking of animals (which I will not discuss in this thesis[[204]](#footnote-204)), I take this to be a plausible and promising moral realism. Kant conceives of morality as unconditionally normative, and argues that rational nature is the ground of this. Rational nature exists as an end in itself, of absolute worth, which stands as the ground of the categorical imperative; elsewhere, Kant tells us that “*Autonomy* is […] the ground of the dignity of a human and of every rational nature” (IV: 436. 6-7).

Rational nature, freedom and autonomy occupy a pivotal role in Kant’s meta-ethics. They also stand at something like a crossroad between formalism, constructivism and realism. They are the ground of the moral law, but they are not external to the will; in a sense, they just are the rational will. Moreover, if there were no rational agents in existence, there would be no rational nature, freedom, or autonomy, and thereby (at least for Kant) also no value.[[205]](#footnote-205)

In what follows, I would like to consider three alternatives to the realist account of the value of rational nature. The first claims that the moral law is constitutive of rational agency, the second that rational nature only has value insofar as it is capable of following the (merely formal) moral law, and the third (much like the ambitious account we considered in the first half of this chapter) attempts something like a proof of morality through freedom. I argue that all three of these alternative accounts fall short, and conclude in favour of a realist treatment of the value of rational nature.

Beginning with constitutivism, the basic claim is that there are certain norms that are constitutive of agency, and that, if we fail to comply with these norms, we will fail to be agents. The hope is then that we can ground moral norms in the norms constitutive of agency. Kantian constitutivists, for instance, offer accounts of how the moral law (typically the universability requirement) is one such norm, and claim that if one fails to act in this way, one fails to be an agent.[[206]](#footnote-206)

There is a powerful objection to this approach though, which concerns the normative significance of agency. David Enoch has famously pushed the constitutivist on this. He points out that one could fail to follow the norms of agency, and thereby fail to be an agent, but be something relatively close, a shmagent – someone who is an agent in all respects apart the fact that they fail to follow certain norms that are constitutive of agency.[[207]](#footnote-207)

Let us grant the Kantian constitutivist that it is constitutive of being an agent that my maxims are universalisable. Now imagine that I am an agent in every respect, apart from my annoying habit of throwing coffee at people when they ask me tough questions (and again, let us assume that this fails the universalisability test). In this, I fail to be an agent, and so am instead a shmagent (perhaps performing a shmaction).

The key issue concerns the normative difference between these two beings. Why is being a certain type of agent so important? Perhaps such agency has moral value, but that would move us away from formalism and constructivism towards recognising the independent value of agency.

There seem to be two options. Either:

(1) We ought to care about being agents (rather than shmagents) because agency is (independently) valuable, or

(2) There is no morally significant reason for us to be agents (rather than shmagents).

Once again, I opt for the realist option (1).[[208]](#footnote-208)

I want to emphasise that Enoch’s point is not a sceptical one. Enoch is happy with the value of agency. His point is rather against the consitutivist, who attempts to ground morality in the norms constitutive of agency. Enoch’s challenge is that this just pushes the question back to whether agency itself is morally valuable: if it is, then we have a realist account of the value of agency; if it is not, then there is no morally significant reason why I should be an agent rather than a shmagent. Again, this is not scepticism about morality, but rather a challenge to a particular attempt to ground moral norms in non-moral norms.

Moving along, another option for the constitutivst is to claim that, if we fail to comply with the norms constitutive of agency, we will not fail to be agents, but instead fail to be good agents. That seems plausible, but a similar question faces this approach, namely what is the value of agency? What is important about being a good, rather than a bad, agent? If good agency is morally valuable, that provides an answer, but moves us towards realism. If good agency is not (independently) morally valuable, then once again, the moral significance of the difference between good and bad agency is missing.

Turning to the second alternative, some commentators argue that rational nature only has value insofar as it is capable of following the moral law. On this understanding, our capacity to act according to the moral law is what gives rational nature value, so it is a mistake to think of rational nature as having moral worth prior to (or independently of) the moral law. Here is Timmermann on this:

What first and foremost distinguished human beings from animals is the capacity for moral choice. […] The only other kind of rational practical principle is a hypothetical imperative, which is concerned with the choice of *means*, not ends; and there is no dignity or purity in that. (Timmermann 2013: 215)

I disagree that there is only dignity in following the (merely formal) moral law. There is choice and dignity involved in all sorts of non-moral matters: projects, relationships and so on (this is part of my interest in a non-moral conception of freedom). However, even aside from this disagreement, it is still not clear why following a merely formal practical principle involves such dignity. There is dignity involved in respecting other people’s agency, but if we treat the moral law as merely formal, we again lose any explanation of why following it (rather than other formal principles of reason) involves dignity.

For our final alternative account of the value of rational nature, we return to Hegel, one last time. In our discussion of Kant’s formalism (§1.4), we saw Hegel characterise two elements of the will: the ability to abstract from everything particular (the universal element), and the ability to determine itself (the particular element). I shared Hegel’s view that these two elements need to be united, but said little about this union itself.

One suggestive answer is that the free will wills freedom. Here are some relevant passages from the Introduction to the *Philosophy of Right*:

Only when the will has itself as its object is it *for itself* what it is *in itself* (PR: §10, p.44)

The will in its truth is such that what it wills, i.e. its content, is identical with the will itself, so that freedom is willed by freedom. (PR: §21Z, p.53)

The will which has being in and for itself is truly infinite, because its object is itself and therefore not something which it sees as other or as a limitation; on the contrary it has merely returned into itself as object. (PR: §22, p.53)

The absolute determination or, if one prefers, the absolute drive, of the free spirit (see §21) is to make its freedom into its object – to make it objective both in the sense that it becomes the rational system of the spirit itself, and in the sense that this system becomes immediate actuality (see §26). This enables the spirit to be for itself, as Idea, what the will is in itself. The abstract concept of the Idea of the will is in general *the free will which wills the free will.* (PR: §27, p.57)

One reading of these passages is as follows. The free will, considered in its universal element, abstracts from everything particular or determinate. But by itself, this is empty. The free will then finds determinacy in the free will itself. In willing freedom, it does not lose its universal element, but also finds determinacy.

Hegel himself attributes this argument to Kant,[[209]](#footnote-209) as does Korsgaard (1996a: 29-31). Korsgaard calls it the Argument from Spontaneity. It is the conclusion to her Regress Argument, where as we saw earlier (§1.3), we back up from some desire, and then keep backing up until only the form of the law remains. Korsgaard thinks that this shows that:

Making the categorical imperative its principle does not require the free will to take an action – it is already its principle. (Korsgaard 1996a: 167)

Kant concludes that the categorical imperative *just is* the law of a free will. It does not impose any external constraint on the free will’s activities, but simply arises from the nature of the will. It describes what a free will must do in order to be a free will. It must choose a maxim that it can regard as a law. (Korsgaard 2008: 320)

The argument is elegant, but unhelpful.

The crucial issue concerns the moral status of freedom. If freedom is valuable, then the free will acts freely in responding to the moral value of freedom. If freedom is not (independently) valuable, then we face a now familiar trouble: the free will acts morally to preserve its own (non-moral) status as free. This makes morality ultimately dependent upon a non-moral end. One could claim that the relation here is constitutive rather than instrumental,[[210]](#footnote-210) but that runs up against other (now familiar) problems, namely that it still does not address the value of freedom.

Another major shortcoming of this approach is that even if, as free I am somehow committed to promoting my own freedom, it is unclear how this commitment extends to the freedom of others. Patten recognises this problem, and proposes an answer on Hegel’s behalf, where mutual recognition is required for free agency, and thus in promoting others’ freedom, we promote our own (1999: 102-3). We can accept Hegel’s claims about the importance of mutual recognition, but the question still remains as to why we should promote others’ freedom. If we should do so because it promotes our freedom, we end up doing the right thing for the wrong reasons. Patten wants to make this relationship constitutive, where it is constitutive of promoting our own freedom that it promotes the freedom of others, but that does not solve the problem. Helping others has value, and that is why we should do it, not because it ultimately promotes (or is constitutive of) our own freedom.

Rational nature occupies a central role in Kant’s ethics. And, as I just hope to have shown, the various attempts to account for this without introducing moral content fall short. Once more, I take this to count in favour of a realist treatment of the value of rational nature.

I would like to now end this chapter by considering one final objection to my modest approach. In moving away from Kant’s formalism and Kantian constructivism, I have not offered a proof of, or foundation for, the moral law. The formalist and constructivist, for all of their flaws, at least attempt this, where I do not. This might appear to be a lacuna in my approach, but it is a lacuna worth leaving be.

In the Introduction (§3), I argued that *Groundwork* III was best read as not attempting a proof of the moral law. The starting point of the *Groundwork* is common human reason. Morality appears to be unconditionally normative (and binding), and in *Groundwork* I and II, Kant identifies and analyses its principle. In *Groundwork* III, he then turns to corroborate it. In doing so, however, Kant does not address the moral sceptic who requires non-moral reasons for acting morally.[[211]](#footnote-211) And this is a good thing; if Kant were to provide non-moral reasons for acting morally, he would renounce the idea of duty for duty’s sake, and reduce morality to hypothetical imperatives – if you want these non-moral things, then you ought to act morally. Instead, Kant is addressing someone who has followed him through the book, but is unsure of whether *we* are capable of standing under the moral law (due to our constitution).

The itch to provide something more than this is understandable, but it is an itch we should refrain from scratching. Korsgaard and Sensen both want to press this issue, but fail to motivate it. Sensen, for instance, claims that:

Kant’s point is […] that even if there were […] an objective value, one would still have to give an account of […] why one should be *motivated* to follow it. […] Kant […] asks in a Humean fashion: Even if there were a value, why should one be motivated to pursue it? (Sensen 2009: 268)[[212]](#footnote-212)

There is a normative question here – why should we pursue objective ends? – and a motivational one – how could we be motivated to pursue objective ends? I have already addressed these issues, but would like to say something more about the normative one, namely why *should* we be motivated to follow or pursue objective values?

Pace Sensen, this is not Kant’s question. And that is a good thing. For one, Sensen’s question comes dangerously close to asking for non-moral reasons to be moral. Secondly, the question has an easy answer. Objective values are precisely things that ought to be pursued. Accordingly, we ought to be motivated to pursue them.

Korsgaard offers a similar argument:

The argument I am making now is, in a way, the ultimate extension of the open question argument. If it is just a fact that a certain action would be good, a fact that you might or might not apply in deliberation, then it seems to be an open question whether you should apply it. (Korsgaard 2008: 317)

This question is not so open. Helping a rational agent in need of help is good. Therefore, one should help a rational agent in need of help. Of course, one might or might not choose to do so, and there might be other considerations that weigh against it. The question is open in this minimal sense, but none of that means that objective values do not prescribe behaviour to us (nor does it mean that they cannot motivate us).

This ties back into our earlier discussion of freedom (from §1.4), where we saw that:

Emancipating our intellectual activity from the pseudo-authority of the merely given cannot be freeing ourselves from the authority of reasons. (McDowell 2009: 104)

In claiming that the fact that an action is good still leaves open the question of whether you should act accordingly, Sensen and Korsgaard appear guilty of this.

At one point, Korsgaard offers the following complaint against the realist:

Having discovered that obligation cannot exist unless there are actions which it is necessary to do, the realist concludes that there are such actions, and that they are the very ones we have always thought were necessary, the traditional moral duties. […] The realist […] places the necessity where he wanted to find it. And the argument cannot even get started, unless you assume that there are some actions which it is necessary to do. (Korsgaard 1996b: 34)

Korsgaard sets the bar too high in this passage.[[213]](#footnote-213) I am happy to accept what Korsgaard (at least in this passage) does not seem happy to, namely that there are some actions which it is necessary to do. Of course, someone might doubt this, but that is another type of moral scepticism that neither Kant, nor I, address. I follow Kant in conceiving of morality as unconditionally normative, and am looking for a meta-ethics that can accommodate this and vindicate this conception of morality.

As we saw earlier (§§1.4-1.5), neither Kant’s formalism nor Kantian constructivism fit this bill. Having now cast doubt on various anti-realist attempts to account for the value of rational nature, I conclude in favour of Kantian moral realism, as a promising a way of vindicating our conception of morality as unconditionally normative.

# Conclusion

Over these first two chapters, I have defended the premises of the following argument:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

I have taken the text of *Groundwork* III as my starting point, and have attempted to develop Kant’s key insights concerning the relationship between reason and freedom, and freedom and the moral law.

In the first chapter, I argued that reason involves spontaneity. In this chapter, I have explored the connections between freedom, autonomy and morality. I distanced myself from an ambitious account of these connections, where freedom in some way establishes morality. In its place, I have offered a modest alternative, where freedom allows us to stand under the moral law. This is how I understand the argument of *Groundwork* III: if we are rational (wills), we are free; and if we are free, we are able to act morality.

Having now set out, modified and defended the first two premises of our argument, we can turn to the two crucial objections.

3 – Practical Reason

We are now half way through this thesis. I have spent the first half outlining and defending the two premises of the following argument:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

In the second half, I will consider what I take to be the two main objections to this argument. In the next chapter, I investigate whether Kant equivocates in his use of the term ‘free’ between the two premises, that is, whether the freedom involved in rational agency is the same as the freedom that the moral law requires. In this chapter, I consider whether this argument applies to us, that is, whether we are rational beings with wills.

The starting point of our argument is a rational will. From here, Kant attempts to move to freedom to our standing under the moral law. One important objection to this whole approach concerns whether we are entitled to think of ourselves as having wills in the first place.

Over the course of this chapter, I will argue that we are entitled to think of ourselves as rational beings with wills. There are two main parts to this. First of all, we need to show that free will is possible. After that, we need to show that it is actual. Kant thinks that transcendental idealism (and only transcendental idealism) makes free will possible. As for what shows that it is actual, in the second *Critique*, the moral law plays this role, whereas in *Groundwork* III, it is something else. I find Kant’s answer in the fifth subsection of *Groundwork* III, where he appeals to our consciousness of the causality of reason. This evidence for our free will is simple – we are conscious of our agency. I claim that this helps entitle us to think of ourselves as rational beings with wills.

After defending *Groundwork* III on this score, I then follow Ameriks to note that Kant moved away from this approach in the second *Critique*, as he came to realise that any consciousness of our agency conflicts with the epistemological limits of transcendental idealism. I think that does not count against the approach of *Groundwork* III, but instead against transcendental idealism itself. I make this case, and offer my own account of how free will is possible.

The chapter falls into two main sections. In the first, I consider the evidence for the claim that we are rational agents. I begin by laying out the problem, detailing Kant’s worry that we might not have wills (§1.1). I then offer some general reflections about doubt (§1.2), before arguing that the phenomenology of our agency reveals to us that we have free wills (§1.3). In the second section, I turn to outline how free will is possible, and distance myself from transcendental idealism in the process (§2).

# Evidence for our Agency

The topic of this chapter is whether we are rational beings with wills. A good place to start is with what Kant means by a ‘will’. Unfortunately, we are not helped by an ambiguity in his use of the term. He famously introduces the will in *Groundwork* II as follows:

Every thing in nature works according to laws. Only a rational being has the capacity to *act according to the representation* of laws, i.e. according to principles, or a *will*. Since *reason* is required for deriving actions from laws, the will is nothing other than practical reason (IV: 412. 26-30).

Here, a will looks like the capacity to act according to principles. Elsewhere though, Kant claims that a will is the capacity “so to act that the principle of actions conforms with the essential constitution of a rational cause, that is, with the condition of universal validity of a maxim as a law” (IV: 458. 12-6). This ambiguity permeates the text (and the surrounding literature), and gives us two basic options as to what Kant means by a rational being with a will: 1) a being who can act on principles (or maxims); 2) a being who has the capacity to conform their principles to the moral law. Otherwise expressed, the two options are: 1) a being for whom reason is practical; 2) a being for whom *pure* reason is practical.[[214]](#footnote-214) Here, I understand ‘will’ in the first sense, and thus take the question at hand to be whether reason is practical.[[215]](#footnote-215) Of course, there is the further question of whether *pure* reason is practical, but that is the topic of our next chapter.

The *Groundwork* is a book about rational beings with wills, beings whose reason is practical. That Kant talks about rational beings *with* wills suggests that he was open to the possibility of rational beings *without* wills.[[216]](#footnote-216) Furthermore, we might be such beings. At the start of the third section of *Groundwork* III, he raises this very worry:

We last traced the determinate concept of morality back to the idea of freedom; which we could not, however, prove as something actual even in ourselves or in human nature; we saw only that we must presuppose it if we want to think of a being as rational and endowed with consciousness of its causality with regard to its actions, i.e. with a will (IV: 448. 25 – 449. 3)

The move from reason to freedom (that we considered in chapter 1) only helps Kant’s case in *Groundwork* III if we are entitled to think of ourselves as having wills.[[217]](#footnote-217)

Henrich has argued that this is pivotal to understanding Kant’s project in *Groundwork* III, and its failure. He insists that, for Kant, there is a crucial difference between a rational being and a rational being with a will (1975: 312-4).[[218]](#footnote-218) Considered as such, our argument begs the question at hand. Its starting point is a rational will, where the first premise states that ‘A rational will is a free will’. From here, we move through freedom to the moral law. Henrich is relatively comfortable with this move, but insists that the crucial issue of whether or not we are the type of rational beings that have wills still remains (1975: 329). This critique has been very influential, and this issue is often taken to count decisively against *Groundwork* III.[[219]](#footnote-219)

In what follows, I will trace the history of how Kant came to worry that we might not have wills. In doing so, I begin with Kant’s pre-critical treatment of rational psychology (§1.1). This brings the exact nature of the problem into focus, and also helps reveal a solution. As we will see, the problem is that we *might* lack wills. I then offer some general epistemological reflections about doubt and inquire as to what provide grounds for us to doubt that we actually do lack wills (§1.2). I find the grounds for doubt in the threat of natural necessity. I then proceed to provide some evidence for our agency (§1.3), before attempting to disarm the threat of natural necessity in the next section (§2).

## 1.1 Kant’s Pre-Critical Argument for Freedom

Rational psychology is a non-empirical science of the soul. The soul is to be studied not through experience, but *a priori*, from reason alone.[[220]](#footnote-220) In his pre-critical lectures on this (in the mid-1770s), Kant sees the sole text for this study as the I:

Now when we speak of the soul *a priori*, then we will talk of it only to the extent we can derive all from the concept of the I […] We will thus cognize *a priori* no more of the soul than the I allows us to cognize. (XXVIII: 266)

In these lectures, Kant thinks that studying the I can take us a long way. We cognise of the soul that it is: 1) substance; 2) simple; 3) single; 4) free (XXVIII: 266-8).

The fourth of these is our concern here. Kant claims that: “The soul is a being which acts spontaneously […] i.e., the human soul is free in the transcendental sense” (XXVIII: 267). When it comes to arguing for this, “the I must again help out” (XXVIII: 268). It does so as follows:[[221]](#footnote-221)

[…] the I proves that I myself act; *I* am a principle and no thing which has a principle, *I* am conscious of determination and actions, and such a subject that is conscious of its determinations and actions has absolute freedom […] When I say: I think, I act, etc., then either the word I is applied falsely , or I am free. Were I not free, then I could not say: I do it, but rather I would have to say: I feel in me a desire to do, which someone has aroused in me. But when I say: I do it, that means spontaneity in the transcendental sense. (XXVIII: 268-9)

Here, Kant engages in a piece of rational psychology, attempting a proof of the (transcendental) freedom of the soul on the basis of the I (and the phenomenology of agency).[[222]](#footnote-222) I think that this strategy is important, and shall return to consider it again shortly (§1.2). For now though, I want to follow Kant through the critical turn.

In the first *Critique*, Kant moves away from rational psychology, to instead present four paralogisms of this doctrine (A 341-405).[[223]](#footnote-223) He complains that rational psychology erroneously inflates the ‘I think’, the mere “vehicle of all concepts” (A 341/B 399), into various ontological claims. The ‘I think’ is a necessary condition of experience (A 106-7[[224]](#footnote-224)), but Kant no longer thinks this establishes anything about a soul.

In his own words, rational psychology is:

[…] falsely held to be a science of pure reason about the nature of our thinking being. At the ground of this doctrine we can place nothing but the simple and in content for itself wholly empty representation **I**, of which one cannot even say that it is a concept, but a mere consciousness that accompanies every concept. Through this I, or He, or It (the thing), which thinks, nothing further is represented than a transcendental subject of thoughts = x (A 346/B 404)

So much for the soul. With the critical turn, Kant ferociously attacks the arguments – and indeed the entire discipline – of rational psychology. The details of this are not our concern here. What is important for our purposes, is that in the paralogisms, Kant no longer considers freedom. As we saw above, in his earlier treatment of rational psychology, Kant offered a proof of our freedom. In Kant’s discussion of the paralogisms however, we neither receive such a proof nor a critique of such a proof.[[225]](#footnote-225)

In the first *Critique*, we do get some discussion of freedom: the understanding is characterised as spontaneity (A 50-1/B 74/5, A 68/B 93, A 126[[226]](#footnote-226)); we get an account of the possibility of freedom in the third antinomy (A 444-51/B 472-9); and we are told in the Canon that practical freedom – “independence of the power of choice from **necessitation** by impulses of sensibility” (A 534/B 562) – can be proved by experience. However, we are no longer offered any proof of our transcendental freedom – “independence of [...] reason itself […] from all determining causes of the world of sense” (A 803/B 831).[[227]](#footnote-227)

While Kant refrained from attempting a proof of (transcendental) freedom in the first *Critique*, this was not an option in the *Groundwork*. If we are not free, then the moral law will be a phantasm for us.

I want to consider in what respects Kant abandons his pre-critical proof of freedom, and why. One key departure is that Kant no longer talks of the soul. Another is that Kant comes alive to the possibility that we might be rational, but lack wills. Recall our lengthy quote from the paralogisms, where Kant mentions “this I, or He, or It (the thing), which thinks” (A 346/B 404). This brief description of the transcendental subject of thoughts forms the basis for an entire paper of Sellars’, where he takes the “It (the thing), which thinks” to express the possibility of “an *automaton spirituale* or *cogitans*, a thinking mechanism” (1971: 24-5). Kant does seem open to the possibility that reason does not have causality in the first *Critique*:[[228]](#footnote-228)

Now that this reason has causality […] (A 547/B 576)

Now let us stop at this point and assume it is at least possible that reason actually does have causality in regard to appearances [… ] (A 548/B 582).

Here we can further clarify why rational psychology no longer provides Kant with a proof of our freedom. Rational psychology is a “putative science, which is built on the single proposition **I think**” (A 342/ B 400); “**I think** is thus the sole text of rational psychology” (A 343/ B 401). With the ‘I think’ as the sole text, we get the spontaneity of the understanding, but still face the possibility that we lack wills; we could be thinking mechanisms.[[229]](#footnote-229)

Following on from his dismissal of rational psychology, and worried by the possibility that we might lack wills, Kant no longer employs his pre-critical rationalist argument for our freedom. In the first *Critique*, he says little about this, instead setting out how transcendental idealism makes (transcendental) freedom possible – I will return to this shortly (§2.1). However, the *possibility* of freedom is not enough, and thus in *Groundwork* III, Kant offers a new argument for the actuality (or necessity – more on this later) of our freedom. Before we move on, it will help to clarify the precise nature of the problem at hand.

I take Allison to have provided the most comprehensive treatment of this problem (including discussion of both Henrich and Sellars[[230]](#footnote-230)). At one point, he characterises it as follows:

[…] the problem is that, whereas the consciousness of possessing reason as a theoretical capacity is arguably self-certifying, on the familiar Cartesian grounds that any doubt concerning the possession of this capacity already presupposes it, […] this immunity to doubt does not extend to reason as a *practical* capacity. (Allison 2011, 329)

I think that this is a fascinating characterisation of the problem, and will attempt to unpack it in the next section. In doing so, I will offer some general reflections upon doubt, such that my solution falls out of the problem.

## 1.2 Grounds for Doubt

In the introduction to this thesis, I characterised Kant’s project in *Groundwork* III as a deduction. I also followed Henrich to note that a deduction has a particular meaning for Kant, namely an attempt to secure an entitlement to something we take ourselves to possess. I remarked that a deduction is only needed when our claim to possession is put in doubt, and quoted Henrich (in a footnote):

A deduction is called for, whether in cognition or in court, when the title to a right is in dispute. If the doubt is not explicit there is no basis for the deduction. The deduction is not undertaken for the sake of amplifying cognition, but for justifying it. (Henrich 1998: 324)

I want to add something to this: A deduction is called for when doubt is grounded. If there are no reasonable grounds for doubt, a deduction is not required. A lot will hinge on this point.

With the case at hand, it is important to ask what the grounds are for doubting that we have wills. Allison remarks that:[[231]](#footnote-231)

Although one might think that Kant could dismiss such a worry as idle, on the grounds that we are conscious of our agency and it can have no effect from the practical point of view, the fact is that he did not. (Allison 2012: 115)

Allison himself follows Henrich to think that this worry counts against the approach of *Groundwork* III.[[232]](#footnote-232)

It is important however, to get clear what the grounds for doubt here are. I think the answer is simple: Kant is worried about the threat of natural necessity. If everything is determined by natural necessity, then it seems like there is no room for rational agency (at least as Kant conceives it).[[233]](#footnote-233) We can see traces of this in the fifth subsection of *Groundwork* III, where Kant talks of a “seeming contradiction” (IV: 456. 7) between freedom and natural necessity, which needs to be “convincingly eliminated” (IV: 456. 7-8). I will say more about this in the final section of this chapter (§2). For now, what matters is that it is the threat of natural necessity that provides reasons to doubt that we might have wills. And that seems like a reasonable doubt to have. If the world was entirely determined by natural necessity, then it is hard to see how we could be rational agents (in any substantial sense). These reasonable grounds for doubt call for a deduction.

Setting things up this way changes the target though. Recall Allison’s characterisation of the problem:

[…] the problem is that, whereas the consciousness of possessing reason as a theoretical capacity is arguably self-certifying, on the familiar Cartesian grounds that any doubt concerning the possession of this capacity already presupposes it, […] this immunity to doubt does not extend to reason as a *practical* capacity. (Allison 2011: 329)

Let us grant Allison that the consciousness of possessing reason as a theoretical capacity is immune to doubt. Let us also leave aside the issue of whether immunity to doubt helps with justification (which we discussed in the first chapter - §1.2). What I want to look at here is the significance of the claim that the possession of reason as a practical capacity is not immune to doubt.[[234]](#footnote-234) By itself, this is of no great significance. Mere lack of immunity to doubt does not itself provide reasonable grounds to actually doubt something; more straightforwardly, the mere possibility of doubt does not make doubt reasonable. Here, I want to distance myself from an unhelpful conception of doubt typically associated with Descartes.

Allison appears to employs something close to a Cartesian conception of doubt, which feeds into his treatment of whether or not we have wills. We can see this clearly in one of his earlier treatments of the issue. Allison begins:

Although Kant, to my knowledge never expresses himself in this way, it seems reasonable to attribute to him the view that the spontaneity of the understanding (or reason in its theoretical capacity) is self-certifying somewhat in the manner of the Cartesian *cogito*.[[235]](#footnote-235) To doubt one’s spontaneity in this sense would be to doubt that one is a thinking being; but this of course, would itself require an act of thought. The ‘I think’ must be able to accompany all my representations, even the thought that I am not really thinking. (Allison 1996: 133)

He then considers the practical analogue of this, and again claims:

By contrast, in the practical domain, not even this degree of self-certification is available. (Allison 1996: 133)

In this sphere, we can only achieve a conditional result:

Here, then, we must conclude that reflection on the ineliminable moment of spontaneity can yield only the conditional result: If I take myself to be a rational agent, […] I must necessarily regard myself as free. (Allison 1996: 133)

I will return to consider what to do with this conditional in the next section (§1.3). Returning to the passage at hand, Allison concludes:

[…] the practical necessity of acting under the idea of freedom leaves in place the epistemic possibility that I am deluded in believing that I am acting, or as Kant sometimes puts it, that my “reason has causality.” Here the Cartesian demon is more difficult to dislodge. In fact, it cannot be exorcized by any theoretical means, although it can be safely ignored from the practical point of view. (Allison 1996: 134)

As a general epistemic point, we do not need to dislodge Cartesian demons. It is possible that our reason does not have causality, just as it is possible that you are not really reading this chapter right now, but instead suffering some (unfortunate) hallucination. However, mere possibility does not provide grounds for doubt. In the case at hand, we do have grounds for doubt though – the threat of a world that operates (solely) according to natural necessity. We need to dispel this threat, and thereby remove the grounds for doubt, which will help entitle us to the claim that we are rational agents. I will attempt just this in the second part of this chapter.

Before I turn to this, I would like to say a little more about doubt. In claiming that one needs grounds for doubt, I align myself with Pragmatism, Hegel, and a certain conception of Kant’s project. The pragmatist connection is relatively clear. Peirce famously wants to distance philosophy from Cartesian doubt:[[236]](#footnote-236)

A person may, it is true, in the course of his studies, find reason to doubt what he began by believing; but in that case he doubts because he has a positive reason for it, and not on account of the Cartesian maxim. Let us not pretend to doubt in philosophy what we do not doubt in our hearts. (Peirce 1868: 29)

Taking my cue from Robert Stern, I also read Hegel in this way. Hegel begins the *Phenomenology* by discussing a “natural assumption [*eine natürliche Vorstellung*]” about philosophy’s starting point, namely that, before it gets about with actual cognition of what is, philosophy should begin by understanding cognition itself (PS: § 73, p.46). Hegel takes the motivation for such an approach to be the fear that without a precise grasp of the nature and limits of cognition, “we might grasp clouds of error instead of the heaven of truth” (PS: § 73, p.46). He objects to this as a natural starting point for philosophy:[[237]](#footnote-237)

[…] if the fear of falling into error sets up a mistrust of Science, which in the absence of such scruples gets on with the work itself, and actually cognizes something, it is hard to see why we should not turn round and mistrust this very mistrust. Should we not be concerned as to whether this fear of error is not just the error itself? (PS: § 74, p.47)

At one point, Stern brings out the point, by discussing a remark of Wittgenstein’s:

Cf. Wittgenstein’s remark in *On Certainty*, with which Peirce’s position is often compared: ‘But what about such a proposition as ‘‘I know I have a brain’’? Can I doubt it? Grounds for *doubt* are lacking! Everything speaks in its favour, nothing against it’ [… *§*4, 2.] Wittgenstein is clearly saying here that I cannot doubt ‘I have a brain’ not because there are some things that must be presupposed in thinking and this is one of them, but because the Cartesian does not give us sufficient *grounds* for questioning this belief, even though it could turn out to be false (‘Nevertheless it is imaginable that my skull should turn out empty when it was operated on’). (Stern 2009: 235-6 n74)

I think Peirce, Hegel, Stern and Wittgenstein are spot on here, and I also think that we can find traces of this line of thought in Kant. For this, I turn once more to what it means to provide a deduction.

With a deduction, we take ourselves to possess something, and something else puts this into doubt. This is not an attempt to answer a Cartesian sceptic.[[238]](#footnote-238) And that is a good thing – once you start playing that game, you are unlikely to emerge.[[239]](#footnote-239) Moreover, such scepticism lacks grounds. Consider, for example, Kant’s deduction of the categories. Kant is not trying to establish that these are immune to any possible doubt. Instead, he starts with our conception of experience, and attempts to show how this is possible. In doing so, he takes himself to have vindicated our conception of experience.

Here, I position Kant alongside Peirce and Hegel, such that he can circumvent some of the prominent early objections to his project. Maimon (1790: 42-3), for instance, accused Kant of not providing an adequate response to the (Humean) sceptic, in that Kant presupposes a rich conception of experience – which involves, for example, objects standing in causal relations – that the sceptic can deny.[[240]](#footnote-240) In this, I think Maimon is right, but more importantly, I think he is wrong to think that this is what Kant actually attempts, or what an adequate response to the sceptic requires.

We can see this by looking at Kant’s argument for freedom in the second *Critique* in a little more detail.[[241]](#footnote-241) Here is the famous (second) gallows case:

But ask him whether, if his prince demanded, 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. (V: 30. 27-35)

He goes on to call this “a fact of reason [*ein Factum der Vernunft*]” (V: 31. 24), and claims that it is “undeniable [*unleugbar*]” (V: 32. 2).

Of course though, this fact is not undeniable.[[242]](#footnote-242) It is just that Kant is not interested in that type of moral scepticism. He wants to take our moral experience at face value, and then provide a system that can vindicate it. That is a perfectly plausible approach, and one that sits well with how I understand Kant’s general strategy.

Returning to *Groundwork* III, I think it is worth considering why Henrich (and those who follow him) take the issue of whether we have wills so seriously. I admit it is *possible* that we lack wills, but these commentators seem to think that this possibility renders the approach of *Groundwork* III hopeless.

Part of the problem stems from how one conceives Kant’s project in *Groundwork* III. If one thinks that Kant is attempting to somehow establish the moral law through considerations of reason and the freedom involved in reason, then any gaps would be very problematic. If we read Kant as trying to move from the spontaneity involved in theoretical judgement to freedom to somehow thereby establish the moral law, then the possibility of possessing theoretical reason without also possessing practical reason becomes pivotal. This is exactly what troubles Henrich (1975: 312-4). But, as I laid out in the introduction (§3) and throughout the last chapter, this is not how I read *Groundwork* III. I think it is best not to read Kant as attempting to somehow prove the moral law through water-tight moves from reason to freedom to morality. Instead, I take it he is attempting to vindicate our conception of ourselves as free, such that morality is not a phantasm for us.

## 1.3 The Phenomenology of Reason

Moving on from the general epistemic discussion, we can now consider whether we have wills. There are two distinct tasks at hand: we need to show that free will is possible, and also that it is actual (or necessary). Kant himself says as much in the fifth subsection of *Groundwork* III:

To presuppose this freedom of the will, moreover, is not only […] quite easily *possible* (as speculative philosophy can show), in the case of a rational being conscious of its causality through reason, and hence of a will (which is distinct from desires) it is also without any further condition *necessary* to presuppose it practically, i.e. in the idea in all the actions he chooses, as their condition. (IV: 461. 17-25)

This is an important passage. Speculative philosophy shows that free will is possible. This is accomplished by transcendental idealism in the first *Critique* (which I will look at in the next section – §2). Free will however, is not only possible, but also necessary, and specifically, necessary for “a rational being conscious of its causality through reason, and hence of a will” (IV: 461. 22-4).[[243]](#footnote-243) I contend that this talk of our *consciousness of the causality of our reason* is crucial. Consider for example, the following related passages from the fifth section of *Groundwork* III:

[…] the legitimate claim even of common human reason to freedom of the will is founded on the consciousness and the granted presupposition of the independence of reason from merely subjective determining causes (IV: 457. 4-7)

[Freedom ...] holds only as a necessary presupposition of reason in a being that believes itself to be conscious of a will (IV: 459. 9-14)

Kant’s claim here is simple, but effective. We are rational agents, and we are conscious of our reason being practical. This might appear too quick, but I think it is the right way to go. Indeed, what better source of evidence could there be for our being rational agents, than the robust phenomenology of agency?[[244]](#footnote-244) If this appears unsatisfactory, then there will have to be reasons to doubt this robust phenomenology. And in this case, there is, namely the threat of natural necessity. I will attempt to disarm this threat in the next section.

Before we get to that though, I would like to say a bit more about the phenomenology of agency – in Kant’s terms, our consciousness of the causality of our reason. With this, we can finally situate the final two non-moral arguments for freedom that I find in *Groundwork* III. The four arguments are:

1. The claim that freedom must be presupposed as a property of the will of all rational beings (IV: 447. 26 – 448. 22).
2. An appeal to transcendental idealism and the activity involved in (discursive) cognition (IV: 451. 1-36).
3. The claim that reason’s capacity for ideas reveals a “spontaneity so pure” (IV: 452. 18) that it thereby takes us far beyond anything sensibility can afford us, and marks us out as members of an intelligible world (IV: 452. 6-22).
4. An appeal to the consciousness of the causality of our reason (IV: 457. 4-7, IV: 457. 22-4, IV 459. 9-14, IV: 461. 17-25).

We looked at A and C in the first chapter, where I claimed that A provided a link between reason and spontaneity. Now, we find ourselves attempting to show that we are rational agents, and for this I think D is the main claim. I think that B also helps.

Here is the relevant passage from B (in the third section of *Groundwork* III):

[…] beyond these characteristics of his own subject that are composed of nothing but sundry appearances he must necessarily assume something else lying at its foundation, namely his I, such as it may be in itself; and with respect to mere perception and receptivity to sensations he must thus count himself as belonging to the *world of sense*, but with regard to what there may be of pure activity in him (what reaches consciousness not by affection of the senses, but immediately) as belonging to the *intellectual world*, with which he is yet no further acquainted. (IV: 451. 28-36)

Here, we see Kant appeal to the activity involved in (discursive) cognition. We are not only receptive in our cognition, but active – “what there may be of pure activity […] what reaches consciousness not by affection of the senses, but immediately” (IV: 451. 33-5). This relates back to his pre-critical argument for freedom, where, as we have just seen, in his lectures on Rational Psychology, he argues that:

*I* am conscious of determinations and actions, and such a subject that is conscious of its determinations and actions has absolute freedom. That the subject has absolute freedom because it is conscious of itself, that proves it is not a subject being acted upon, but rather [one] acting. (XXVIII: 268-9)

Kant makes a clear appeal to our consciousness of our agency. And in the first *Critique*, he makes a similar claim:[[245]](#footnote-245)

Yet the human being, who is otherwise acquainted with the whole of nature solely through sense, knows himself also through pure apperception, and indeed in actions and inner determinations which cannot be accounted at all among impressions of sense; he obviously is in one part phenomenon, but in another part, namely in regard to certain faculties, he is a merely intelligible object, because the actions of this object cannot at all be ascribed to the receptivity of sensibility. (A 546-7/B 574-5)

In these passages, we see phenomenology do some work for Kant. We are conscious of the activity of our reason, in both the theoretical and the practical case. [[246]](#footnote-246)

There is a crucial objection to this approach though, one that is pressed by Ameriks, Grenberg, and that we can find Kant himself make in the second *Critique*. The objection is that any talk of consciousness of the causality of our reason violates the epistemological limits of transcendental idealism. For this, we turn to the second section of the chapter.

# The Possibility of Freedom

I want to now introduce transcendental idealism, which Kant thinks makes freedom possible (§2.1). I then go on to outline how the epistemological restrictions that come with transcendental idealism preclude the phenomenological approach I have just offered, but will conclude “so much the worse for transcendental idealism”. I will argue that, while transcendental idealism makes freedom possible, it precludes other essential aspects of a satisfactory account of freedom, which I bring out through considering the possibility of recognising other rational agents (§3.2). I conclude that transcendental idealism fails to provide to provide an adequate account of freedom. In its place, I sketch an alternative picture of how freedom is possible, one that locates freedom within, rather than outside of nature (§3.3).

## 2.1 Transcendental Idealism

I began the previous section with the following quote from the fifth subsection of *Groundwork* III:

To presuppose this freedom of the will, moreover, is not only (without falling into contradiction with the principle of natural necessity in the connection of appearances in the world of sense) quite easily *possible* (as speculative philosophy can show), in the case of a rational being conscious of its causality through reason, and hence of a will (which is distinct from desires) it is also without any further condition *necessary* to presuppose it practically, i.e. in the idea in all the actions he chooses, as their condition. (IV: 461. 17-25)

In the last section, I attempted to flesh out the second half of this passage, discussing the consciousness of the causality of our reason, which, I contend, shows us that our freedom is actual. In this section, I will turn my attention to the first part of the passage, where Kant claims that to presuppose freedom of the will is “quite easily *possible* (as speculative philosophy can show)” (IV: 461. 20-1), “without falling into contradiction with the principle of natural necessity in the connection of appearances in the world of sense” (IV: 461. 18-20). For this, we need to introduce transcendental idealism.

Transcendental idealism enters the *Groundwork* in the third subsection of *Groundwork* III. We have already seen that it plays a role in two of Kant’s arguments for freedom – the activity involved in discursive cognition (B) and reason’s capacity for ideas (C) – but it also makes freedom possible.

In the third subsection of *Groundwork* III, Kant wonders whether his argument hitherto might be circular.[[247]](#footnote-247) There are two parts to this. The first is his thought that he might be “presupposing the idea of freedom only for the sake of the moral law” (IV. 453. 5-6). As I noted in the introduction (§2), in *Groundwork* III, Kant wants a non-moral proof of our freedom. The second concern – Kant’s main worry – is ours in this chapter, namely that he has merely assumed that we have wills.[[248]](#footnote-248)

Kant seems to offer transcendental idealism as the solution to both of these worries. He writes that:

[…] there still remains for us one way out, namely to try: whether when, through freedom, we think of ourselves as causes efficient a priori we do not take up a standpoint that is different from when we represent ourselves according to our actions as effects that we see before our eyes. (IV: 450. 30-4)

Kant appears to think that transcendental idealism helps show that freedom is both necessary – arguments B and C – and possible. He takes up the issue of its possibility in the fifth subsection of *Groundwork* III. In the remainder of this section, I want to work my way carefully through this text to get Kant’s claim into clear sight.

Kant begins the fifth subsection of *Groundwork* III, by detailing a “dialectic of reason” (IV: 455. 28) that arises between freedom and natural necessity. We have seen why freedom is necessary, but Kant reminds us that:

On the other side it is equally necessary that everything that happens should without fail be determined according to laws of nature (IV: 455. 16-18)[[249]](#footnote-249)

Kant continues:

Even though there arises from this a dialectic of reason, as with regard to the will the freedom ascribed to it seems to be in contradiction with natural necessity and, at this intersection, reason for *speculative purposes* find the route of natural necessity much more even and useful than that of freedom: yet for *practical purposes* the footpath of freedom is the only one on which it is possible to make use of one’s reason in our behaviour; which is why it is just as impossible for the subtlest philosophy as for the commonest human reason to rationalize freedom away. It must therefore presuppose: that no true contradiction can be found between freedom and natural necessity of just the same human actions, for it cannot give up the concept of nature, any more than that of freedom. (IV: 455. 28 – 456. 6)

Here, Kant seems to suggest that both natural necessity and freedom are equally necessary, and that we cannot give either up. However, he immediately follows this passage with the important qualification:

However, this seeming contradiction must at least be convincingly eliminated, even if one should never be able to comprehend how freedom is possible. For if even the thought of freedom contradicts itself or nature, which is equally necessary, it would have to be given up altogether in favour of natural necessity. (IV. 456. 7-11)

It is not enough to just show that freedom is practically necessary. Kant recognises that we also need to eliminate the seeming contradiction between freedom and natural necessity.

He then puts forth the following famous passage:

But it is impossible to steer clear of this contradiction if the subject who deems himself free were to think of himself *in the same sense*, or *in just the same relation* when he calls himself free, as when he takes himself to be subject to the law of nature with respect to the same action. That is why it is an indispensable task of speculative philosophy: at least to show that its deception concerning the contradiction rests in this, that we think a human being in a different sense and relation when we call him free from when we take him, as a piece of nature, to be subject to its laws, and that both not only can very well coexist, but also must be thought as *necessarily united* in the same subject, because otherwise no ground could be stated why we should burden reason with an idea that, even if it can *without contradiction* be united with another that is sufficiently validated, still entangles us in a business that puts reason in its theoretical use in a very tight corner. (IV: 456. 12-27)

I will return to discuss how we are to unite these two things in the same subject shortly. For now, I just want to note that Kant thinks that this is “an indispensable task of speculative philosophy” (IV: 456. 16-7). He goes on to explain why:[[250]](#footnote-250)

This duty, however, is incumbent upon speculative philosophy only so that it may clear the way for practical philosophy. Thus it is not left to the philosopher's discretion whether he wants to remove the seeming conflict, or leave it untouched; for in the latter case the theory about this is a bonum vacans [or vacant good[[251]](#footnote-251)], of which the fatalist can with good reason seize possession and chase all moral science from its supposed property as possessing it without title. (IV: 456. 29-33)

This is the sceptical challenge that Kant addresses in *Groundwork* III. He has to show that freedom is possible. Otherwise, morality will be nothing more than a phantasm for us.

Returning to the text, Kant remarks that:

[T]he settlement of that dispute does not belong to it [practical philosophy] at all (IV: 456. 34-5)

Instead, practical philosophy:

[…] demands of Speculative Reason that she put an end to the discord in which she entangles herself in theoretical questions, so that Practical Reason may enjoy rest and security from external attacks that could bring into dispute the ground on which she wants to settle. (IV: 456. 36 – 467. 3)

After setting up the problem, Kant gives his solution:

But the legitimate claim even of common human reason to freedom of the will is founded on the consciousness and the granted presupposition of the independence of reason from merely subjectively determining causes, which together one and all constitute what merely belongs to sensation, and hence under the general label of sensibility. (IV: 457. 4-9)

This aligns with the reading that I have offered. We are conscious of the causality of our reason (the previous section of this chapter), and reason involves an independence from subjectively determining causes (the result of chapter 1). This is what reveals that we are free.

Kant then sets out how transcendental idealism makes this possible:[[252]](#footnote-252)

A human being who considers himself in this way as an intelligence thereby puts himself in a different order of things and in a relation to determining grounds of an entirely different kind, when he thinks of himself as an intelligence endowed with a will, and consequently with causality, than when he perceives himself as a phenomenon in the world of sense (which he actually is as well) and subjects his causality, according to external determination, to laws of nature. Now, he soon becomes aware that both can, and indeed even must, take place at the same time. For that *a thing in the appearance* (belonging to the world of sense) is subject to certain laws from which just the same *as a thing* or a being *in itself* is independent, contains not the least contradiction; however, that he must represent and think of himself in this twofold way rests, as regards the first on consciousness of himself as an object affected through the senses, as regards the second on the consciousness of himself as an intelligence, i.e. as independent of sensuous impressions in the use of reason (hence as belonging to the world of understanding). (IV: 457. 9-24)

Here, we have two different orders of things: the world of sense, which is subject to the laws of nature; and the world of understanding, which is independent of such laws.[[253]](#footnote-253) And we must represent and think of ourselves in both of these ways, because we are conscious of ourselves as an object affected through the senses, and also conscious of ourselves as intelligence, that is “independent of sensuous impressions in the use of reason” (IV: 457. 23-4).

Kant begins the next paragraph: “That is why a human being presumes for himself a will” (IV: 457. 25). We now have Kant’s answer to the question of this chapter: What entitles us to think of ourselves as having wills? Transcendental idealism makes it possible that there are two orders of things: one determined by natural necessity, and one not. And our consciousness of the causality of our reason entitles us to think of ourselves as belonging to the latter.

This is the heart of Kant’s argument for freedom in *Groundwork* III. Transcendental idealism makes freedom possible, and the consciousness of the causality of our reason entitles us to hold that it is actual.

In what follows, I will consider why Kant abandoned the approach of *Groundwork* III. I follow Ameriks to claim that Kant came to worry that the argument I have just presented conflicts with the epistemological limits of transcendental idealism. In the end though, I argue that this counts against transcendental idealism instead of the approach of *Groundwork* III.

## 2.2 Against Transcendental Idealism

### Kant’s Theory of Mind

Karl Ameriks famously argues that Kant’s views on freedom “betray very strong and important rationalist sympathies well into the 1780s” (2000: 189). Perhaps most strikingly, he makes the case that in *Groundwork* III, Kant attempted a “very strong rationalist argument for freedom” (Ameriks 2000: 191). He reads this as Kant’s last attempt to construct a rationalist argument for our (transcendental) freedom, and diagnoses it as follows:[[254]](#footnote-254)

[Kant] felt compelled in [*Groundwork* III … ] to attempt a very strong rationalist argument for freedom. I will argue that once that attempt was worked out as far as it could be, it suffered shipwreck for it conflicted with the critical strains that were being developed simultaneously in Kant’s theory of mind and self-knowledge (Ameriks 2000: 191).

Recall the following passages (D):

[…] the legitimate claim even of common human reason to freedom of the will is founded on the consciousness and the granted presupposition of the independence of reason from merely subjective determining causes (IV: 457. 4-7)

[…] on the consciousness of himself as an intelligence, i.e. as independent of sensuous impressions in the use of reason (IV: 457. 22-4)

[Freedom ...] holds only as a necessary presupposition of reason in a being that believes itself to be conscious of a will (IV: 459. 9-14)

To presuppose this freedom of the will, moreover, is not only […] quite easily *possible* (as speculative philosophy can show), in the case of a rational being conscious of its causality through reason, and hence of a will (which is distinct from desires) it is also without any further condition *necessary* to presuppose it practically, i.e. in the idea in all the actions he chooses, as their condition. (IV: 461. 17-25)

I have argued that in these passages Kant makes a crucial appeal to the consciousness of the causality of our reason.

Ameriks agrees, but contends that this appeal violates the strict epistemological limits of transcendental idealism.[[255]](#footnote-255) In claiming that we are conscious of the causality of our reason, Kant seems to be assuming that we are somehow acquainted with the transcendental subject, that we can (in some sense) experience our activity. Ameriks argues that Kant became aware of this after the publication of the *Groundwork*, and thus:

[I]n the second *Critique* (1788) Kant had to recast his treatment of freedom radicallyso as to be in line with the more severe limits on self-knowledgethat he had come to stress in the second edition revisions of the first *Critique* (1787). (Ameriks 2000: 191)

As for *Groundwork* III, Ameriks writes:

I think it possible and preferable to take the arguments of [*Groundwork* III …] at face value, and to see them merely as a slightly more self-conscious continuation of some relatively crude beliefs about freedom that Kant had held for some time and simply had not gotten around to submitting to a thorough critique (2000: 214)

I agree with Ameriks that Kant has not yet seen that this approach conflicts with transcendental idealism. I also follow Ameriks in thinking that Kant became clear of this shortly after the publication of the *Groundwork*.

### The Second *Critique*

We can see Kant’s change of mind in the second *Critique* (1788). Here is the so-called ‘great reversal’:

I ask instead from what our *cognition* of the unconditionally practical *starts,* whether from freedom or from the practical law. It cannot start from freedom, for we can neither be immediately conscious of this, since the first concept of it is negative, nor can we conclude to it from experience, since experience lets us cognize only the law of appearances and hence the mechanism of nature, the direct opposite of freedom. It is therefore the *moral law,* of which we become immediately conscious (as soon as we draw up maxims of the will for ourselves), that *first* offers itself to us and, inasmuch as reason presents it as a determining ground not to be outweighed by any sensible conditions and indeed quite independent of them, leads directly to the concept of freedom. (V: 29. 28 – 30. 3)

Ameriks diagnoses Kant’s reversal as follows:

[…] out of necessity he had to renounce the attempt metaphysically to ground freedom or morality; and that, finally, it was only to make a virtue of necessity that (given his deepest beliefs of longest standing) he felt it was proper to announce that without the moral law 'we would never have been justified in assuming anything like freedom’ (Ameriks 2000: 211)

The necessity Ameriks speaks of concerns the epistemic limits of transcendental idealism. For Kant, we can have no experience or intuition of the world of understanding (the noumenal). Moreover, he conceives the world of sense as entirely determined by natural necessity, and thus we cannot experience freedom there either. Kant is clear on both of these points throughout the second *Critique*.

This is especially clear in the following passage:

That this is the true subordination of our concepts and that morality first discloses to us the concept of freedom, so that it is *practical reason* which first poses to speculative reason, with this concept, the most insoluble problem so as to put it in the greatest perplexity, is clear from the following: that, since nothing in appearances can be explained by the concept of freedom and there the mechanism of nature must instead constitute the only guide (V: 30. 9-14)

The claim that we cannot experience freedom in nature does not conflict with the approach of the *Groundwork*. In the *Groundwork*, Kant also claims that we can have no experience of freedom.[[256]](#footnote-256) However, as we have seen, in *Groundwork* III (as well as his pre-critical lectures on rational psychology and the A-edition of the first *Critique*), Kant makes a crucial appeal to our consciousness of the causality of reason. In the second *Critique*, he retreats from this position.

He also insists that we can have no intuition of our freedom. We can see this in the fact of reason passage:

Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical, although it would be analytic if the freedom of the will were presupposed; but for this, as a positive concept, an intellectual intuition would be required, which certainly cannot be assumed here. (V: 31. 24-31)

Here, Kant is clear that we have no consciousness of freedom, which he claims would require an intellectual intuition.[[257]](#footnote-257) Elsewhere, he claims that we are not conscious of the freedom of our will independently of our consciousness of the moral law, and that any such consciousness (of our own freedom) would require a “special intuition of itself [*einer besondern Anschauung seiner selbst*]” (V: 42. 15-6).

This seems right, on Kant’s set-up. Any awareness of the activity (or causality) of reason is not going to be found in the world of experience.[[258]](#footnote-258) The activity of reason occurs in the noumenal realm, and to claim that we are conscious of this activity would involve some sort of access – “a special intuition of itself” (V: 42. 15-6) – to our noumenal self.

This is a crucial development in Kant’s thought. Here is Ameriks on this, one last time:

In saying Kant definitely moved away from an attempt at a deduction [in *Groundwork* III] I again mean not to commend his move but only to show how it was forced on him once he chose squarely to face both the full consequences of his theoretical philosophy with respect to the self and the implications of his deepest beliefs, his principles of practical philosophy. Although he could no longer believe he had a theoretical proof of our transcendental freedom close at hand, Kant still chose to reject compatibilism and to hold to the moral philosophy to which he had been committed longer and more intensely than his theoretical system. Without denying the temptation to wish Kant had not taken these moves, my aim here has been to show that we must recognize that in fact he did take them, and that some familiarity with the history of his views can help to make the fact less surprising than most believe. (Ameriks 2000: 219-20)

I agree with Ameriks that Kant makes this move, but I think that this ultimately counts against Kant’s treatment of freedom. I will shortly make this case, through considering the problem of recognition. I want to now set the scene for this by considering Jeanine Grenberg’s recent book on phenomenology in Kant’s moral philosophy.

### Grenberg and Moral Phenomenology in Kant

Grenberg’s recent work emphasises the importance of *first-personal moral phenomenology* in Kant, which she sees at the heart of his moral philosophy. The moral phenomenology in question is the fact of reason – our experience of an unconditional obligation. In what follows, I lay out her position, before suggesting that she, along with Kant, overlooks the importance of the third-person in moral philosophy.

Grenberg opens her book as follows:

The central claim of this book is that to engage in practical philosophy, Kantian style – indeed, in order to be *entitled* to the pursuit of cognitions beyond the limits of theoretical reason – we must set aside third-person, theoretical concerns and enter first into phenomenological reflection upon the common, first-personal experience of ourselves as agents. (Grenberg 2013: 15)

Grenberg complains about the “disfigurement of practical philosophy” (2013: 2), which becomes (at best) “a theoretical, scientific, third-personal reflection upon practical experience” (2013: 3).[[259]](#footnote-259) She is not alone in this. It is one of Korsgaard’s central claims that practical philosophy is first-personal, and theoretical philosophy third-personal.[[260]](#footnote-260) I follow these thinkers in the thought that we should not look to reduce practical questions – about what ought to be done – to theoretical questions about what in fact is done. However, this does not mean that there is no important third-person work to be done in moral philosophy. Indeed, I will argue that Grenberg runs together third-personal, theoretical and scientific claims in a way that is unhelpful, and furthermore that certain theoretical and third-personal thoughts are crucial to the enterprise of moral philosophy.

I think Grenberg adopts this exclusively first-personal approach for the same reasons that Kant abandons the approach of *Groundwork* III. She is acutely aware that any awareness of our activity would conflict with the epistemic limits of transcendental idealism. She then carefully follows Kant in the second *Critique* to offer an account of moral phenomenology that does not violate these limits. Like Kant himself, Grenberg offers an ingenious solution to a hopeless set up.

We can see this by thinking about third-person issues in ethics. Martin Sticker has recently responded to Grenberg’s work, by pointing to important instances of second and third person moral phenomenology in Kant himself (Sticker 2014: §4.2.1).[[261]](#footnote-261) Sticker is right that Kant does talk this way, but I think transcendental idealism precludes any such talk. Grenberg also recognises that Kant does make third-personal appeals to experience in his practical philosophy, but wants to distance herself and Kant from this (2013: 128-9). As such, Grenberg sticks closer to what Kant’s system can allow for, but I do not think that this counts in her favour. Without the third person, Kant’s moral philosophy is hopeless, or so I shall now argue.[[262]](#footnote-262)

### The Problem of Recognition

Transcendental Idealism is an ingenious solution to a hopeless set up. If we think that nature is entirely determined by natural necessity, but want to preserve a substantial conception of freedom, we must locate this freedom somehow outside nature. We then run into the problem of how we can be aware of our freedom. For a while, Kant thought he could provide a simple answer to this puzzle. In *Groundwork* III, he argues that we are conscious of the causality of our reason, but as we have seen, he then came to realise that this conflicts with the epistemological limits of transcendental idealism. If we can have no intuitions of, or epistemic access to, things in themselves, then we can have no consciousness of the causality of our reason. In its place in the second *Critique*, Kant proposes that we are conscious of an unconditional obligation, which commands us to do something, and we thus infer that we can. Any more than this would violate the set up. Even from this brief sketch, we can see the ingenuity of Kant’s solution. In what follows, I will reveal the hopelessness of the set up, and claim that, ultimately, Kant’s solution fails too. I argue that transcendental idealism precludes Kant from accounting for the recognition of other rational agents.

Kant’s moral philosophy contains an asymmetry between the first and the third person.[[263]](#footnote-263) In the second *Critique*, I experience the moral law as binding and this is what reveals *my* freedom to *me*. Grenberg follows suit to insist that phenomenology in Kant’s account is exclusively first-personal. As we have seen, she contrasts this with the third-personal approach, where everything is determined by natural necessity. The basic thought is that we can accept this conception of nature, but preserve freedom, as that is something that is exclusively revealed through first-person phenomenology, something that third-person explanations could never undercut. Again, this is an ingenious solution to a hopeless set up. If we think of nature as entirely determined by natural necessity, then we will find no freedom in it, nor the natural sciences. This can lead to the further thought that neither theoretical reason, nor the third-person standpoint, will be able to find freedom.

One major problem with such an exclusively first-personal approach involves the recognition of other rational agents. I am conscious of the moral law, which tells me that I ought to act in a certain way, and this reveals that I can. Once more, this is how *my* freedom is revealed to *me*. The problem concerns how this applies to the freedom of others:[[264]](#footnote-264) How can I know which other parts of nature are rational agents? Kant views the world of sense as entirely determined by natural necessity. As such, I cannot even experience (nor intuit) my own agency, so, *a fortiori*, I cannot experience (or intuit) the agency of others. Otherwise expressed, if our agency is exclusively first-personal, then there is no room for any meaningful third-personal judgements about other agents.

### Empirical Character

We find one possible response to this problem in the first *Critique*, where Kant draws a distinction between the intelligible and empirical character of a free being. He writes that:

[…] one can consider the causality of this being in two aspects, as intelligible in its action as a thing in itself, and as sensible in the effects of that action as an appearance in the world of sense. (A 538/B 566)

In his (so-called) ‘Clarification of the cosmological idea of a freedom in combination with the universal natural necessity’, Kant attempts to clarify how this is possible.

He is clear that we cannot experience intelligible character immediately (A 540/B 569). He remarks that:

[…] the intelligible character, which is the transcendental cause of the former [empirical character], is passed over as entirely unknown, except insofar as it is indicated through the empirical character as only its sensible sign. (A 546/B 574)

He says something similar shortly thereafter:

We are not acquainted with the latter [intelligible character], but it is indicated through appearances, which really give only the mode of sense (the empirical character) for immediate cognition. (A 551/B 579)

Kant’s claim is that human beings (and presumably finite rational agents in general), exhibit an empirical character.[[265]](#footnote-265) That is, our behaviour “exhibits a rule, in accordance with which one could derive the rational grounds and the actions themselves according to their kind and degree, and estimate the subjective principles of his power of choice” (A 549/B 577). The thought seems to be that our behaviour exhibits patterns, which allows us to estimate what maxims are in play, and furthermore, offers some indication or sensible sign of intelligible character.[[266]](#footnote-266) And Kant appeals to something similar in the second *Critique,* where he talks of *“*traces” [*Spuren*] (V: 85. 21) that suggest that an action was done wholly from respect for duty(V: 85. 21-2).[[267]](#footnote-267)

In the first and second Critiques, Kant operates with an entirely mechanistic conception of nature, which might seem to preclude the possibility of an empirical character. However, I want to leave this objection aside, as in the third *Critique* at least, Kant expands his conception of nature to include organisms (which operate according to ends and thereby require a different type of explanation than the merely mechanistic – I will return to say more about this in §3.2). This entitles Kant to talk of empirical character, but does not help with the problem of recognition as it still leaves us unable to differentiate between organisms and rational organisms.

Empirical character buys Kant a way to differentiate between certain parts of nature, namely that some are determined by what we might call ‘inner forces’ – desires, beliefs and so on. However, he is still stuck with an empirical psychology that explains *all actions* as entirely determined by natural necessity. As such, the most that we could ever get from observation is the “freedom of a turnspit” (V: 97. 19) that Kant famously derides as a “wretched subterfuge” (V: 96. 1). The fact that certain creatures are determined by desires does not reveal any significant form of agency.

Consider for instance, the difference between a dog, a crude cyborg, and a person. Each of these creatures can exhibit rules of behaviour, but only one of them is a rational agent. If we are confined to reading their rules of behaviour off an empirical world *entirely determined by natural necessity*, we will not be able to differentiate which of these patterns of behaviour exhibit an intelligible character.[[268]](#footnote-268)

By accepting a conception of nature entirely determined by natural necessity, Kant has painted himself into this corner. Consider for example, the following passage from the second *Critique*:

One can therefore grant that if it were possible for us to have such deep insight into a human being's cast of mind, as shown by inner as well as outer actions, that we would know every incentive to action, even the smallest, as well as all the external occasions affecting them, we could calculate a human being's conduct for the future with as much certainty as a lunar or solar eclipse and could nevertheless maintain that the human's being's conduct is free. (V: 99. 12-9)

I think Kant is too cavalier about his account of freedom here. His thought seems to be that, everything in nature is entirely determined and (in principle) predictable, but freedom is nevertheless possible. I experience the moral law, and this reveals *my* freedom to *me*. Once more, I accept that this approach makes freedom possible, but object that it precludes any recognition of other rational agents. I want to attempt to bring this out with the use of another example.

Imagine that you are a super-scientist that can predict the behaviour of everything under the sun. Given Kant’s setup, how could you possibly determine which things were rational agents or not? Everything in nature is determined and predictable, which precludes our discovering which parts of nature are free in the noumenal realm.

As a super-scientist, you have a complete understanding of how nature works, you understand everything that has happened and can predict everything that will happen. But still, you need to eat. So, 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, live animals and humans. How can you differentiate between these different types of things? They are all entirely predictable in accordance with natural laws. Some of these things might be more internally determined, but that is nothing more than the freedom of the turnspit. So, there you are looking at various different entirely predictable parts of nature. 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, 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 and predictable?

One possible response is that I am adopting the wrong perspective here. The super-scientist for instance, when they see the predictability of nature, is operating from the theoretical standpoint. The response is that, when the super-scientist comes to act, they adopt the practical standpoint, and from here, things are fine. However, even from this perspective, the super-scientist will have to be able to differentiate between the types of things that it is okay to eat, and the types of things that it is not okay to eat. And that involves a theoretical judgement about the nature of those things. We need to be able to judge which things are rational agents, and which are not. And when Kant claims that nature is entirely determined by natural necessity and predictable, he precludes this possibility.[[269]](#footnote-269)

Another possible response available to Kant involves an argument from analogy.[[270]](#footnote-270) The moral law reveals to me that I am a rational agent, but I also see my behaviour as exhibiting certain patterns in the world of experience. I then see other beings exhibit similar sorts of patterns, and can then infer that they are also rational agents. This seems like a promising avenue, but, once again, I think that Kant’s setup precludes it. The problem lies in the type of behaviour that I observe of myself in the world of experience – recall that on his set-up, I am unable to experience (or intuit) my own agency. When it comes to the world of experience, all of my behaviour is entirely determined by natural necessity. And in viewing the world of experience this way – both my behaviour and that of others – there is no way to tell which parts of it might have free causes in the noumenal.

One might suggest that, with human beings, we experience behaviour that requires explanations in terms of reasons, or agency. And that seems right. However, once again, Kant is not entitled to this, given his conception of experience. I am not being unfair on him here. This is his position. As we have just seen, he repeatedly insists that we could have no experience of freedom whatsoever. For instance, in the second *Critique*, he tells us that:[[271]](#footnote-271)

It is […] absolutely impossible to give anywhere in experience an example of it, since among the causes of things as appearances no determination of causality that would be absolutely unconditioned can be found (V: 48. 23-7)

This problem infects Kant’s whole theory of freedom. For instance, in the second subsection of *Groundwork* III, Kant famously argues that we must lend the idea of freedom to every rational being with a will, which some commentators take to be his crucial insight about freedom.[[272]](#footnote-272) But this is too quick, as once again, we still need to be able to determine which rational beings have wills, and which rational beings do not. Indeed, we need to be able to determine which bits of nature are rational beings at all.

Here, we touch upon the problem of other minds. Carol Van Kirk has argued that Kant does not face this problem, as there is no asymmetry between self-knowledge in the first personal and the third-personal case (1986: 41). However, she is talking exclusively about empirical character (1986: 42; 56). When it comes to our intelligible character, and especially our agency, as we have seen, there is an important asymmetry between the first and third person, and it is here where Kant’s account is problematic. Morality requires us to be able to recognise other rational agents. On this note, I want to return to Grenberg, and her emphasis on the first-person.

### Grenberg and the Third Person

In the introduction to her book, Grenberg contends that:

When we set aside theoretical modes of pursuing knowledge and turn instead to our practical experiences, we are no longer simply interested in knowing something; we are, more centrally, interested in that knowledge which will secure our status and efficacy as agents. Indeed, it is only by reflecting upon ourselves as agents that we find the very possibility of expanding our knowledge beyond the limits asserted in theoretical philosophy. (Grenberg 2013: 6)

The strict division between practical and theoretical philosophy is unhelpful. Indeed, it fails to secure precisely what Grenberg wants, namely, our status (and efficacy) as agents.

I have argued that Kant’s approach in the *second* *Critique* precludes recognition of other rational agents. I think that *Groundwork* III offers some hope on this score, in that we are at least conscious of the causality of reason in our own case. Grenberg recognises that in *Groundwork* III, Kant also seems to allow some third person talk of freedom. But she thinks he is mistaken to do so:[[273]](#footnote-273)

One could, perhaps, use the language of “experience” to refer to this rational reflection upon one’s actions. But this is an inference a third-personal spectator upon my action could make as easily as I could. Kant is thus not, in this passage, entering fully into the felt, first-personal experience of freedom; rather, there remains this remnant of a third-personal, even nonmoral or theoretical, starting point for his argument. (Grenberg 2013: 110)

As we have seen, Grenberg adopts Kant’s approach in the second *Critique*, where it is the experience of the moral law which reveals *my* freedom to *me*. Grenberg wants to stick to this exclusively first-personal approach.

She also resists Kant’s suggestions in the *Groundwork* that we might be conscious of the causality of our own reason, as this would conflict with the epistemic limits of transcendental idealism. She worries that:

[…] were Kant to be arguing here in *Groundwork* III that the felt experience of activity is the felt experience of positive freedom, he would be arguing to just this illicit noumenal content of felt phenomenological experience: he would be saying that we have a felt experience of being a noumenally rational causal force. (Grenberg 2013: 112)

Grenberg thinks that *Groundwork* III fails. However, unlike Ameriks, she does not think that it is just a re-hashing of Kant’s pre-critical account of the consciousness of our activity (2013: 119). She thinks that “Kant is at least *trying* to do something different” (Grenberg 2013: 119).

Here, Grenberg pulls apart Kant’s claim that we are conscious of our agency. She offers two options (Grenberg 2013: 114): We are conscious of either *negative* freedom (freedom from constraint) or *positive* freedom (which she equates with the freedom involved in reason). As we have just seen, she rules out the latter, as it would conflict with the epistemic limits of transcendental idealism. In its place, she suggests reading Kant’s talk of the phenomenology of our agency in *Groundwork* III as concerning *negative* freedom (Grenberg 2013: 112).

Here, Grenberg allows something close to consciousness of agency, but still insists that it is exclusively first-personal. I want to continue to push her on this. She offers a beautiful example of this feeling of activity:

Think, for example, of a child chasing seagulls on a beach: “My mother isn’t doing this; nor is she preventing me from doing this. I am causing those birds to fly away!” Even a child can distinguish the phenomenological nature of this experience from one in which her mother makes her brush her teeth, or in which her inability to walk prevents her from running. She is experiencing herself as being active in relation to her mental representations of the birds. (Grenberg 2013: 113)

This is a great example, but I am not a child, and (sadly) nowhere near a beach. What I can do though is picture this scenario: a child racing through the waves in delight as seagulls scatter at their every turn. I do not need to be that child, and I also do not need to imagine myself as that child.[[274]](#footnote-274) Instead, I can imagine watching that child experience its first glimpses of agency.

Grenberg offers another revealing example:

Psychologists also describe even very young children as having an experience of being free from constraints and of having an effect on the world when they smush their fingers happily in mud. The joy in such experiences comes from the child’s sense of being free from both external and internal constraints, and thus of being active in relation to her world. This does seem, then, a plausible common experience: sometimes I feel like I’m moved by things, but other times I feel like I’m moving or causing other things. (Grenberg 2013: 124-5)

This is another beautiful example. And I think it is a common experience. I have not smushed any mud in a while, but I do enjoy a sense of freedom when I cycle for instance, weaving left and right between the dotted parking lines. But what is revealing about Grenberg’s discussion is that psychologists pick up on this behaviour in children. Of course, this could be understood in an exclusively first-person sense: the children experience the activity of mud-smushing and then report this to the psychologists. But it seems more than this, it seems like the psychologists can observe this type of activity.[[275]](#footnote-275) Anecdotally, I do actually remember seeing a child discover part of their agency. They were going through a phase where all they did was lift up things and drop them, which brought them a lot of joy. This was something observable, and it was fairly clear what was going on. It seems to be a robust experience that psychologists can (and do) observe.

In general, we can observe agency. Of course, the knowledge that we gain on the basis of this is fallible. Perhaps we can never be completely *certain* whether something is an agent (or some act an act of agency) or not, but we can be more or less sure about various cases, with a good claim to knowledge.

This is not the only problem that transcendental idealism causes for freedom. Two other famous issues concern interaction and the (supposed) timelessness of reason. I will leave aside the timelessness of reason here,[[276]](#footnote-276) but want to say something about the problem of interaction, as I think it relates to the problem of recognition. After doing so, I turn to offer my own account of how freedom is possible

### The Problem of Interaction

Locating freedom outside of nature makes interaction between freedom and nature problematic. Consider this coffee I am now drinking. I drink it and my thoughts speed up. I wake up, start to tap my leg quickly, and begin to think a little faster. After a while, the coffee will wear off and I will feel tired. My thoughts will slow down considerably, and I will probably go home for a nap. Here we have interaction between the world, my thoughts and my actions. I want to get some work done this afternoon, so decide to grab a coffee. I drink it, and it helps. Transcendental idealism struggles to make sense of this series of interactions. This is a trivial example though. There are more serious ones. And here, transcendental idealism is not only unhelpful, it can also be offensive.

In the first *Critique* Kant tells us that:

[…] one may take a voluntary action, e.g. a malicious lie, through which a person has brought about a certain confusion in society; and one may first investigate its moving causes, through which it arose, judging on that basis how the lie and its consequences could be imputed to the person. With this first intent one goes into the sources of the person's empirical character, seeking them in a bad upbringing, bad company, and also finding them in the wickedness of a natural temper insensitive to shame, partly in carelessness and thoughtlessness; in so doing one does not leave out of account the occasioning causes. In all this one proceeds as with any investigation in the series of determining causes for a given natural effect. Now even if one believes the action to be determined by these causes, one nonetheless blames the agent, and not on account of his unhappy natural temper, not on account of the circumstances influencing him, not even on account of the life he has led previously; for one presupposes that it can be entirely set aside how that life was constituted, and that the series of conditions that transpired might not have been, but rather that this deed could be regarded as entirely unconditioned in regard to the previous state, as though with that act the agent had started' a series of consequences entirely from himself. This blame is grounded on the law of reason, which regards reason as a cause that, regardless of all the empirical conditions just named, could have and ought to have determined the conduct of the person to be other than it is. 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: now, in the moment when he lies, it is entirely his fault; hence reason, regardless of all empirical conditions of the deed, is fully free, and this deed is to be attributed entirely to its failure to act. (A 554-5/B 582-3)

In the second *Critique*, he repeats this claim:

[…] every action […] is to be regarded in the consciousness of his intelligible existence as nothing but the consequence and never as the determining ground of his causality as a noumenon. So considered, a rational being can now rightly say of ever unlawful action he performed that he could have omitted it (V: 97. 37 – 98. 8)

This is Kant’s dualism at its worst. He views the natural world as entirely mechanistically determined, and the noumenal world as entirely free, and is unable to say anything about their interaction. Kant claims that:

The judicial sentences of that wonderful capacity in us which we call conscience are in perfect agreement with this. (V: 98. 13-4)

It is more useful to think of actual judicial sentences. Judges navigate a difficult terrain all the time. People are accused of crimes and judges (and juries) determine whether they are guilty or not, but they also account for mitigating circumstances. There are various things that can impair our judgements (drugs, for instance), and various circumstances that excuse our behaviour (to a lesser or greater degree).[[277]](#footnote-277) By separating the noumenal world and the phenomenal word in the way he does, Kant cannot accommodate this within his system.[[278]](#footnote-278) This comes through in his treatment of moral motives.

In a footnote in the first *Critique*, after noting that we only have immediate cognition of the world of appearances, Kant claims that:

The real morality of actions (their merit and guilt), even that of our own conduct, therefore remains entirely hidden from us. Our imputations can be referred only to the empirical character. How much of it is to be ascribed to mere nature and innocent defects of temperament or to its happy constitution (*merita fortunae* [merit of fortune]) this no one can discover, and hence no one can judge it with complete justice. (A 551/B 579n)

We see an echo of this in the start of *Groundwork* II, where Kant famously claims that:

In fact, it is absolutely impossible by means of experience to make out with complete certainty a single case in which the maxim of an action that otherwise conforms with duty did rest solely on moral grounds and on the representation of one’s duty (IV: 407. 1-4)

It is important to differentiate between two claims here. The first is the plausible thought that we can never know with “complete certainty” or judge with “complete justice” what someone’s maxim is. And that might be right. But Kant is saddled with a bigger problem than this. On his set-up, it is not just that we cannot know with complete certainly what someone’s maxim is, but rather that we cannot know, with any degree of confidence, what someone’s maxim is.[[279]](#footnote-279)

In the *Metaphysics of Morals*, Kant appears to change his tune:

*Subjectively,* the degree to which an action *can be imputed (imputabilitas)* has to be assessed by the magnitude of the obstacles that had to be overcome. – The greater the natural obstacles (of sensibility) and the less the moral obstacle (of duty), so much the more merit is to be accounted for a *good deed,* as when, for example, at considerable self-sacrifice I rescue a complete stranger from great distress.

On the other hand, the less the natural obstacles and the greater the obstacle from grounds of duty, so much the more is a transgression to be imputed (as culpable). – Hence the state of mind of the subject, whether he committed the deed in a state of agitation or with cool deliberation, makes a difference in imputation, which has results. (VI: 228. 11-22)

This seems like a much better position. We no longer blame everyone, in every circumstance, equally for their actions. However, it is unclear how Kant can vindicate these judgements given his conception of experience.[[280]](#footnote-280) The passage from the *Metaphysics of Morals* makes it sound like we can account for interaction between the obstacles of sensibility and the motive of duty. But it is hard to see how this is possible on Kant’s set-up, where, as he outlined in the first (and second) *Critique*, *every* action is to be regarded in either two ways: 1) as an action in the world of experience, where it is entirely determined by natural necessity; or 2) as an action in the noumenal world, where it is entirely free. This setup leaves little room for any meaningful account of interaction.

Kant concludes his discussion of freedom in the Transcendental Dialectic as follows:

[To show] that this antinomy rests on a mere illusion, and that nature at least **does not conflict with** causality through freedom – that was the one single thing we could accomplish, and it alone was our sole concern. (A 558/B 586)

I think this is telling. Kant sets about showing how freedom is possible, in the face of the threat of a world entirely determined by natural necessity. And he accomplishes that. However, he does so at the price of precluding the possibility of recognising other rational agents, or a workable account of interaction. Of course, he could just insist that these are not his concerns. He might try to relegate these problems to the issue of explaining how freedom is possible; at the end of *Groundwork* III, Kant famously declares that we cannot explain how freedom is possible (IV: 459. 4-9).[[281]](#footnote-281) However, it is one thing to have a bit of metaphysical mystery in one’s approach, and quite another to offer a set-up that precludes the possibility of recognising other rational agents or a meaningful account of interaction. And Kant cannot afford to dismiss these as not his concerns – as I hope to have shown, recognition and interaction are crucial for our moral practices.

## 2.3 Nature and Freedom

I have argued that transcendental idealism does not provide an adequate account of freedom. In this final section, I want to sketch an alternative account of the possibility of freedom, one that locates freedom within, rather than outside of, nature.

In locating freedom outside of nature, Kant ruled out any experience of freedom, and thus was saddled with the problem of recognition. As we have seen, Kant did move beyond an entirely mechanistic conception of nature in the third *Critique*. In the Critique of Teleological Judgement, he offers a teleological treatment of organisms, detailing the ways in which organisms require explanations in terms of ends or purposes, rather than merely mechanical explanations.[[282]](#footnote-282) I have argued that, by itself, this does not allow Kant to overcome the problem of recognition as he still cannot differentiate between an organism and a rational organism from the third-person perspective.

Following Kant’s lead however, the German Idealists (in particular Schelling and Hegel) made progress on this front. As Kant himself realised, organisms require explanation in terms of ends, or purposes – final, rather than efficient, causes.[[283]](#footnote-283) For Kant though, all of this remained merely regulative;[[284]](#footnote-284) ultimately, we view nature this way, but this says nothing about the determination of the objects themselves.[[285]](#footnote-285) The German Idealists found this unsatisfying. Schelling, for instance, asks why certain parts of nature call out to be explained in this way.[[286]](#footnote-286) The simple answer is that they are different sorts of things. Organisms are different from non-organic parts of nature, and accordingly they require different types of explanation. Here is a simple example: It is a hot day, and a dog is drinking water. Why? Well, it is thirsty. It is hot, the dog has been running around, and now it needs to drink to sustain itself. Here we have moved away from mechanism to a conception of nature that involves beings that operate according to ends (and perhaps even representations).

This does not yet get us to us, and our freedom. But what it does is help disarm the thought that nature is entirely alien to freedom. At least before the third *Critique*, Kant sets up nature as entirely mechanistic, totally determined (and in principle predictable) in terms of efficient physical causes. Thinking about organisms enriches our conception of nature, and moves us away from this.

Moving on from this, with the advent of society, culture and language, some organisms become rational. And not only do rational beings act according to ends, but we also have the capacity to reflect upon these ends. Much like the introduction of organic matter, this requires different types of explanations – social, rational – because there are now different things. McDowell is spot on when he insists upon two points here: 1) this can all be thought of as natural; 2) we should not shy away from thinking that the introduction of rational beings makes a metaphysical difference.

These two points might seem in tension, but I do not think they are. McDowell famously invokes second nature, as a type of social upbringing [*Bildung*] that allows us to understand how we came to occupy the space of reasons without falling into rampant Platonism.[[287]](#footnote-287)

His diagnosis of Kant is spot on here. He writes that:

Since he does not contemplate naturalism of a second nature, and since bald naturalism has no appeal for him, he cannot find a place in nature for this required real connection between conceptions and intuitions [namely, spontaneity]. And in this predicament, he can find no option but to place the connection outside nature in the transcendental framework.

Kant is peculiarly brilliant here. Even though he has no intelligible way to deal with it (McDowell 1994: 98)

Second nature provides a way of accommodating our capacity to reason within nature. Without this:

We fall into rampant Platonism if we take it that the structure of the space of reasons is *sui generis*, but leave in place the equation of nature with the realm of law [or natural necessity]. That makes our capacity to respond to reasons look like an occult power, something extra to our being the kind of animals we are, which is our situation in nature. (McDowell 1994: 83)

McDowell also notes that we should not shy away from thinking that this makes a metaphysical difference. Once we have beings that are brought up in such a way that they can enter into a space of reasons, we are dealing with a different type of being.[[288]](#footnote-288)

Of course, this is all rather suggestive.[[289]](#footnote-289) But I think it is adequate for the task at hand. The task is to show that freedom is possible, and that we are entitled to think of ourselves this way. We have a lot of evidence for our freedom – the consciousness of our activity (and the experience of an unconditional obligation) – but something puts this into doubt. The grounds for doubt concern an entirely mechanistic conception of nature, a world that operates solely according to mechanistic laws. This threat can be dealt with, as there is more to science than (Newtonian) physics. There is also biology for instance, which studies things that operate in ways other than mere mechanism. They operate according to ends, and accordingly require explanations in terms of final, rather than merely efficient causes. This enriches our conception of nature in a way that makes it not entirely alien to freedom. The notion of second-nature does this even further, in that it allows us to account for our social practices and the space of reasons within nature.

So, pace Kant, transcendental idealism is not necessary for freedom. And moreover, it does not help. A major problem, as we have seen, involves the strict epistemological limits it imposes, which precludes the possibility of recognising other rational agents. This is especially unsatisfying given that securing freedom is supposed to be one of the main benefits of transcendental idealism. But what transcendental idealism gives with one hand, it takes away with the other.

# Conclusion

Once more, I have begun with an important insight that I find in *Groundwork* III. Kant recognises that we need to dispel the threat of determinism. Something has to make freedom possible, and we also need to show that it is actual (or necessary). Unfortunately, Kant signs up for a conception of nature that is entirely determined by natural necessity, and thus has to locate freedom outside of it. This precludes him from providing a workable account of recognition or interaction. For this, we need a richer conception of nature. However, we can still retain an important part of Kant’s approach in *Groundwork* III, namely that, we are conscious of the causality of our agency. In moving away from Kant’s system though, we can also vindicate our (third-person) awareness of others’ agency.

4 – Rational Agency and the Moral Law

In this final chapter, I turn to the second major objection to Kant’s argument for freedom in *Groundwork* III. So far, we have seen that reason involves spontaneity (ch.1), and that we are entitled to think of our reason as practical (ch.3). The objection at hand grants all of this, but insists that the spontaneity involved in rational agency is not the same as the freedom required for the moral law.

Kant is thus thought to equivocate in his use of the term ‘free’ between the two premises.

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

The worry is that we could be spontaneous rational agents, but unable to act for the sake of the moral law, which would render the above argument invalid.

This objection is perhaps most prominently pushed by Allison, who accepts (versions of[[290]](#footnote-290)) both of our premises, but rejects the argument.[[291]](#footnote-291) Allison’s critique has been influential, and this issue is often taken to count decisively against the approach of *Groundwork* III.[[292]](#footnote-292) I think I can overcome this objection. Indeed, I contend that with the ground we have laid in the previous chapters, the objection drops away. For this, I need to lay out the objection in a little more detail.

# A Gap

In *Kant’s Theory of Freedom*, Allison makes out the problem in terms of a distinction between practical and transcendental freedom (1990: 227-8). He thinks of practical freedom as independence from any particular influence of sensibility, and transcendental freedom as independence from the influence of sensibility as such (Allison 1990: 65); another way of putting this distinction is to view practical freedom as the ability to back up from *any* particular desire, and transcendental freedom as the ability to back up from *all* desires.

The objection then runs as follows. Kant’s claim that reason involves spontaneity only seems to entitle us to practical freedom. Allison notes that this involves:

[…] a genuine, albeit limited, spontaneity and, therefore, a capacity to act on the basis of imperatives, although the incentives for obeying these imperatives would ultimately be traceable to our sensuous nature. (Allison 1990: 65)

This conception of freedom seems sufficient for rational agency. We can thus accept Kant’s claim that reason involves spontaneity, but resist the thought that this establishes the specifically *transcendental* freedom that the moral law requires; as we saw in chapter 2 (§1.3), Allison insists that the moral law requires transcendental freedom.

Putting all of this together, Kant’s argument in *Groundwork* III looks something like the following:

1.\* A rational will is a *practically* free will

2.\* A *transcendentally* free will stands under the moral law

3. Therefore, a rational will is a will under moral law.

This argument contains an equivocation between two different conceptions of freedom, and accordingly falls short.

I think we can overcome this supposed gap. As rational agents, we are spontaneous and possess the capacity to act on the basis of imperatives. Allison’s worry is that this is not enough, as it is still possible that “the incentives for obeying these imperatives would ultimately be traceable to our sensuous nature” (1990: 65). This needs to be unpacked.

Crucially, how are we to understand the claim that the incentives for our acting on the basis of imperatives are “ultimately traceable to our sensuous nature”? As rational agents, we must take our obeying imperatives as up to us. This is one of Allison’s central claims – the Incorporation Thesis.[[293]](#footnote-293) As we saw in chapter 1 (§3.2), he characterises this feature of rational agency as follows:

[…] justification goes all the way down […] one cannot appeal to a desire, no matter how strong, in justification of a course of action, […] to do so is to remove oneself from the space of reasons; it is to takes one’s behaviour as caused (and therefore as perhaps excused), but not as rationally justified. In that case, one is no longer acting under the idea of freedom, because one is no longer *acting*. (Allison 2012, 93)

It is unclear how rational agency, so conceived, could bottom out in sensuous nature. Engstrom (1993) offers a similar objection to Allison’s (1990) account of the (supposed) gap between the freedom involved in rational agency and the freedom involved in the moral law.

In response, Allison has changed his terms.[[294]](#footnote-294) He still insists that there is a gap in *Groundwork* III, but has clarified the exact nature of the problem. The issue is motivational, and Allison now expresses it through the simple claim that: *spontaneity does not entail autonomy*. He glosses spontaneity as causal independence from our desires, and autonomy as motivational independence from our desires. So conceived, to be spontaneous is to be not causally determined by one’s desires, whereas to be autonomous is to be able to determine oneself independently of one’s desires. Allison’s claim is then that:[[295]](#footnote-295)

Given this distinction, it follows that we must allow at least conceptual space for the notion of an agent that possesses genuine spontaneity but not autonomy, that is, one that is both free in an indeterminist sense and heteronomous. (Allison 1996: 111)

[…] it seems perfectly possible that a will might be free in the contra-causal sense of not being causally necessitated by antecedent conditions and yet ineluctably heteronomous in the sense that its menu of incentives (or motives) all stem from its sensuous nature. (Allison 2012: 116)

Hill offers a similar objection to *Groundwork* III:[[296]](#footnote-296)

[…] though acting for reasons is not to be understood as being causally determined by one’s given desires, nevertheless one’s capacity to act for reasons might be limited, perhaps even by the concept of reasons, to policies aimed at the satisfaction of some of the desires and inclinations one happens to have. We might speak of such a will as incapable of motivation by anything but inclination and desire, where “motivation” refers not to what causes the willing but rather to the range of things the will can count as reasons or rational objectives of its policies. (Hill 1998: 257)

The problem is that we could be spontaneous rational agents, yet incapable of acting for the sake of the moral law; otherwise expressed, it is possible that we are heteronomous rational agents, rather than the autonomous beings that the moral law requires.[[297]](#footnote-297) This is the crux of the objection of this chapter. I want to now defuse this threat to our argument.

# Reason and Motivation

In chapter 2 (§1.5), we considered a motivational objection to moral realism. The objection ran that, even if there were ends of unconditional value, we would be unable to be motivated by them. I (briefly) argued that we should move beyond the idea of needing pre-existing desires to motivate us.[[298]](#footnote-298) I claimed that this strong internalism was implausible, and I think this holds for both moral and non-moral motivation.

Once more, it is crucial to pull apart the claim that we need a desire to motivate us to action. This can be understood in either a weak way (something like action requires an affective state) or a strong way (action requires an actual pre-existing desire). Understood in the weak way, the claim is compatible with the thought that reasons can motivate us, through allowing that recognising reasons can generate desires. And understood in the strong sense, the claim is no longer plausible.

Allow me to share two classical treatments of this point.

Nagel argues that:

[…] it may be admitted as trivial that, for example, considerations about my future welfare or about the interests of others cannot motivate me to act without a desire being present at the time of action. That I have the appropriate desire simply follows from the fact that these considerations motivate me; if the likelihood that an act will promote my future happiness motivates me to perform it now, then it is appropriate to ascribe to me a desire for my own future happiness. But nothing follows about the role of the desire as a condition contributing to the motivational efficacy of those considerations. It is a necessary condition of their efficacy to be sure, but only a logically necessary condition. It is not necessary either as a contributing influence, or as a causal condition. (Nagel 1970: 29-30)

Nagel thus disarms the weak claim that we need desires to motivate us. He also follows Aristotle to deny the strong claim, pointing out that “many desires, like many beliefs, are *arrived at* by decision and after deliberation” (1970: 29).[[299]](#footnote-299)

McDowell makes a similar point:

Why should it not be the case [...] that the agent's conception of the situation, properly understood, suffices to show us the favourable light in which his action appeared to him? If we credit him with a suitable desire, then, [...] that need be no more than a consequence of the fact that we take his conception of the circumstances to have been his reason for acting as he did; the desire need not function as an independent component in the explanation, needed in order to account for the capacity of the cited reason to influence the agent's will. (McDowell 1998: 79)

I agree with Nagel and McDowell. This is how we talk about our motivation, and I view Humean scepticism about practical reason as unmotivated.

We saw a similar move in chapter 3 (§1.2), in our discussion of doubt and moral motivation. In the second *Critique*, Kant claims that we experience an unconditional obligation, which tells us that we ought to act independently of our inclinations, thus revealing that we can. He calls this the fact of reason, and claims that it is “undeniable [*unleugbar*]” (V: 31. 24). As I noted previously though, this is not undeniable. A sceptic about pure practical reason would claim that the obligation is not really unconditional, but that we are led to think so due to whatever their favourite debunking story is.[[300]](#footnote-300) Kant however is uninterested in this type of scepticism. He takes our moral obligations at face-value, and attempts to provide a system that can vindicate them.

I want to do something similar for practical reason in general. Timmermann insists that Kant is best thought of as entirely Humean about non-moral motivation.[[301]](#footnote-301) But I want to resist this general scepticism about practical reason. In the second *Critique*, Kant resists it in the moral case, and I want to resist it in the case of action generally. We can be motivated by reasons in non-moral matters. But I will not belabour this point here. Instead, I want to show how this approach can overcome the objection at hand.

The objection in this chapter is that all human action might require pre-existing desires, which would mean that, even if we are spontaneous, we are still incapable of acting for the sake of the moral law. However, as McDowell and Nagel highlight, we have good reasons to reject that picture of motivation, even in the non-moral case.

Of course, it is *possible* that we are the type of rational agents who always need pre-existing desires to act, and can thus never act for the sake of the moral law. And it is this possibility that worries Allison:

[…] we must allow at least conceptual space for the notion of an agent that possesses genuine spontaneity but not autonomy, that is, one that is both free in an indeterminist sense and heteronomous. (Allison 1996: 111)

[…] it seems perfectly possible that a will might be free in the contra-causal sense of not being causally necessitated by antecedent conditions and yet ineluctably heteronomous in the sense that its menu of incentives (or motives) all stem from its sensuous nature. (Allison 2012: 116)

However, this mere possibility need not trouble us. We are rational agents, and can be motivated by reasons. That is a perfectly plausible description of our motivation, in both the moral and non-moral cases.

Allison seems excessively worried about a mere possibility. And this, once again, ties in to the important issue of how one conceives the project of *Groundwork* III. If one thinks that Kant is trying to move from the bare idea of reason to rational agency to the moral law, then any possible gaps are very problematic. But this is not the best way to understand *Groundwork* III. As I have argued, Kant is trying to vindicate our standing under the moral law through appealing to a non-moral conception of rational agency. And, in the case of non-moral action, I think we are perfectly entitled to think of ourselves as capable of being motivated by reasons. The threat of a gap between rational agency and the moral law then drops away. Indeed, the gap only opens up if one subscribes to a (strong) internalist theory of motivation. And, as I have argued, we do not have to do that.

I take this to be a virtue of my approach. If we read *Groundwork* III in the way I suggest, and reject the strong reading of the claim that we need desires to motivate us, then we can disarm the threat of a gap between rational agency and the moral law. We thereby defend Kant’s argument against this objection.

Conclusion

Here is Kant’s argument for freedom in *Groundwork* III, one last time:

1. A rational will is a free will
2. A free will stands under the moral law
3. Therefore, a rational will stands under the moral law

Stepping back from the specifics of Kant’s argument, the basic thought is that, in thinking about reason and agency, we can vindicate our conception of ourselves as free, and capable of being moral. In this thesis, I hope to have defended this thought.

I have argued that we are entitled to think of ourselves as rational agents, and thereby as capable of standing under the moral law.

In doing so, I have departed from Kant (and Kantians) at several points. Most significantly, I have argued against a non-metaphysical account of freedom (ch.1), a formalist or constructivist meta-ethics (ch. 2), and transcendental idealism itself (ch.3).

I have kept with the spirit of Kant’s philosophy, but often departed from the letter. In this, I find my project to be post-Kantian. And this came through in the thesis. I began with Kant, but often ended up alongside Kant’s successors (primarily Hegel) and their present-day followers (most notably McDowell and Stern).

I hope to contribute to this movement. Analytic Philosophy began as a reaction against Hegel and the British Idealists.[[302]](#footnote-302) It underwent a long Humean phase, before Kant came back into fashion. Primarily due to the influence of Rawls and his students, Kant’s views on meta-ethics and agency have recently received an excellent airing. I follow Kant (and Kantians) in wanting to vindicate our conception of morality as unconditionally normative, and of ourselves as free. However, as I hope to have shown, there are some serious limitations to Kant’s (and Kantians’) attempt(s) to do this. Kant hits upon the right basic insights, but his system cannot sustain them.

I think we need to move beyond transcendental idealism to embrace a more encompassing conception of nature. And for this, the post-Kantians (primarily Schelling and Hegel) are the prime candidates. Of course, this approach might have problems of its own, but that is another story.

Bibliography

I quote Kant according to the standard *Academy Edition*. Translations from the *Groundwork* are from Timmermann (2011). Other works by Kant are quoted from the *Cambridge Edition of the Works of Immanuel Kant* edited by Paul Guyer and Allen Wood.

Allison, H (1969) ‘Transcendental Idealism and Descriptive Metaphysics’ in *Kant-Studien*, 60.2, pp. 216-33.

– (1989) ‘Justification and Freedom in the Critique of Practical Reason’ in Förster (1989), pp. 114-30.

– (1990) *Kant’s Theory of Freedom* (Cambridge University Press).

– (1996) *Idealism and Freedom: Essays on Kant’s Theoretical and Practical Philosophy* (Cambridge University Press).

– (1998) ‘Morality and Freedom: Kant’s Reciprocity Thesis’ in Guyer (1998), pp.273-303.

– (2004) *Kant's Transcendental Idealism: An Interpretation and Defense* (Yale University Press) (2nd edition).

– (2011) *Kant’s Groundwork of the Metaphysics of Morals. A Commentary* (Oxford University Press).

– (2012) *Essays on Kant* (Oxford University Press).

Ameriks, K (2000) *Kant’s Theory of Mind* (Oxford University Press) (2nd edition).

– (2000b) *The Fate of Autonomy. Problems in the Appropriation of the Critical Philosophy* (Cambridge University Press).

– (2003) *Interpreting Kant’s Critiques* (Oxford University Press).

– (2006) *Kant and the Historical Turn. Philosophy as Critical Interpretation* (Oxford University Press).

– (2012) *Kant’s Elliptical Path* (Oxford University Press).

Aristotle, *The Nicomachean Ethics.*

Bagnoli, C. ‘Morality as Practical Knowledge’ in *Analytic Philosophy*, 53.1, pp. 61-70.

– (ed) (2013) *Constructivism in Ethics* (Cambridge University Press).

Beck, L. (1960) *A Commentary on Kant’s Critique of Practical Reason* (The University of Chicago Press).

– (1975) *The Actor and the Spectator* (Yale University Press).

Beiser, F. (2002) *German Idealism: The Struggle against Subjectivism*, *1781-1801* (Harvard University Press).

Bennett, J. (1984) ‘Kant’s Theory of Freedom’ in Wood (1984).

Blöser, C. (2015) ‘Degrees of Responsibility in Kant's Practical Philosophy’ in *Kantian Review*.

Bojanowski, J. (2012) ‘Is Kant a Moral Realist?’ in *Kantian Yearbook*, 4, pp. 1-22.

Boyle, M. (2011) ‘Making up your Mind and the Activity of Reason’ in *Philosophers’ Imprint*, 11.17.

Brandom, R. (1997) ‘Study Guide’ in Sellars (1997), pp. 119-81.

Brooks, T (ed). (2012) *Hegel’s Philosophy of Right: Essays on Ethics, Politics and Law* (Wiley-Blackwell).

Brown, J. (1998) *The Self* (McGraw-Hill).

Callanan, J. (2014) ‘The Role of the Holy Will’ in *Hegel Bulletin* 35, pp. 163-184.

Cohen, G. (1996) ‘Reason, Humanity, and the Moral Law’ in Korsgaard (1996b), pp. 167-88.

Darwall, S. (2006) *The Second-Person Standpoint. Morality, Respect, and Accountability* (Harvard University Press).

Deligiorgi, K. (2012) *The Scope of Autonomy: Kant and the Morality of Freedom* (Oxford University Press).

Enoch, D. (2006) ‘Agency, Shmagency: Why Normativity Won’t Come from What is Constitutive of Action’, in *Philosophical Review*, 115.2, pp. 169-98.

– (2011) ‘Shmagency Revisited’ in Brady, M (ed) (2011) *New Waves in Metaethics* (Palgrave Macmillan), pp. 208-33.

Engstrom, S. (1993) ‘Allison on Rational Agency’, in *Inquiry*, 36, pp. 405-18.

– (2009) *The Form of Practical Knowledge* (Harvard University Press).

– (2012a) ‘Summary of The Form of Practical Knowledge’ in *Analytic Philosophy*, 53.1, pp. 58-60.

– (2012b) ‘Bringing Practical Knowledge into View: Response to Bagnoli, Hill, and Reath’ in *Analytic Philosophy*, 53.1, pp. 89-97.

– (2013) ‘Constructivism and Practical Knowledge’, in Bagnoli (2013), pp. 133-52.

Formosa, P. (2011). ‘Is Kant a Moral Constructivist or a Moral Realist?’ in *European Journal of Philosophy*, 21.2, pp. 170-96.

Förster, E. (ed). *Kant’s Transcendental Deductions: The Three ‘Critiques’ and the ‘Opus Postumum’* (Stanford University Press).

Fichte, J. (1797) *Foundations of Natural Right*, translated by (2000) Baur, M. (Cambridge University Press).

Freyenhagen, F. (2008) ‘Reasoning Takes Time: On Allison and the Timelessness of the Intelligible Self’ in *Kantian Review*, 13.2, pp. 67-84.

Fugate, C. (2012) ‘On a Supposed Solution to the Reinhold/Sidgwick Problem in Kant’s Metaphysics of Morals’ in *European Journal of Philosophy*.

Geiger, I. (2010) ‘What it is the Use of the Universal Law Formula of the Categorical Imperative?’ in *British Journal for the History of Philosophy,* 18.2, pp. 271-95.

Grenberg, J (2009) ‘The Phenomenological Failure of *Groundwork* III’ in *Inquiry*, 52.4, pp. 335-56.

– (2013) *Kant's Defense of Common Moral Experience. A Phenomenological Account* (Cambridge University Press).

Guyer, P (ed). (1998) Kant’s Groundwork of the Metaphysics of Morals. *Critical Essays* (Rowman & Littlefield).

– (2005) *Kant’s System of Nature and Freedom* (Oxford University Press).

– (2007a) *Kant’s Groundwork for the Metaphysics of Morals: A Reader’s Guide* (Continuum)

– (2007b) ‘The Derivation of the Categorical Imperative: Kant’s Correction for a Fatal Flaw’ in *Harvard Review of Philosophy*, 10, pp. 64-80.

– (2009) ‘Problems with Freedom: Kant’s Argument in Groundwork III and its Subsequent Emendations’ in Timmermann (2009), pp. 176-202.

Guyer, P and Wood, A (1995) (eds). *The Cambridge Edition of the Works of Immanuel Kant*, (Cambridge University Press).

Hammer, E (ed). *German Idealism: Contemporary Perspectives* (Routledge)

Hegel, G. (1977) [PS] *Phenomenology of Spirit* (Oxford University Press), translated by Miller, A V.

– (1991) [EL] *The Encyclopedia Logic* (Hackett), translated by Geraets, W, Harris, H, and Suchting, W.

– (1991)[PR] *Philosophy of Right* (Cambridge University Press), translated by Nisbet, H B.

– (2001) [PH] *The Philosophy of History* (Batoche Books), translated by Sibree, J.

– (2006) [LHP] Lectures on the History of Philosophy (Oxford University Press), translated by Brown, R.

Henrich, D (1975) ‘The Deduction of the Moral Law: The Reasons for the Obscurity of the Final Section of Kant’s *Groundwork of the Metaphysics of Morals*’ in Guyer (1998), pp.303-341.

– (1989) ‘Kant’s notion of a Deduction and the Methodological Background to the First *Critique*’ in Förster (1989), pp.29-46.

– (1994) *The Unity of Reason: Essays on Kant’s Philosophy* (Harvard University Press).

Hill, T. (1992) *Dignity and Practical Reason* (Cornell University Press).

– (1998) ‘Kant’s Argument for the Rationality of Moral Conduct’, in Guyer (1998), pp. 249-72.

– (2013) ‘Kantian autonomy and contemporary ideas of autonomy’ in Sensen (2013a), pp. 15-31.

Hookway, C (1999) ‘Modest Transcendental Arguments and Sceptical Doubts: A Reply to Stroud’ in Stern (1999), pp. 173-87.

Horn, C and Schönecker, D (eds). (2006): *Groundwork for the Metaphysics of Morals* (de Gruyter).

Korsgaard, C. (1996a) *Creating the Kingdom of Ends* (Cambridge University Press).

– (1996b) *The Sources of Normativity* (Cambridge University Press).

– (2008) *The Constitution of Agency* (Oxford University Press).

– (2009) *Self-Constitution: Agency, Identity, and Integrity* (Oxford University Press).

Kosch, M. (2006) *Freedom and Reason in Kant, Schelling and Kierkegaard* (Oxford University Press).

Land, T. (2006) ‘Kant’s Spontaneity Thesis’ in *Philosophical Topics*, 34, pp. 189-220.

Langton, R. (2007) ‘Objective and Unconditioned Value’, in *Philosophical Review*, 116.2, pp. 157-85.

Laurence, S and Margolis. (2013) ‘In Defence of Nativism’ in *Philosophical Studies*, 165.2, pp. 693-718.

Lenman, J and Shemmer Y. (2012) ‘Introduction’ in Lenman, J and Shemmer Y (eds). (2012) *Constructivism in Practical Philosophy* (Oxford University Press), pp. 1-17.

Lowe, E. (2008) *Personal Agency: The Metaphysics of Mind and Action* (Oxford University Press).

Mackie, J. (1977) *Inventing Right and Wrong* (Viking Press).

Maimon, S. (1790) *Essay on Transcendental Philosophy*, translated by (2010) Midgley, N, Somers-Hall, H, Welchman, A, and Reglitz, M. *(Continuum).*

McCarthy, M. (1982) ‘Kant’s Rejection of the Argument of *Groundwork* III’ in *Kant-Studien,* 73.2, pp. 169-190.

*–* (1985) ‘The Objection of Circularity in *Groundwork* III’ in *Kant-Studien*, 76.1, pp. 28-42.

McDowell, J. (1994) *Mind and World* (Harvard University Press).

– (1998) *Mind, Value, and Reality* (Harvard University Press).

– (2000) ‘Responses’ in Willaschek (ed) (2000), pp. 91-114.

– (2006) ‘Response to Graham Macdonald’ in Macdonald C, and Macdonald G (eds). (2006) *McDowell and his Critics* (Blackwell), pp. 235-39.

– (2009) *Having the World in View* (Harvard University Press).

Mieth, C and Rosenthal, J. (2006) ‘Freedom Must be Presupposed as a Property of the Will of all Rational Beings’ in Horn and Schönecker (2006), pp. 247-84.

Moore, A. (2012) *The Evolution of Modern Metaphysics* (Cambridge University Press).

Nagel, T. (1970) *The Possibility of Altruism* (Princeton University Press).

– (1996) ‘Universality and the Reflective Self’ in Korsgaard (1996b).

Nelkin, D. (2000) ‘Two Standpoint and the Belief in Freedom’ in *The Journal of Philosophy*, 97, pp. 564-576.

O’Neill, O. (1989) *Constructions of Reason: Exploration of Kant's Practical Philosophy*(Cambridge University Press).

Paton, H. (1947) *The Categorical Imperative: A study in Kant’s Moral Philosophy* (Hutchinson University Library).

– (1948) *The Moral Law* (Hutchinson University Library).

Patten, A. (1999) *Hegel’s Idea of Freedom* (Oxford University Press).

Peirce, C. (1868) ‘Some Consequences of Four Incapacities’ in *The Essential Peirce, Volume1* (Indiana University Press), pp. 28-55.

(1877) ‘The Fixation of Belief’ in *The Essential Peirce, Volume1* (Indiana University Press), pp. 109-123.

– (1902) ‘On Science and Natural Classes’ in *The Essential Peirce, Volume1I* (Indiana University Press), pp. 115-132.

Pippin, R. (1987) ‘Kant on the Spontaneity of Mind’ in *Canadian Journal of Philosophy*, 17.2, pp. 449-475

– (2009) ‘Natural and Normative’ in *Daedalus*, 138, pp. 35-43.

Quarfood, M. (2006) ‘The Circle and the Two Standpoints’ in Horn and Schönecker (2006), pp. 285-301.

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

Reath, A. (1993) ‘Intelligible Character and the Reciprocity Thesis’ in *Inquiry* 36, pp. 419-30.

– (2006) *Agency and Autonomy in Kant’s Moral Theory. Selected Essays* (Oxford University Press).

– (2013) ‘Kant’s conception of autonomy of the will’ in Sensen (2013a), pp. 32-53.

Redding, P (2007) *Analytic Philosophy and the Return of Hegelian Thought* (Cambridge University Press).

Reinhold, K (1792) Briefe über die Kantische Philosophie (Zweyter Band).

Saunders, J. (2014) ‘Kant, Rational Psychology and Practical Reason’ in *Kant Yearbook*, 6,pp. 115-131.

– (2016) ‘The Value of Practical Knowledge’ in *Hegel Bulletin*.

Saunders, J and Williams, N. (draft) ‘Practical Reasons to Believe: Kant and James on Religion’

Schelling, F. (1803) *Ideas for a Philosophy of Nature*, translated by (1988) Harris, E and Heath P. (Cambridge University Press).

Schönecker, D (1999) *Kant: Grundlegung III. Die Deduktion des kategorischen Imperativs* (Karl Alber – Verlag).

– (2006) ‘How is a Categorical Imperative Possible?’ in Horn, Schönecker (2006), pp. 301-23.

– (2012) ‘Review of Allison, H Kant’s Groundwork for the Metaphysics of Morals: A Commentary’ in *Notre Dame Philosophical Reviews*.

– (2013) ‘A free will and a will under moral laws are the same’ in Sensen (2013a) pp. 225-45.

Schopenhauer, A. (1818) *The World as Will and Representation*, translated by (2014) Norman, J and Welchman, A. (Cambridge University Press).

Searle, J (1984): *Minds, Brains and Science* (Harvard University Press).

Sellars, W. [1956] (1997): *Empiricism and the Philosophy of Mind* (Harvard University Press).

– (1971): ‘...this I or He or It (The thing) which thinks...’ in *Proceedings and Addresses of the American Philosophical Association*, 44, pp. 5-31.

Sensen, Oliver (2009) ‘Kant’s Conception of Inner Value’ in *European Journal of Philosophy*, 19.2, pp. 262-80.

– 2009b ‘Dignity and the formula of humanity’, in Timmermann (2009), pp.102-19.

– (2013a) (ed) *Kant on Moral Autonomy* (Cambridge University Press).

– (2013b) ‘The moral importance of autonomy’ in Sensen (2013a).

Shell, Susan (2009) *Kant and the Limits of Autonomy*, Harvard University Press: Cambridge, Mass.

Shemmer, Y (2007) ‘Desires as Reasons’ in *Philosophy and Phenomenological Research*, 75. 2, pp. 326-48.

Sidgwick, H. (1874) *The Method of Ethics*.

Smith, J and Sullivan, P (eds). (2011) Transcendental Philosophy and Naturalism (Oxford University Press).

Stern, R. (1990) *Kant, Hegel and the Structure of the Object* (Routledge).

– (1999) (ed) *Transcendental Arguments: Problems and Prospects* (Oxford University Press).

– (2000) *Transcendental Arguments and Scepticism* (Oxford University Press).

– (2007) ‘Freedom, Self-Legislation and Morality in Kant and Hegel: Constructivist vs Realist Accounts’, in Hammer, E (2007), pp. 245-66.

– (2009) *Hegelian Metaphysics* (Cambridge University Press)

– (2010) ‘Moral Scepticism and Agency: Kant and Korsgaard’, in *Ratio*, 23, pp. 451-74.

– (2011) **'The Value of Humanity: Reflections on Korsgaard's Transcendental Argument'**  
in Smith, J and Sullivan, P. (2011), pp. 74-95.

– (2012a) ‘On Hegel’s Critique of Kant’s Ethics: Beyond the “Empty Formalism” Objection’, in Brooks (2012), pp. 73-99.

– (2012b) *Understanding Moral Obligation* (Cambridge University Press).

– (2013a) *Hegel’s Phenomenology of Spirit* (2nd ed) (Routledge).

– (2013b) ‘Moral Skepticism, Constructivism, and the Value of Humanity’, in Bagnoli (2013), pp. 22-40.

– (forthcoming) 'Round Kant or through him? On James' Arguments for Freedom, and their relationship to Kant's'.

Strawson, G. (2010): *Freedom and Belief*, Oxford.

Sticker, M. (2014). *Common Human Reason in Kant: A Study in Kant’s Moral Psychology and Philosophical Method* (PhD Thesis).

– (2014b) ‘How can Common Human Reason confirm the Correctness of the Deduction in Groundwork III?’ in *Hegel Bulletin*, 35.2, pp. 228-51.

Sticker, M and Van Ackern. (2014) ‘Kant and Moral Demandingness’, in *Ethical Theory and Moral Practice.*

Street, S. (2010) ‘What is Constructivism in Ethics and Metaethics?’ in *Philosophy Compass*, 5, pp. 363-84.

Sussman, D (2008) ‘From Deduction to Deed: Kant's Grounding of the Moral Law’ in *Kantian Review*, 13.1, pp. 52-81.

Taylor, C. (1975) *Hegel* (Cambridge University Press).

Tenenbaum, S. (2012) ‘The Idea of Freedom and Moral Cognition in *Groundwork* III’ in *Philosophy and Phenomenological Research,* 84.3, pp. 555-89.

Timmermann, J. (2006) ‘Value without Regress: Kant’s ‘Formula of Humanity’ Revisited’ in *European Journal of Philosophy*, 14.1, pp. 69-93.

– (2007) *Kant's Groundwork of the Metaphysics of Morals. A Commentary* (Cambridge University Press).

– (2009) ‘Acting from duty: inclination, reason and moral worth’ in Timmermann (2009b), pp.45-63.

– (2009b) (ed) *Kant's Groundwork of the Metaphysics of Morals. A Critical Guide* (Cambridge University Press).

– (2010) ‘Reversal or retreat? Kant’s deductions of freedom and morality’ in Reath and Timmermann (2010), pp. 73-89.

– (2011) *Groundwork of the Metaphysics of Morals* (Translated by Jens Timmermann and Mary Gregor) (Cambridge University Press).

– (2013) ‘Autonomy and moral regard for ends’, in Sensen (2013a), pp.212-24.

Tolley, C (2006) ‘Kant on the Nature of Logical Laws’, *Philosophical Topics*, 34, pp. 371-407.

Van Kirk, C (1986), ‘Kant and the Problem of Other Minds’ in [*Kant-Studien*](http://philpapers.org/asearch.pl?pubn=Kant-Studien)*,* 77, pp. 41-58.

Velkley, R. (1994) ‘Introduction: Unity of Reason as Aporetic Ideal’ in Henrich (1994), pp. 1-15.

Willaschek, M (ed). (2000) *John McDowell: Reason and Nature. Lecture and Colloquium in Mϋnster* 1999 (LIT – Verlag).

– (2006): ‘Practical Reason. A Commentary on Kant’s Groundwork of the Metaphysics of Morals (GMS II, 412-417) in Horn and Schönecker (2006), pp. 121-38.

– (2010) 'The Primacy of Practical Reason and the Idea of a Practical Postulate' in Reath and Timmerman (2010), pp. 168-96.

Wood, A (ed). (1984) *Self and Nature in Kant’s Philosophy* (Cornell University Press).

– (1998) ‘Humanity as End in Itself’ in Guyer (1998).

– (1999) *Kant's Ethical Thought* (Cambridge University Press).

– (2008) *Kantian Ethics* (Cambridge University Press).

1. Jeanine Grenberg is more flattering; she calls it “the most beloved flawed argument in the history of philosophy” (2009: 335). [↑](#footnote-ref-1)
2. Korsgaard (1996a: 24) offers an alternative three line reconstruction of the argument. I discuss my departure from her reconstruction in both chapters 1 (§3.3) and 2 (§§1.3-1.5; §2.2). [↑](#footnote-ref-2)
3. Gregor translates ‘*Aufsuchung’* as ‘search for’, and ‘*Festsetzung’* as ‘establishment’. Timmermann departs from Gregor here, noting that ‘*Aufsuchung’* is a success term, and as such, is better captured by ‘identification’: “an investigator has not been engaged in Aufsuchung unless or until he has found what he was looking for” (2011: 162). As for *Festsetzung*, Timmermann notes that “this has connotations of making determinate and secure” (Timmermann 2011: 162). [↑](#footnote-ref-3)
4. Cf. IV: 403. 34-7: “[…] we have progressed in the moral cognition of common human reason to reach its principle, which admittedly it does not think of as separated in this way in a universal form, but yet always actually has before its eyes and uses as the standard of its judging.” [↑](#footnote-ref-4)
5. This is typically known as the formula of universal law. [↑](#footnote-ref-5)
6. These include: the law of nature formula: “act as if the maxim of your action were to become by your will a **universal law of nature**” (IV: 421. 18-20); the formula of humanity: “So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means” (IV: 429. 10-2); and the formula of autonomy: “the idea of the will of every rational being as a will giving universal law” (IV: 431. 17-8). [↑](#footnote-ref-6)
7. Kant notoriously claims that these are “at bottom only so many formulae of the very same law” (IV: 436. 8-9). [↑](#footnote-ref-7)
8. I will return to consider the relationship between the first and second formulations in the second chapter (§§1.4-1.5; §2.2). [↑](#footnote-ref-8)
9. Another aspect of this is that moral demands *override* what we desire; cf. Darwall (2006: 226-7). [↑](#footnote-ref-9)
10. He sets out the crucial contrast as follows: “[…] hypothetical imperatives [… say]: I ought to do something *because I want something else*. By contrast the moral and hence categorical imperative says: I ought to act in such or such a way, even if I did not want anything else.” (IV: 441.10-3) Kant sometimes makes this contrast out in other ways, which I will discuss when I criticise constructivism in chapter 2 (§1.4). [↑](#footnote-ref-10)
11. Although, as Allison recognises, Kant “is notorious for not keeping strictly to this distinction and for treating them as virtually equivalent” (1990: 278n1). Paton also notes that, “[t]hough Kant does not always keep the distinction clear, we should not forget that the principle of autonomy need not take the form of a categorical imperative” (1947: 199). This ambiguity (understandably) permeates the secondary literature. [↑](#footnote-ref-11)
12. See, for example: IV: 403. 18-34, IV: 411. 9-10, IV: 424 .34-5; Timmermann (2007: 149n59). [↑](#footnote-ref-12)
13. I will return to consider the relation between the moral law and the categorical imperative in more detail in chapter 2 (§2.2). [↑](#footnote-ref-13)
14. Other such passages can be found at: IV: 428. 34 – 429. 9, IV: 429n, IV: 431. 32-4, IV: 440. 20-8. The end of *Groundwork* II also sets the stage for *Groundwork* III (IV: 444. 35 – 445. 15). [↑](#footnote-ref-14)
15. Henrich (1988: 30-40; 1975: 322-9); cf. Allison (2011: 274). [↑](#footnote-ref-15)
16. The third reference is ambiguous. It could refer to Kant’s deduction of freedom, or the deduction of the categorical imperative, qua imperative, or both; cf. Allison (2011: 275). [↑](#footnote-ref-16)
17. There is also a third concern. Kant is troubled by the interest that we take in morality. He wonders how it is that we – rational animals that we are – take an interest in the purely rational claims of morality (IV: 449. 11-32). I shall address this issue in chapter 4. [↑](#footnote-ref-17)
18. Cf. Korsgaard (1996a: 160-71; 1996b: 92-8) and Hill (1998: 249-50). [↑](#footnote-ref-18)
19. See Allison (2011: 331-47), Schӧnecker (2006), Stern (2012: 68-99) and Timmermann (2007: 139-44). [↑](#footnote-ref-19)
20. Here, I opt for what Allison has called a ‘double-deduction’ reading of *Groundwork* III. Allison (2011: 274-84) defends this reading against Schӧnecker’s (2006; 2012) claim that Kant attempts only one deduction in *Groundwork* III, that of the categorical imperative, qua imperative. I depart from Allison slightly in that I take the first deduction to be a deduction of freedom, rather than a deduction of the moral law, but I think this might end up being a distinction without a difference: I claim that Kant is trying to entitle us to regard ourselves as free (and thereby standing under the moral law), where Allison claims that Kant is trying to entitle us to regard ourselves as standing under the moral law (because we are free). [↑](#footnote-ref-20)
21. See note 19. [↑](#footnote-ref-21)
22. IV: 447. 22-3. [↑](#footnote-ref-22)
23. Cf. Guyer (2009: 179): “Not what but how Kant intended to argue in Section III of the *Groundwork* is as controversial as any issue in Kant interpretation.” [↑](#footnote-ref-23)
24. Jens Timmermann finds another argument in the fourth subsection of *Groundwork* III, where Kant claims that even the most hardened scoundrel, when presented with examples of goodness wishes that they were so disposed (IV: 454. 21-7). Kant then claims that the scoundrel cannot easily bring this about, but wishes that he was free from his inclinations (IV: 454. 27-9), and concludes: “By this he therefore proves that, with a will free from impulses of sensibility, he transfers himself in thought into an order of things quite different from that of his desires in the field of sensibility” (IV: 454. 30-2). Timmermann thinks that Kant is arguing from moral consciousness to freedom here, and claims that the similarities with the approach of the second *Critique* are obvious (2007: 143-4). I go on to discuss the respective merits of Kant’s moral and non-moral arguments for freedom in the next section (§2). [↑](#footnote-ref-24)
25. Cf. Wood (2008: 124): “Kant’s theory of freedom presents us with a moving target throughout his writings, as is evicted by insuperable difficulties from each successive position he comes to occupy. Perhaps none of Kant’s forced peregrinations on the subject of free will ever takes him to a dwelling place that is even minimally inhabitable. I say these things, however, not to condemn Kant but rather because they may mark him out as the philosopher who understood the problem of freedom better than any other. In the end, Kant’s greatest insight regarding the problem of freedom may be that it is insoluble and a source of permanent torture to philosophy.” I agree with Wood that Kant’s views on freedom present a moving target, but I do not think the situation is as hopeless as he makes out. [↑](#footnote-ref-25)
26. I will return to discuss this shortly (§2), and in more depth in chapter 3 (§1.2; §2.2). [↑](#footnote-ref-26)
27. Michelle Kosch begins her book, *Freedom and Reason in Kant, Schelling and Kierkegaard*, by noting that (2006: 15): “The point of these chapters is not to present a unified interpretation of Kant’s thinking on freedom. Were such a unified interpretation readily available, the post-Kantian development could not have taken the form that it did.” Richard Velkley makes a similar point about Henrich’s work on Kant (1994: 10 : “Many of Henrich’s studies of Kant locate within his principal arguments areas of tension, or “fractures”, which are revelatory not of Kant’s logical negligence but of his penetration into the fundamental and necessarily problematic sources of human questioning. […] The most central passages of a philosopher’s thought, Henrich avers, are necessarily the most difficult and resistant to interpretation. They are also the passages that the most fertile soil for future philosophical developments.” By these standards, Kant’s deduction of freedom in *Groundwork* III is soil as fertile as any. [↑](#footnote-ref-27)
28. Cf. Henrich (1975: 324): “A deduction is called for, whether in cognition or in court, when the title to a right is in dispute. If the doubt is not explicit there is no basis for the deduction. The deduction is not undertaken for the sake of amplifying cognition, but for justifying it.” I will say a lot more about this in chapter 3 (§1.2). [↑](#footnote-ref-28)
29. Cf. Timmermann (2007: 146n55). [↑](#footnote-ref-29)
30. Gregor describes this as “something that belongs to no one” (1996: 103n), Timmermann as “a good that has no apparent owner” (2007: 146) [↑](#footnote-ref-30)
31. We can find a hint of this worry at the start of *Groundwork* II, where Kant talks of “philosophers in every age who have absolutely denied the actuality of this [moral] disposition in human actions, and attributed everything to more or less refined self-love, without however calling into doubt the correctness of the concept of morality because of this; rather, with intimate regret they made mention of the frailty and impurity of human nature that is indeed noble enough to take an idea so worth of respect as its prescription, but at the same time too weak to follow it” (IV: 406. 14-22). [↑](#footnote-ref-31)
32. In the second *Critique*, Kant famously remarks that freedom is the “*ratio essendi* [the ground of being, or existence] of the moral law” (V: 4n). [↑](#footnote-ref-32)
33. Kant expresses something close to this worry in his derivation of the formula of humanity: “without [objective ends …] nothing whatsoever of *absolute worth* could be found; but if all worth were conditional and hence contingent, then for reason no supreme practical principle could be found at all.” (IV: 428. 30-3) [↑](#footnote-ref-33)
34. In the introduction to the second *Critique*, Kant calls freedom “the *keystone* of the whole structure of a system of pure reason” (V: 3. 25 – 4. 1). Allison begins *Kant’s Theory of Freedom* with the claim that, “it is not exaggeration to claim that, at bottom, Kant’s critical philosophy is a philosophy of freedom” (1990: 1). And at one point, McDowell remarks that, “on Kant’s view, freedom was *the* great problem of modern thought” (2009: 43). [↑](#footnote-ref-34)
35. Henrich contends that Kant attempts a theoretical proof of freedom and the moral law throughout the 1770s and the first half of the 1780s culminating in *Groundwork* III (1994: 72-82). I agree with much of his assessment of this period, but want to distance myself from the talk of a theoretical proof of the validity of the moral law, especially in *Groundwork* III. I do so in the next section. Ameriks also finds this argument in *Groundwork* III: “[I]n the course of the composition of the *Groundwork* […] Kant ventures nevertheless to construct an argument toward spontaneity that is still based on a supposed implication of the phenomenon of judgment” (2003: 228). I will return to consider Ameriks’ diagnosis of *Groundwork* III in chapter 3 (§2.2). [↑](#footnote-ref-35)
36. I consider some of these passages in chapter 1 (§1.1, §3.1). [↑](#footnote-ref-36)
37. Cf. The second *Critique* (V: 29. 28 – 30. 30) and the *Religion* (VI: 26n). [↑](#footnote-ref-37)
38. I return to briefly consider this in reference to his *Review of Schulz* (VIII: 13-4) in chapter 1 (§1.1). Kosch notes that the moral argument is “discernible in the *Groundwork* and is dominant in the second *Critique*” (2006: 30). [↑](#footnote-ref-38)
39. Cf. Grenberg (2009: 335, 337, 344, 354n12), Henrich (1975: 330-5), and Timmermann (2007: 127, 127-8n17, 130; 132, 137, 137n38, 139, 143-4). [↑](#footnote-ref-39)
40. Allison (2011: 330n56). [↑](#footnote-ref-40)
41. He continues: “However, I could not yet bring it to such completeness here without introducing considerations of a wholly different kind and confusing the reader. On account of this I have availed myself of the label of a *Groundwork of the Metaphysics of Morals*, and not of a *Critique of Pure Practical Reason* instead.” (IV: 391. 29-33) [↑](#footnote-ref-41)
42. Reinhold (1792: 262-308). This objection was repeated later by Henry Sidgwick (1874: 511-16). [↑](#footnote-ref-42)
43. See Kosch (2006: 44-65) for a development of this criticism of Kant’s moral conception of freedom. [↑](#footnote-ref-43)
44. This is perhaps the centrepiece of his seminal work, *Kant’s Theory of Freedom* (1990: 1, 5-6, 133-6). [↑](#footnote-ref-44)
45. Although there are some passages in Kant’s later writings that suggest a non-moral conception of freedom; the most notable of these is the (so-called) Incorporation Thesis (VI: 23. 3 – 24. 5), which I will discuss in chapter 1 (§3.2). [↑](#footnote-ref-45)
46. For a detailed treatment of Reinhold’s (and Sidgwick’s) objection, see Fugate (2012). [↑](#footnote-ref-46)
47. I should note that certain parts of the *Groundwork* do suggest such a reading (see, for instance, IV: 449. 24-31). For those who read *Groundwork* III in this way, see: Allison (1998: 273), Ameriks (2000b: 75; 2003: 170), and Kosch (2006: 16; 32). It is worth noting that Ameriks (2010: 46) and Allison (2011: 309-10) have recently changed their mind on this score. [↑](#footnote-ref-47)
48. If Kant did attempt to address that sceptical challenge, he would fall prey to Prichardian worries; cf. Stern (2010) for a treatment of *Groundwork* III, Prichard and Korsgaard. [↑](#footnote-ref-48)
49. Allison notes that: “this interlocutor is concerned with the question of how the moral law can be binding or a categorical imperative possible, rather than why one should subject oneself to its dictates. Kant has no interest in the latter question because it implicitly reduces morality to a matter of hypothetical imperatives.” (2011: 310-11). [↑](#footnote-ref-49)
50. IV: 447. 6-7, IV: 450. 24; V: 29. [↑](#footnote-ref-50)
51. Like *Groundwork* III itself, my target is the curious and sympathetic meta-ethicist. [↑](#footnote-ref-51)
52. Cf. IV: 389. 11-5; Allison (2011: 303). [↑](#footnote-ref-52)
53. Cf. Allison (2011: 303). [↑](#footnote-ref-53)
54. Consider the following two sentences. Kant claims that when: “I ask only what I now have to do, and then freedom is a necessary practical presupposition and an idea under which alone I can regard commands of reason as valid.” (VIII: 13). Here, I read ‘commands of reason’ as involving both prudential and moral matters. He also writes that “the most confirmed fatalist, who is a fatalist as long as he gives himself up to mere speculation, must still, as soon as he has to do with wisdom and duty, always act *as if he were free”* (VIII: 13). I take it that Kant’s reference to wisdom *and* duty signals that he is talking about practical reason in general as well as morality. [↑](#footnote-ref-54)
55. Both Allison (1996: 98-104, 123-8, 2011: 302-9; 2012: 90-3, 111-4) and Korsgaard (1996a: 162-3; 1996b: 92-7, 100, 228-9, 252; 1996b: 92-7) mount compelling and extensive defences of this thought. Mieth and Rosenthal devote an entire paper to considering Kant’s argument in the second subsection of *Groundwork* III (2006). Ameriks (2003: 242-8), Beck (1960: 197-8), Henrich (1998: 311), Hill (1998: 263-7), Paton (217-22), Timmermann (2007: 138), and Tenenbaum (2012: 570-5) also discuss the argument. Almost all of these accounts (with the notable exceptions of Ameriks and Tenenbaum) are sympathetic; I will return to consider their objections in due course. McDowell also discusses something similar to this in his article ‘Two Sorts of Naturalism’ (especially 1998: 169-73). [↑](#footnote-ref-55)
56. Cf. Lowe (2008: 85-91) and Searle (1984: 94-9). Lowe argues that it would be irrational to believe in determinism – I will return to consider the plausibility of this strategy shortly (§1.2). [↑](#footnote-ref-56)
57. Cf. Peirce (1902: 113): “I do not admit that indispensability is any ground of belief. It may be indispensible that I should have $500 in the bank—because I have given checks to that amount. But I have never found that the indispensability directly affected my balance, in the least […] when we discuss a vexed question, we hope that there is some ascertainable truth about it, and that the discussion is not to go on forever and to no purpose. A transcendentalist would claim that it is an indispensible “presupposition” that there is an ascertainable true answer to every intelligible question. I used to talk like that, myself; for when I was a babe in philosophy my bottle was ﬁlled from the udders of Kant. But by this time I have come to want something more substantial” [↑](#footnote-ref-57)
58. Cf. McDowell (1996: 8). [↑](#footnote-ref-58)
59. Cf. Stern (2000: 85-6) and (Galen) Strawson (2010: 64): “The idea that a certain view may be inescapable and therefore somehow permissible or even correct, despite the fact that we are able to get into a position in which we can see it to be false – or so we think – is very problematic. Surely ‘irresistibility does not entail truth’ – even species-wide irresistibility?” [↑](#footnote-ref-59)
60. IV 447. 23. [↑](#footnote-ref-60)
61. Cf. A 84-7/B 116-9. [↑](#footnote-ref-61)
62. It is worth noting that for Kant, it seems that the necessity in the claim that we must act under the idea of freedom somehow satisfies both the *quid facti* and the *quid juris*. Consider the following from our lengthy passage (A): “Now I say: every being that cannot act otherwise than *under the idea of freedom* is actually free, in a practical respect, precisely because of that” (IV: 448. 4-6). Here, it seems that the necessity of acting under the idea of freedom is doing the work. It is “precisely because” we cannot but act under the idea of freedom that, in a practical respect, we are actually free. I will return to consider this claim (and the famous footnote – IV: 448n – that accompanies it) towards the end of this chapter (§3.3). [↑](#footnote-ref-62)
63. Allison pushes an analagous point against (P.F.) Strawson in an early paper of his (1969: 231). Thanks to Wayne Martin for pointing this out to me. [↑](#footnote-ref-63)
64. Ameriks also considers several different varieties of the ‘must’ involved here (2003: 247). He dismisses logical and metaphysical readings of it out of hand, before also dismissing a naturalistic reading of it. Here, he parallels our objection to note that: “it would be absurd for Kant, of all philosophers, to insists that wills as such are naturally *forced* to think of themselves in a particular way” (ibid). He also dismisses a deontological reading, as he also thinks that Kant “has just indicated that he is seeking an argument that precisely does not already have a moral premise” (ibid), before finally considering an epistemological reading of the ‘must’. On this reading, we must act under the idea of freedom because “the evidence for such a claim makes it the appropriate thought for a rational being” (ibid). Ameriks notes that this raises the question as to what this evidence is, before concluding: “I see now answer to this question, but I think that ultimately this is an embarrassment not for Kant but for contemporary apologists for his *Groundwork* argument. Kant himself went on to argue quite differently elsewhere, and this is itself a kind of evidence that ultimately he would agree with the weakness of this particular argument. One easily imagines that he came to see the frog smiling at the bottom of the mug precisely at the same time that he was working on the project of developing an even more Critical discussion of claims of our spontaneity in the massive revisions of his discussion of mind in the second edition of the first *Critique*. (Ameriks 2003: 247-8) I see no frog in the bottom of the mug, and go on to attempt to provide an answer to Amerik’s question throughout this chapter (especially in §3.1 and §3.3). [↑](#footnote-ref-64)
65. As a side note, I do not think that this counts against Korsgaard, as there are two different ways in which our freedom might be necessary. I think that, for Korsgaard, it is brute that we happen to be the type of creatures who must act under the idea of freedom, but given that we are such creatures, it is not brute that we must act under the idea of freedom. Indeed, this – despite what Korsgaard sometimes seems to say – is a necessary feature of reason itself. I shall return to this shortly (§3.1). [↑](#footnote-ref-65)
66. Tenenbaum (2012: 563-5) offers an alternative account of the problems with this strategy. [↑](#footnote-ref-66)
67. In a footnote in the second *Critique*, Kant writes: “Wizenmann […] disputes the authorization to conclude from a need to the objective reality of its object and illustrates the point by the example of a *man in love*, who, having fooled himself into an idea of beauty that is merely a chimera of his own brain, would like to conclude that such an object really exists somewhere.” (V: 143n) Kant grants Wizenmann this point, but wants to distance his postulates of pure reason from this approach as they concern “*a need of reason* arising from an *objective* determining ground of the will, namely the moral law” (V: 143n). As noted in the Introduction (§2), the argumentative strategy of the second *Critique* is not my concern in this thesis. Wizenmann’s objection does however parallel our own, and this parallel is instructive. [↑](#footnote-ref-67)
68. Willashek (2010) offers a discussion of Kant and James, where he claims that it can be rational to adopt certain beliefs on practical grounds. I agree with him on this, but still want to insist that necessity does not entail justification; cf. Saunders and Williams (draft), where we discuss Willashek’s account of practical grounds for belief, arguing that these grounds count as evidence. [↑](#footnote-ref-68)
69. Cf. Hookway (1999: 178). [↑](#footnote-ref-69)
70. Moore describes Kant’s claim that we must act under the idea of freedom as a “compulsion” (2012: 140) [↑](#footnote-ref-70)
71. Stern writes the following here (2000: 111): “[A]lthough the principle of ‘ought implies can’ means that people cannot be blamed for thinking or doing what they could not have thought or done differently, it does *not* mean that questions of right and wrong, rationality and irrationality, stop with what they are capable of controlling”. [↑](#footnote-ref-71)
72. Lewis White Beck puts forward an argument to this effect (1975: 14-29). He introduces this discussion as follows: “I propose now to show that if human beings are machines, no argument suffices to show that they are. […] If the machine theory is true, we cannot rationally argue for it. […] The machine theory is self-stultifying.” (Beck 1975: 14) [↑](#footnote-ref-72)
73. Cf. Enoch (2011: 218-20). [↑](#footnote-ref-73)
74. Sussman (2008: 54-6) and Tenenbaum (2012: 570-1) ascribe normative force to the ‘must’ in Kant’s claim that we must act under the idea of freedom. This approach appears to get things backwards, in that it seems to already presuppose the ability to act towards (or for the sake of) an ideal. Tenenbaum (2012: 575) acknowledges this, but seems happy to move away from the attempt to find a non-moral proof of freedom. He agrees with Kant’s later strategy that the moral law reveals our freedom to us, and then reads the claim we must act under the idea of freedom as a prescription that we ought to strive to act independently of sensible influences (Tenenbaum 2012: 572-3). [↑](#footnote-ref-74)
75. I will use capitalisation to refer to Reason as the specific capacity; cf. Korsgaard (2008: 2). [↑](#footnote-ref-75)
76. Cf. (A 314/B 370-1): “Plato noted very well that our power of cognition feels a far higher need than that of merely spelling out appearances according to a synthetic unity in order to be able to read them as experience, and that our reason naturally exalts itself to cognitions that go much too far from any object that experience can ever give to be congruent, but that nonetheless have their reality and are by no means merely figments of the brain.” [↑](#footnote-ref-76)
77. Here we see a clear parallel to (C): “As pure self-activity, it is elevated even above the understanding […] reason under the name of ideas shows a spontaneity so pure that thereby he goes far beyond anything that sensibility can ever afford him.” (IV: 452. 9-19). [↑](#footnote-ref-77)
78. See, for example, his discussion of the transcendental ideal (A 571-83/B 579-611). [↑](#footnote-ref-78)
79. Timmermann also thinks that this approach (along with (A) and any attempt at a non-moral proof of freedom in general) fails, as it does not establish that we possess wills (2007: 136, 137n38). I will return to address this in depth in chapter 3. [↑](#footnote-ref-79)
80. For a detailed treatment of these issues, see Laurence and Margolis (2013). [↑](#footnote-ref-80)
81. In his later treatment of *Groundwork* III, Allison ends up attempting a defence of Kant’s argument here (2011: 324-30). However, Reason’s capacity for ideas does not end up doing the work (Allison 2011: 325; cf. note 7), but instead it is the (supposed) self-certifying consciousness of possessing theoretical reason that involves spontaneity and thereby requires us to view ourselves as members of an intelligible world (Allison 2011: 329). This intelligible world turns out to be nothing more than a point of view assumed in order to conceive of our reason as practical, but nonetheless, from this point of view, we are entitled to presuppose that our reason is practical (2011: 329-30). This is an interesting and complicated line of thought, and I will return to in chapter 3. [↑](#footnote-ref-81)
82. Allison does not seem to think that the appeal to reason in IV: 452 (C) adds anything new of significance to Kant’s earlier appeal to reason in IV: 448 (A). He notes that in IV: 452 (C), Kant is considering the spontaneity of reason, whereas earlier (A) he was considering the spontaneity of the understanding (cf. Allison: 307-8). Allison recognises that this is a key distinction, but claims that it cannot be the salient difference between the two passages, as “Kant could easily have included the contrast between the two levels of epistemic spontaneity in the earlier discussion without materially changing things” (2011: 325). [↑](#footnote-ref-82)
83. Kant makes a related claim in *Reflexionen* 4220 (XVII: 462-3); cf. Henrich (1994: 80). [↑](#footnote-ref-83)
84. Timmermann translates ‘*Lenkung empfienge’* as to ‘receive guidance’, Gregor as ‘receive direction’. Quite literally, ‘Lenkung’ means something like ‘steering’ (which Timmermann points out in his editorial notes (2011: 172)). [↑](#footnote-ref-84)
85. Here, I must apologise to Thomas Land (2006), who has also offered a Spontaneity Thesis in Kant. I will say a little more about my departure from Land shortly. [↑](#footnote-ref-85)
86. Cf. B 129-30: “[…] the **combination** of a manifold in general can never come to us through the senses, and therefore cannot already be contained in the pure form of sensible intuition; for it is an act of the spontaneity of the power of representation”. [↑](#footnote-ref-86)
87. Thomas Land (2006) offers a clear account of this distinction. He argues that: “There are two distinct types of exercise of spontaneity, viz. judgment and sensible synthesis”, and calls this the Two-Species View of the Exercise of Spontaneity (Land 2006: 194). I agree with Land on this, but want to re-name things a little. I suggest we opt for two theses: The Spontaneity Thesis and the Synthesis Thesis. The former concerns the spontaneity involved in judgement, and the latter the synthesis of the manifold. [↑](#footnote-ref-87)
88. In the A-deduction, Kant makes a similar claim: “Now no cognitions can occur in us, no connection and unity among them, without that unity of consciousness that precedes all data of the intuitions, and relation to which all representations of objects is alone possible. Thus pure, original, unchanging consciousness I will now name **transcendental apperception**.” (A 107) [↑](#footnote-ref-88)
89. For an acute criticism of Kant here, see McDowell (2009: 148-53). [↑](#footnote-ref-89)
90. Allison (1990: 37-40) also notes the link between ‘taking as’ and the ‘I think’ – or a full treatment of ‘taking as’ in Kant, see Allison (1990: 37-8; 1996: 59-64; 94-5; 2012: 91, 113-4). [↑](#footnote-ref-90)
91. One route would be to claim that my mug, computer, and MS Word (so conceived) do not make judgements at all. This might well be the case, but it will hinge on what we mean by ‘judgement’. I wish to remain neutral on this, as even if my mug, computer, and MS Word can, in some sense, judge, these judgements do not stand as such *for them*. [↑](#footnote-ref-91)
92. Korsgaard seems attuned to this. At one point she claims that: “Self-consciousness is the source of reason” (Korsgaard 2009: 119), in that it frees us from instinct and creates the need for principles of reason (Korsgard 2009: 119-20); cf. Korsgaard (1996b: 92-4). [↑](#footnote-ref-92)
93. Cf. Allison (1996: 127-8) and Ameriks (2000: 190, 195). [↑](#footnote-ref-93)
94. They also do not allow Kant any direct intuition of our agency, which he appeals to surrounding this proof (XXVIII: 268-9). I will return to consider this and related issues in chapter 3. [↑](#footnote-ref-94)
95. He continues: “Attempts at simplification have usually produced only faint and distorted echoes of Kant’s meaning, and most of them, if they succeed in making him seem simple, do so only by making his views seem silly” (Beck 1960: 4). [↑](#footnote-ref-95)
96. Land recognises that Allison needs to be careful here, and ties this caution to his Two-Species View of the Exercise of Spontaneity (2006: 211): “[A] position like Allison’s […] is in danger of giving ‘the skeptic what he most desires’ (B168), […] if spontaneity is tied exclusively to judgment, then whatever structure intuition is said to have, this structure will not reflect objective requirements of cognition, but rather merely our subjective constitution.” [↑](#footnote-ref-96)
97. Here is another: “For to say that a certain experience is a *seeing that* something is the case, is to do more than describe the experience. It is to characterize it as, so to speak, making an assertion or claim, and – which is the point I wish to stress – to *endorse* that claim.” (Sellars 1956: 39; §16) [↑](#footnote-ref-97)
98. Brandom famously cashes out Sellars’ claim in terms of the ability to locate one’s judgments within inferences: “[…] to understand a sentence, to grasp a propositional content (a necessary condition of having a belief) is to place it in the space of reasons, to assign it an inferential role in the game of giving and asking for reasons, as entailing some other contents and being incompatible with others.” (Brandom 1997: 153) [↑](#footnote-ref-98)
99. Another Pittsburgh thinker, Engstrom, makes out this connection in terms of the self-sustaining nature of judgements (2009: 98-104). [↑](#footnote-ref-99)
100. He makes a similar move in his *Review of Schulz*: “Although he would not himself admit it, he has assumed in the depths of his soul that understanding is able to determine his judgement in accordance with objective grounds that are always valid and is not subject to the mechanism of merely subjectively determining causes […] hence he always admits freedom to think, without which there is no reason. In the same way he must also assume freedom of the will in acting, without which there would be no morals […]” (VIII:13-4). In this last sentence, we can, once again, read Kant as employing both moral and non-moral arguments for freedom in his *Review of Schulz*. [↑](#footnote-ref-100)
101. I read this as a re-statement of the conclusion of the argument from the previous sentence, and thus understand ‘principles’ here in a minimal sense. [↑](#footnote-ref-101)
102. Cf. The *Metaphysic of Morals*: “A rule that the agent himself makes his principle on subjective grounds is called his *maxim*.” (VI: 225. 2-3); “A *maxim* is a *subjective* principle of action, a principle which the subject himself makes his rule (how we wills to act).” (IV: 225. 34-6) [↑](#footnote-ref-102)
103. See Allison (1990: 35-42; 1996: 130-4; 2011: 114-20; 2012: 144-5). [↑](#footnote-ref-103)
104. In this passage, Allison overstates his case against desires. Cf. Shemmer (2007), who makes the case that some desires can give us reasons for action. It is worth noting though, that Shemmer’s account shares some of the features of the Spontaneity Thesis. Consider, for example, the following passage (2007: 328-9): “[…] a desire cannot be a mere impulse if it is to be reason giving […] For a desire to be reason giving it must have a minimal functional complexity; in particular, it must be such that it presents itself (or could present itself) as a consideration in the process of deliberation, and that means that its effect on one's action could be, at least temporarily, blocked by one's deliberative mechanisms. If an agent could not, at least temporarily, prevent a desire from moving her, then there would not be sufficient time for her to take the desire as a consideration in her practical deliberation. And that, I will claim, prevents a desire from being reason giving.” [↑](#footnote-ref-104)
105. Cf. McDowell (1998: 189): “[W]e cannot escape the burden of reflective thought – the obligation to weigh, by the best lights we have, the credentials of considerations purporting to appeal to reason – by a fantasy of having some suitable first-natural facts force themselves on us in a way that would bypass the need for thought.” [↑](#footnote-ref-105)
106. Cf. Ameriks (2003: 246): “Heavy weather is often made here of the fact that Kant ends his argument by saying that he is speaking of what is said ‘from a practical point of view’ (IV: 448). This sounds exciting, but I believe it may mean little more than that his discussion *concerns* what are in part practical issues, and that he is not committed to any theoretical *explanation of how* freedom actually works”. [↑](#footnote-ref-106)
107. Allison tries to downplay Kant’s talk of ‘proof’ here, suggesting that it should be replaced by the language of ‘presupposition’ (2011: 302-3). [↑](#footnote-ref-107)
108. For an excellent account of what metaphysics is, see Moore (2012: 1-22). [↑](#footnote-ref-108)
109. Cf. Allison (1996: 95): “[…] the “must” involved here is conceptual; the point being simply that unless one is aware of taking x as an F (recognizing it in a concept), one has not in fact taken it as such.” [↑](#footnote-ref-109)
110. McDowell complains that “Kant’s half-heartedness […] spoils his attempt at an equipoise between subjective and objective” (2009: 152), and in a footnote to this remarks: “Hence the unhappy addition of “for me” at B138” (2009: 152n13); Cf. McDowell (2009: 150n8). As we have just seen however (§3.1), the addition of “for me” is anything but unhappy – it signals the importance of self-consciousness. But McDowell is right to worry that Kant is too half-hearted. We can go beyond Kant (and especially Korsgaard’s and Allison’s Kant) to assert a strong link between reason and freedom. [↑](#footnote-ref-110)
111. See Allison (2011: 304; 304n3), and Korsgaard (1996a: 162): “The point is not that you must *believe* that you are free, but that you must choose *as if* you were free” [↑](#footnote-ref-111)
112. Kant offers a similar argument in the *Critique of Practical Reason* (V: 29. 12-22), which I will return to discuss shortly (§1.4). [↑](#footnote-ref-112)
113. Kant is less ambiguous on this in the second *Critique*, where he explicitly equates laws with objective principles in the strong sense. He opens the Analytic as follows: “Practical principles […] are subjective, or maxims, when the condition is regarded by the subject as holding only for his will; but they are objective, or practical laws, when the condition is cognized as objective, that is, as holding for the will of every rational being.” (V: 19. 7-12). This makes sense, given that, from the second *Critique* onwards, Kant primarily operates with an exclusively moral conception of freedom. [↑](#footnote-ref-113)
114. Timmermann (2007: 122) also points to Kant’s claim in the preface to the *Groundwork*, that physics concerns the laws of nature, whereas morality concerns the laws of freedom (IV: 387. 14-6). [↑](#footnote-ref-114)
115. Korsgaard has an elegant take on this: “[W]e may say that since the will is practical reason, it cannot be conceived as acting and choosing for no reason. Since reasons are derived from principles, the free will must have a principle.” (1996b: 97-98) For some critical discussion of this, see Nagel (in Korsgaard 1996b: 202-3). [↑](#footnote-ref-115)
116. There are two important related ambiguities, which concern Kant’s use of the terms ‘autonomy’ and ‘will’. I will discuss autonomy shortly (§1.3), and wills in chapter 3 (§1). [↑](#footnote-ref-116)
117. IV: 447. 6-7, IV: 450. 24, V: 29. 24-5. [↑](#footnote-ref-117)
118. Cf. Allison (1986; 1990: 201-213; 2011: 283-300). [↑](#footnote-ref-118)
119. Timmermann agrees (2007: 121): “A will is autonomous *by virtue of* being radically free and independent of external determination.” [↑](#footnote-ref-119)
120. Cf. Paton (1948: 39): “If the laws of freedom cannot be other-imposed (if we may use such an expression), they must be self-imposed.” [↑](#footnote-ref-120)
121. This argument also figures in both Korsgaard’s earlier (1996a: 160-7) and later works (2008: 11-2; 120-1; 319-20; 2009: 79-80, 153-4). [↑](#footnote-ref-121)
122. One worry is that Korsgaard’s regress appears unmotivated. Just because I can question everything does not mean that I am rationally required to do so. Right now, I desire a sandwich. I back up and consider why I desire a sandwich. It turns out that I am hungry, that I like sandwiches, and am the type of being for whom sandwiches satiate hunger. Do I have to back up from this? Of course, in a way, I can. I can reflect upon what it means to be such a being and question my liking of sandwiches, but in this case that does not appear rationally required – I just want a sandwich. Korsgaard’s regress seems unmotivated in this case. [↑](#footnote-ref-122)
123. In this reconstruction, Korsgaard draws more upon Kant’s formulation of the argument from freedom to autonomy in the second *Critique* rather than the *Groundwork*. It is implicit in a footnote (Korsgaard 1996b: 98) that she reads these formulations as equivalent, and for her purposes, this seems the case. I will return to consider the second *Critique* argument shortly (§1.4). [↑](#footnote-ref-123)
124. It is worth noting that whilst Korsgaard does not (explicitly) mention transcendental freedom, she does seem to operate with such a conception. Consider the following passage from *The* *Sources of Normativity*: “[Kant] defines a free will as a rational causality which is effective without being determined by any alien cause. Anything outside of the will counts as an alien cause, including the desires and inclinations of the person. The free will must be entirely self determining.” (Korsgaard 1996b: 97) Furthermore, in her earlier paper ‘Morality as Freedom’ (which she references in *The* *Sources of Normativity* as a place where the move from freedom to autonomy is “defended in greater detail” (Korsgaard 1996b: 98n7)), she calls the move from freedom to autonomy ‘the argument from spontaneity’ and remarks (in a endnote) that, “the important point here is just being uninfluenced by anything” (Korsgaard 1996a: 185n9). And finally, in an introductory chapter on Kant, she notes that, for him, a free will “is completely spontaneous” (Korsgaard 1996a: 25). [↑](#footnote-ref-124)
125. Both Timmermann (2006) and Langton (2007) criticise Korsgaard on this. For a (limited) defence of Korsgaard’s argument, see Stern (2011). [↑](#footnote-ref-125)
126. Cf. Allison (2011: 261); Hill (2013: 24-6) also offers an illuminating account of how Kant’s conception of autonomy relates to various contemporary conceptions. [↑](#footnote-ref-126)
127. Allison emphasises this conception of autonomy (2011: 261-6). [↑](#footnote-ref-127)
128. We will encounter a similar distinction in chapter 3 (§1), when we discuss the term ‘will’. [↑](#footnote-ref-128)
129. Cf. Ameriks (2003: 238): “[…] autonomy would take on such a general meaning (signifying any action of a being insofar as it is rational at all) that it would be especially hard to see how it must express anything as specific as the ‘principle of morality’”. [↑](#footnote-ref-129)
130. Wood points to a related problem with Kant’s passage, which he identifies as an equivocation between two senses of universalizability (1990: 164-5); cf. Patten (1999: 87-93). [↑](#footnote-ref-130)
131. See Allison (2011: 1–3, 56, 133, 237, 240, 262–3, 266, 288). Allison shifts between arguing that only a self-imposed law could be unconditionally normative and that only a self-imposed law could be unconditionally binding. These however, are separate issues. It might be the case that only a self-imposed law could be unconditionally binding, but that does not mean that only a self-imposed law could be unconditionally normative; cf. Stern (2012b: 1-99). [↑](#footnote-ref-131)
132. Cf. Timmermann (2007: 121): “It is *because* we are free that we must impose on ourselves our very own rational law […] Oddly enough, ‘can’ implies ‘ought’.” [↑](#footnote-ref-132)
133. We will shortly encounter a passage where Kant claims that the moral law must contain “*merely the form of willing*” (IV: 444. 29-30), and later another where Kant talks of us thinking of the world of understanding “merely according to its *formal* condition” (IV: 458.30). In the second *Critique*, Kant constantly talks of the mere form [*bloße Form*] of the law; see, for example: V: 27.14; V. 28. 31-4; V: 31. 12; V: 33. 14. [↑](#footnote-ref-133)
134. Similar passages can be found throughout Kant’s mature moral philosophy. See, for example: V: 27. 3-19; V: 33. 8-33; V: 41. 30-8; V: 62. 8-35. [↑](#footnote-ref-134)
135. He does the same thing elsewhere. Consider for instance the following footnote from the *Critique of Judgemen*t: “The final end, the promotion of which is imposed upon us by the moral law, is not the ground of duty; for this lies in the moral law, which, as a formal practical principle, guides us categorically, regardless of the object of the faculty of desire (the matter of the will), hence regardless of any end. This formal property of my actions, in which alone their inner moral worth consists […]” (V: 471n) [↑](#footnote-ref-135)
136. Even Rawls recognises that Kant’s argument suffers from a neglected alternative (2000: 229; 235-6); cf. Stern (2012b: 21-2), who provides an excellent response to Rawl’s attempt to buttress Kant’s argument, and Guyer (2005: 152; 2007b: 47-8; 75). We will cover similar ground shortly. [↑](#footnote-ref-136)
137. See also his discussion of ends that are also duties in the *Metaphysics of Morals* (VI: 384. 33 – 385. 29). [↑](#footnote-ref-137)
138. Furthermore, Kant says that we are to advance this end (IV: 430.17; cf. VI: 384. 33 – 388. 30). [↑](#footnote-ref-138)
139. Guyer (2005: 152-4) thinks that this makes Kant’s formalism consistent with the independent unconditional value of rational nature. In one sense he is right – a weak account of Kant’s formalism is certainly consistent with this – but I am interested in the plausibility of a strong anti-realist account of Kant’s formalism. [↑](#footnote-ref-139)
140. Cf. McDowell (2009: 139-40; 201-3). [↑](#footnote-ref-140)
141. There is a third element – the will’s individuality – which I will say a little more about shortly. [↑](#footnote-ref-141)
142. Cf. EL: §7, p.41: “The first two moments – that the will can abstract from everything and that it is also determined (by itself or something else) – are easy to grasp, because they are in themselves, moments of the understanding and devoid of truth. But it is the third moment, the true and speculative (and everything true, in so far as it is comprehended, can be thought of only speculatively), which the understanding refuses to enter into, because the concept is precisely what the understanding always describes as incomprehensible”. [↑](#footnote-ref-142)
143. Cf. PR: §5Z, p.38-9: “This negative freedom or freedom of the understanding is one-sided, but this one-sidedness always contains within itself an essential determination and should therefore not be dismissed; but the defect of the understanding is that it treats a one-sided determination as unique and elevates it to supreme status.” [↑](#footnote-ref-143)
144. Cf. Allison (1996: 188), who claims that Hegel rejects Kant’s doctrine of transcendental freedom. I will return to consider this, and the complicated issue of moral motivation, in more detail in chapter 4. [↑](#footnote-ref-144)
145. Cf. Stern (1990). [↑](#footnote-ref-145)
146. Cf. Stern (2013: 185-6). [↑](#footnote-ref-146)
147. Cf. Charles Taylor (1975: 373): “The dilemma of radical freedom can be restated succinctly as follows: if freedom is to renounce all heteronomy, any determination of the will by particular desires, traditional principle of external authority, then freedom seems incompatible with any rational action whatsoever. For there do not seem to be any grounds of action left, which are not wholly vacuous, that is which would actually rule some actions in and others out, and which are not also heteronomous.” Patten discusses the problems this might cause Kant and also Hegel himself (1999: 86-7). [↑](#footnote-ref-147)
148. Elsewhere Hegel notes that: “With the Germans, [this conception of freedom] remained tranquil theory; but the French wished to put it into practice” (PH: 463). The Germans, and primarily Kant, end up with a limited conception of freedom. Hegel thinks that if we try to actualise this, the results will be disastrous, which he sees in the French Revolution. A wholly universal conception of freedom views all particularity and determinacy as in need of elimination. Hegel worries that: “The sole work and deed of universal freedom is therefore *death*, a death too which has no inner significance or filling, for what is negated is the empty point of the absolutely free self. It is thus the coldest and meanest of all deaths, with no more significance than cutting off a head of cabbage or swallowing a mouthful of water.” (PS: §590, p.360) As interesting as it is though, Hegel’s critique of the French Revolution is not our concern here; cf. PR: §5, pp.5-6; PS: §§582-595, pp.355-363; Stern (2013: 184-190). [↑](#footnote-ref-148)
149. He continues to note (PS: §590, p.359): “But the supreme reality and the reality which stands in the greatest antithesis to universal freedom, or rather the sole object that will still exist for that freedom, is the freedom and individuality of actual self-consciousness itself.” [↑](#footnote-ref-149)
150. We will see shortly (§1.5) that Timmermann also appears subject to Hegel’s criticism, when he contends that (2007: 125): “Acting *for the sake of the good* would be a mark of heteronomy”. [↑](#footnote-ref-150)
151. Cf. Stern (2010b: 8-9), Street (2010: 370-1) and Lenman and Shemmer (2012: 2-3). [↑](#footnote-ref-151)
152. Amongst others, Bojanowski (2012), Korsgaard (2008: 319-26) and Sensen (2009) argue in favour of reading Kant as a constructivist, where Guyer (2005: 152; 2007: 47-8), Langton (2007), Stern (2012b: 28-9) and Wood (1999: 113-4) argue in favour of reading him as a realist. [↑](#footnote-ref-152)
153. Cf. Formosa (2011). [↑](#footnote-ref-153)
154. Cf. Korsgaard (1996b: 35-7; 112). [↑](#footnote-ref-154)
155. When I refer to ‘constructivism’, I am talking about constructivism *all the way down*, that is, the position that there is no moral value prior to the procedures of construction. [↑](#footnote-ref-155)
156. Cf. Hegel (PR: §135, pp.162-163; PS: §§ 430-1, pp.257-259). [↑](#footnote-ref-156)
157. Cf. Stern (2012a: 75-7)on the Hegelian side, and Allison (2011: 139-40), Engstrom (2009: 184-239), and Korsgaard (1996a: 86-7, 95) on the Kantian. [↑](#footnote-ref-157)
158. See Geiger (2010) for another attack on the idea that the formula of universal law can generate moral content by itself. [↑](#footnote-ref-158)
159. In the *Reason as testing laws* section of the *Phenomenology*, Hegel remarks that: “It would be strange, too, if tautology, the maxim of contradiction, which is admitted to be only a formal criterion for the cognition of theoretical truth, i.e. something which is quite indifferent to truth and falsehood, were supposed to be more than this for the cognition of practical truth.” (PS: §430, p.259) The principles of non-contradiction and universalizability are only formal criteria for the cognition of theoretical truth. Alongside the traditional issue, part of the empty-formalism objection is a general challenge to the Kantian who takes these formal criteria to be more than that for the cognition of practical truth. Hegel also writes that: “It is not […] because I find something is not self-contradictory that it is right; on the contrary, it is right because it is what is right.” (PS: §437, p.262) Of course, this is still early days in the *Phenomenology*. We are yet to pass through the various guises of *Geist*, not to mention Religion, and there is always the question of whether we are encountering Hegel’s voice and not just that of another form of consciousness; in this section (PS: §437, pp.261-262), we appear to be encountering the voice of the Greeks. Nevertheless, this textual wrinkle aside, in this section of the *Phenomenology*, we can find a realist challenge to Kant. [↑](#footnote-ref-159)
160. Cf. Stern (2012a: 77): “In addition to these ways in which Hegel’s empty formalism objection may continue to be pressed, it can also be argued that there is a yet deeper worry underlying it, which is that the FUL is inadequate as the supreme principle of morality taken on its own, because something more substantive is required if we are to understand why there is any moral significance in acting on maxims that are universalisable – why this matters from a moral point of view. The problem might be put as a dilemma for the Kantian: on the one hand, he could answer this question by relating the FUL to considerations such as equality, fairness, or free-riding, but then it is not clear why “treat others fairly” is not the supreme moral principle and the FUL merely a test for whether or not in acting a certain way one would be doing so; or he could treat the FUL as somehow prior in itself, but then make its moral relevance mysterious.” [↑](#footnote-ref-160)
161. Cf. Guyer (2005: 152-4), Langton (2007), Stern (2012b: 28-9) and Wood (1999: 107, 111-4). [↑](#footnote-ref-161)
162. It is sometimes claimed that the hypothetical imperative is not a principle of practical reason, but instead a principle of theoretical reason; cf. V: 172. 23-36. This would make the categorical imperative distinctive, but it still does not help explain why it is *normatively* distinctive. There are various requirements of reason, which we can violate in our actions: I can contradict myself, I can fail to take the means to an end I desire. And however we classify these requirements of reason, we still lack a way to differentiate them *normatively* if we treat them all as merely formal. [↑](#footnote-ref-162)
163. I think Kant is right in thinking that immoral conduct is irrational, but of course that does not mean that irrational conduct is immoral. [↑](#footnote-ref-163)
164. Stern (2012a: 90n14) points out that even Kantian Constructivists sometimes write this way. Korsgaard, for instance claims that (1996a: 92): “What the test show to be forbidden are just those actions whose efficacy in achieving their purposes depend upon their being exceptional”. O’Neill also argues that (1989: 156): “The reason why a universalizability criterion is morally significant is that it makes of our case no special exception”. [↑](#footnote-ref-164)
165. For a detailed attempt to criticise Engstrom’s constructivism, see Saunders (2016). [↑](#footnote-ref-165)
166. Cf. Engstrom (2013: 141): “[moral realists] ask: Are these procedurally determined judgments correct because they agree with what we really should do, or is what we really should do decided by these judgments? Constructivists do address this question, but their answers have left realists unsatisfied.” [↑](#footnote-ref-166)
167. Cf. Engstrom (2009: 132). [↑](#footnote-ref-167)
168. I am not reading the ‘to us’ in this question in an egoist sense, as this would fall afoul of the Prichardian worries we discussed in §3 of the Introduction. I will return to say a bit more about this at the end of this chapter (§2.2). [↑](#footnote-ref-168)
169. There is the additional issue of the binding nature of the moral law. As noted in the introduction (§1), Kant views as us sensibly affected rational beings, who thereby experience the moral law as binding – a categorical *imperative*. This raises the question of where this binding comes from: is it internal or external? However, as Stern has recently argued, we can combine a realist account of value with a constructivist account of the bindingness of the moral law (2012b: 1-99). [↑](#footnote-ref-169)
170. Cf. O’Neill (1989: 29): “If the standards of practical reasoning are fundamental to all human reasoning, then any vindication of these standards is either circular (since it uses those very standards) or a failure (since it is not a vindication in terms of the standards said to be fundamental.” [↑](#footnote-ref-170)
171. In *Groundwork* II, Kant claims that “in moral judging it is better always to proceed by the strict method, and make the foundation the universal formula of the categorical imperative” (IV: 436. 29-32). [↑](#footnote-ref-171)
172. Cf. Lenman and Shemmer (2012: 2): “We would, however, suggest that one feature of Rawls's constructivism that has appealed to many who have since styled themselves constructivists is his attempt to make good sense of moral justification in a way that is maximally free of any controversial metaphysical suppositions.” [↑](#footnote-ref-172)
173. Korsgaard (1996b: 37-8) also invokes Mackie against the realist. [↑](#footnote-ref-173)
174. It is not so clear that we can only look upon this with suspicion. I will leave this aside here though, as arguably Kant would. [↑](#footnote-ref-174)
175. One option here is to plump for an ontologically light reading of Kant, but that creates more problems than it solves. I will make this case in chapter 3. [↑](#footnote-ref-175)
176. Cf. Bojanowski (2012: 2) and Sensen (2013: 277-8). [↑](#footnote-ref-176)
177. For a full realist treatment of this argument, see Stern (2013b: 29-37). [↑](#footnote-ref-177)
178. Cf. Engstrom (2009: ix): “Kant breaks sharply with the received view that reason’s function in the practical sphere is to guide us in action on the basis of its antecedent apprehension of a final end that has an independent footing in nature. Kant rejects this picture, claiming that practical reason’s most basic imperatives, those of morality, are categorical rather than hypothetical in form. Human reason must accordingly be conceived as autonomous, as the sole source of its principle of action.” Sensen makes a similar claim (2013: 270-9). [↑](#footnote-ref-178)
179. See Allison (2011: 1–3, 56, 133, 237, 240, 262–3, 266, 288). [↑](#footnote-ref-179)
180. Cf. Paton (1948: 39): “If the laws of freedom cannot be other-imposed (if we may use such an expression), they must be self-imposed.”; and Timmermann: “A will is autonomous by virtue of being radically free and independent of external determination” (2007: 121). [↑](#footnote-ref-180)
181. I think this thought emerges from a blurring of the two ways of making out the distinction between categorical and hypothetical imperatives, which results in a parallel blurring of the distinction between heteronomy and autonomy. [↑](#footnote-ref-181)
182. Something like this also appears to be Kant’s strategy in the second *Critique*; see V: 21. 14 – 22. 3; V: 33. 8-33; V: 62. 8-35. [↑](#footnote-ref-182)
183. Timmermann goes on to claim that (2013: 220): “In the former case his action is morally and immediately good; in the latter it is not because the value of the action depends on whether the end is in fact realized. Unlike duty, interest that rests on inclination is essentially ‘consequentialist’.” The neglected alternative disarms this complaint. One can act for the sake of an independently valuable end, and the moral worth of this action can reside in one’s intentions, and not whether the end is realised or not.” [↑](#footnote-ref-183)
184. For the classical treatments of this, see McDowell (1998: 79-80) and Nagel (1970: 29-30). [↑](#footnote-ref-184)
185. Moreover, even if the internalist about motivation were right, this would generate a peculiar result for the Kantian. It would mean that the claim that the moral law must be merely formal only applies for beings like us. It is due to our particular constitution, the (supposed) fact that we can only either act upon desires or a merely formal law. This would not apply to other beings that could be motivated by objects of unconditional value – if the motivational issue is doing the work, then for these beings, the moral law does not have to be merely formal. And then, whether the moral law needs to be formal for creatures like us would be turn out to be a contingent issue, which depends on whether we can only be motivated by desires (and a merely formal law), or not. [↑](#footnote-ref-185)
186. I know of one other Kantian argument against realism, which concerns the distinction between theoretical and practical reason. Engstrom (2009: 119; 2013: 139; 145)) and Korsgaard (2008: 325-6) have both pushed the idea that acknowledging that reason is practical somehow counts in favour of constructivism. Here is Engstrom (2013: 139): “Constructivism rests on the idea that reason has a practical as well as theoretical use.” The basic idea seems to be that the relationship between knowledge and its object is the opposite when it comes to theoretical and practical knowledge: “In the theoretical case the knowledge depends for its actuality on the actuality of its object; hence the object must, in order to be known, be "given from elsewhere" by affecting the mind. In the practical case the relation is the reverse: here the actuality of the object - as determined - depends on the actuality of the knowledge.” (Engstrom 2009: 119). We can grant that practical knowledge involves bringing about its object, but that is not the point in contention between realism and constructivism. Realists can accept this claim, but still ask where the value of the object comes from. Kant’s distinction between ends to be effected and independently existentent/self-sufficient ends (IV: 437. 21-7) seems important here. Rational agents have independent value, they are existent ends. I do not bring that into being, but when I act morally towards them, I bring a certain treatment of them into being. However, that is not something that is going to settle the realist/constructivist issue. The question remains as to where this treatment gets its value from – whether it is valuable because it respects an independently existing end, or not. That is precisely the point in question, and one that is not settled by the respective existential relationships between practical and theoretical knowledge, and their objects. For a full treatment of Engstrom’s position, see Saunders (2016: §4). [↑](#footnote-ref-186)
187. Schӧnecker (2006: 303; 2013: 230, 238-9). [↑](#footnote-ref-187)
188. Cf. IV: 414. 1-3: “Thus a perfectly good will would just as much stand under objective laws (of the good), but it could not be represented as thereby necessitated actions that conform with laws”. [↑](#footnote-ref-188)
189. Here, I side with Allison (2011: 162) and Wood (1999: 173; 379-80n25) against Tolley (2006: 377-80). [↑](#footnote-ref-189)
190. Cf. VI: 222. 35 – 223. 3: “A categorical imperative, because it asserts an obligation with respect to certain actions, is a morally practical *law*. But since obligation involves not merely practical necessity (such as a law in general asserts) but also *necessitation*”. [↑](#footnote-ref-190)
191. Cf. Schӧnecker (2013: 231): “The moral law of a holy will […] is descriptive, not prescriptive.” [↑](#footnote-ref-191)
192. Again, cf. IV: 414. 1-3. [↑](#footnote-ref-192)
193. Elsewhere, Kant gives indirect arguments for promoting one’s own happiness (IV. 399. 3-6; V: 93. 15-9; VI: 388. 17-30). [↑](#footnote-ref-193)
194. In his treatment of the highest good in the second *Critique*, Kant seems to suggest as much; cf. V: 110. 18-26: “That *virtue* (as worthiness to be happy) is the *supreme condition* of whatever can even seem to us desirable and hence of all our pursuit of happiness and that it is therefore the *supreme* good has been proved in the Analytic. But it is not yet, on that account, the whole and complete good as the object of the faculty of desire of rational finite beings; for this, *happiness* is also required, and that not merely in the partial eyes of a person who makes himself an end but even in the judgment of an impartial reason.” [↑](#footnote-ref-194)
195. Schönecker complains that everyone has overlooked this: “This thesis has been widely misunderstood; as a result, Kant’s overall argument (deduction) in GMS III has been misunderstood too (often not recognised at all, as a matter of fact). His overall argument in GMS III has always been reconstructed as follows: A free will is a will under moral law; freedom must be presupposed as a quality of the will of all rational beings; human beings are rational beings; therefore, the human will as a free will is under the moral law, which is to say the categorical imperative is valid.” (Schönecker 2006: 302). He overstates his case against the literature. Kant’s argument has not always been reconstructed in this way (see, for example, Allison 1990: 216; 278n1). The argument often follows this trajectory, but also often ends with us standing under the moral law, and not with the validity of the categorical imperative. [↑](#footnote-ref-195)
196. This is a source of disagreement between Schönecker and Allison, I am unsure of how much hangs on it. The interesting question seems to be whether the moral law is normative or prescriptive, rather than whether it is analytic or synthetic – whatever that means. For what it is worth, I do not think it is at all clear that the moral law itself is analytic; cf. Allison (2011: 280): “[…] it does not follow from the fact that morality together with its principle (the moral law or principle of autonomy) follows from an analysis of the concept of freedom that the moral law itself is analytic.” [↑](#footnote-ref-196)
197. This is probably grossly unfair on the actual Leibniz, but hopefully the point still holds. [↑](#footnote-ref-197)
198. Cf. VI: 222. 35 – 223. 3: “A categorical imperative, because it asserts an obligation with respect to certain actions, is a morally practical *law*. But since obligation involves not merely practical necessity (such as a law in general asserts) but also *necessitation*”. [↑](#footnote-ref-198)
199. On this point, in a famous footnote in the *Critique of Judgement*, Kant tells us that: “I deliberately say ‘**under** moral laws.’ The final end of creation is not the human being **in accordance with** moral laws, i.e., one who behaves in accordance with them. For with the latter expression we would say more than we know, namely, that it is in the power of an author of the world to make it the case that the human being always behaves in accordance with moral laws” (V: 448-9n). Of course, there are passage in Kant that pull towards a different reading, but I am happy to claim that the moral law is normative and prescriptive for us. For an alternative reading of this passage (and others similar), see Fugate (2010: 3-6). [↑](#footnote-ref-199)
200. Cf. Reath (2006: 1): “Not much is gained by saddling a great philosopher with a philosophical view that has little to recommend it.” [↑](#footnote-ref-200)
201. Recall Timmermann (2007: 121): “It is *because* we are free that we must impose on ourselves our very own rational law […] Oddly enough, ‘can’ implies ‘ought’.” [↑](#footnote-ref-201)
202. Cf. Allison (2011: 299-300). [↑](#footnote-ref-202)
203. Again, cf. Stern (2013b: 29-37). [↑](#footnote-ref-203)
204. For what it is worth, I think that animals also serve as the ground of unconditionally normative claims, in that the ways in which we ought to treat them do not entirely depend upon what we desire. I hope to make this case elsewhere. [↑](#footnote-ref-204)
205. Sensen makes some similar points (2009: 264). [↑](#footnote-ref-205)
206. Cf. Korsgaard (2008: 109-10; 2009: 76-80). [↑](#footnote-ref-206)
207. Here is Enoch himself (2006: 179): “Classify my bodily movements and indeed me as you like. Perhaps I cannot be classified as an agent without aiming to constitute myself. But why should I be an agent? Perhaps I can’t act without aiming at self-constitution, but why should I act? If your reasoning works, this just shows that I don’t care about agency and action. I am perfectly happy being a shmagent—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency but not of shmagency) of self-constitution. I am perfectly happy performing shmactions—nonaction events that are very similar to actions but that lack the aim (constitutive of actions but not of shmactions) of self-constitution.” [↑](#footnote-ref-207)
208. We find a similar challenge in Hegel. In the *Philosophy of Right*, after he lays the empty-formalism charge, he repeats his challenge about property, but also mentions human life: “The fact that *no property* is present is in itself no more contradictory than is the non-existence of this or that individual people, family, etc., or the complete *absence of human life*. But if it is already established and presupposed that property and human life should exist and be respected, then it is a contradiction to commit theft or murder; a contradiction must be a contradiction with something, that is, with a content which is already fundamentally present as an established principle.” (PR: §135, p.162) Moreover, this seems important to Kant. One of the four examples presented in *Groundwork* II involves suicide, something he wants to show is immoral (IV: 421. 24 – 422. 14 and IV: 429. 15-28, respectively). When it comes to deriving this from the formula of humanity, he tells us that: “[…] a human being is not a thing, hence not something that can be used *merely* as a means, but must in all his actions always be considered as an end in itself. Thus the human being in my own person is not at my disposal, so as to maim, to corrupt, or to kill him” (IV: 429. 20-5). This aligns nicely with a realist reading of the formula of humanity. So conceived, our rational nature has absolute worth. It is an end in itself, and our agency is thereby to be respected. (A similar point has recently been pressed hard by Rae Langton (2007).) [↑](#footnote-ref-208)
209. Hegel (LH: 462-3); cf. Patten (1999: 98). [↑](#footnote-ref-209)
210. Cf. Patten (1999: 99). [↑](#footnote-ref-210)
211. I also think that Kant adopts a similar approach in the second *Critique* with the fact of reason, and will return to say something more about this in the next chapter. [↑](#footnote-ref-211)
212. In the full passage, Sensen also mentions Divine Command Theory, which I have removed for ease of reading. [↑](#footnote-ref-212)
213. Cf. Cohen (1996: 178-183). [↑](#footnote-ref-213)
214. Cf. IV: 427. 19-20. This ambiguity relates to the ambiguity in Kant’s use of the term ‘law’ that we discussed at the beginning of the previous chapter. Ameriks (2012: 185-6) draws attention to a similar ambiguity. [↑](#footnote-ref-214)
215. For a fuller treatment of this, see Willaschek (2006: 123-9). [↑](#footnote-ref-215)
216. At one point in *Groundwork* I, Kant mentions the possibility that reason might not “break forth into *practical* use” (IV: 395. 22-3). [↑](#footnote-ref-216)
217. This is one reason why the second subsection of *Groundwork* III is sometimes called the ‘preparatory argument’, as it leaves open the question of whether we are rational beings with wills. Cf. IV: 447. 25, and Allison (1990: 216-7), who coined this term. [↑](#footnote-ref-217)
218. Henrich does not make this point lightly: “It is one of his most fundamental premises that the faculty of desire has an independent origin from that of cognition, so that reason does not imply will” (1975: 313). [↑](#footnote-ref-218)
219. Allison (1990: 227-9; 2011: 309, 324-30; 2012: 91-2, 115), Ameriks (2000: 203-4; 2003: 171-4), Korsgaard (1996a: 170), Timmermann (2007: 136-7, especially 137n38), and Tenenbaum (2012: 572-5) all seem strongly influenced by Henrich here. [↑](#footnote-ref-219)
220. Cf. XXVIII: 263. [↑](#footnote-ref-220)
221. Cf. Allison (1996: 127-8) and Ameriks (2000: 190, 195). [↑](#footnote-ref-221)
222. Kant goes on to argue from the fact that there are imperatives to our freedom (XXVIII: 269). Once again, we find traces of both a moral and a non-moral argument for freedom in Kant. [↑](#footnote-ref-222)
223. Here, I focus on the first edition of the *Critique* [1781], as it pre-dates the *Groundwork*. For Kant’s treatment of the paralogisms in the second edition of the *Critique* [1787], see B 399-432. [↑](#footnote-ref-223)
224. Cf. B 131-6. [↑](#footnote-ref-224)
225. Cf. Ameriks (2000: xvii-xix, 213-4; 2003: 227-8). [↑](#footnote-ref-225)
226. Cf. B 130, 132, 150. [↑](#footnote-ref-226)
227. Elsewhere, Kant describes transcendental (or cosmological) freedom as follows (A 533/B 561): “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.“ [↑](#footnote-ref-227)
228. Allison also mentions these passages (2011: 327n53; 2012: 115n10). [↑](#footnote-ref-228)
229. Drawing upon one of Kant’s Reflections (R 5442; XVIII: 183. 13-8),Henrich makes this out in terms of a shift in confidence concerning the relationship between logical and transcendental freedom. He contends that in his pre-critical days, Kant thought the former – the spontaneity of the understanding – entailed the latter, but came to doubt this, due to the possibility that we might lack wills (Henrich 1975: 311-6). [↑](#footnote-ref-229)
230. Allison (1990: 63-5; 2011: 327-30). [↑](#footnote-ref-230)
231. Elsewhere, Allison refers to the claim that we have wills as a “seemingly innocuous proposition” (2011: 309). [↑](#footnote-ref-231)
232. At one point, he does attempt an answer on Kant’s behalf (2011: 329-30). This involves a complicated appeal to reason’s capacity to form ideas and our membership in the intelligible world (the argument we discussed in §1 of chapter 1). I do not think this argument works, as it assumes that we are *entitled* to take ourselves as occupying the practical point of view. [↑](#footnote-ref-232)
233. Cf. Allison (2011: 317). [↑](#footnote-ref-233)
234. In a previous paper, I argued that our having wills was immune to doubt from the practical standpoint, and concluded that, in this sense, practical reason was also self-certifying (Saunders 2014). I now think this approach is misguided, as I take the crucial issue to concern what *entitles* us to the practical standpoint in the first place. [↑](#footnote-ref-234)
235. The following remark comes close to this: “Nothing worse could happen to these labours than that someone should make the unexpected discovery that there is and can be no a priori cognition at all. But there is no danger of this. It would be tantamount to someone's wanting to prove by reason that there is no reason.” (V: 12. 6-10). [↑](#footnote-ref-235)
236. Cf. Peirce (1877: 114-5). [↑](#footnote-ref-236)
237. Stern draws out the connection with Peirce here as follows (2009: 220): “The point here does appear to be a Peircean one: The Cartesian philosopher insists on placing a burden on the inquirer to reflect on his capacities prior to inquiry (what Hume identified as Descartes’s ‘antecedent’ scepticism), while in fact it is pointless to feel any such burden: we would do better to just get on with inquiring, and if this goes well we will know our cognitive capacities are in order anyway, and if it goes badly, there is no reason to think this prior investigation would have helped. As the Cartesian cannot really tell us in advance either way, or do anything to improve our prospects of success, it seems fruitless to be moved by his concerns, and better to just ‘get on with the work itself ’.” [↑](#footnote-ref-237)
238. Cf. Ameriks (2003: 10-12). One place where Kant is sometimes thought to address the Cartesian sceptic is the Refutation of Idealism. I follow Ameriks (2003: 17-20) and Stern (2000: 142-64) in thinking that this is not the best way to read Kant’s argument. [↑](#footnote-ref-238)
239. Cf. Hookway (1999: 175). [↑](#footnote-ref-239)
240. Maimon wants to push the issue of the *quid facti*. He challenges Kant thus: “[T]he question is not whether we can use it [the concept of cause] legitimately, which is the question: *quid juris*?, but whether the fact is true, namely that we do use it with actual objects. Yes, it will be said, the fact is indubitable. We say, for example: the fire warms (makes warm) the stone and this signifies not merely perception of the succession of the two appearances in time, but the necessity of this succession. But to this **David Hume** would reply: it is not true in this case that I perceive a necessary succession” (Maimon 1790: 42). He concludes with Hume and against Kant: “**David Hume** (or his representative) can gladly admit this. He will say: the concept of cause is not in the nature of our thought in general such that it would also occur [*stattfinden*] in symbolic cognition, and it is also not grounded in experience in the sense in which **Kant** uses this word; consequently there are no experiential propositions properly so called [*eigen*] (expressing necessity) and if I say this concept is taken from experience, I understand by this mere perception containing a merely subjective necessity (arising from habit) that is wrongly passed off as an objective necessity. So, in order to prove the fact itself against **David Hume**, we would have to be able to show that when they first have this perception, children too instantly make the judgement that the fire is the cause of the warming of the stone. But this will be difficult to do.” (Maimon 1790: 43). [↑](#footnote-ref-240)
241. Cf. Ameriks (2003: 20). [↑](#footnote-ref-241)
242. Cf. Allison (2012: 118). [↑](#footnote-ref-242)
243. There are two different modalities in this claim: Kant is claiming that freedom is *speculatively* possible, and *practically* necessary. I want to claim that this evidence shows that freedom is *actual*; as I argued towards the end of the first chapter (§3.3), we should not shy away from a metaphysical account of freedom, and I will return to make this point again towards the end of this chapter. [↑](#footnote-ref-243)
244. This is not to affirm the consequent: if we were free, things would look like this; things look like this; therefore, we are free. Instead, I mean to point towards what the evidence would look like, if we were free. I think it is important to ask: If we were rational agents, what type of evidence would there be for this? A robust phenomenology of agency would be as good as any. [↑](#footnote-ref-244)
245. Cf. Ameriks (2000: 190). [↑](#footnote-ref-245)
246. Engstrom goes so far as to claim that this lies at the heart of the distinction between the understanding and receptivity (2009: 29): “Kant bases this distinction in the contrast between spontaneity and receptivity, or between a cognizing subject’s capacity to be self-consciously active and its capacity to be aware of the affections it undergoes”; “Ultimately, this distinction lies in a distinction between two forms of consciousness: consciousness that, being aware of itself as arising from itself, necessarily recognizes itself as the source of its object and thereby understands the latter as its product, and consciousness the subject recognizes as arising from a source outside itself and so regards as an affection of itself by something acting upon it” (Engstrom 2009: 29n7). [↑](#footnote-ref-246)
247. IV: 450. 18 – 453. 15. For excellent treatment of the circle, see Quarfood (2006). [↑](#footnote-ref-247)
248. In this, I agree with Allison’s (2011: 312-6) account of the circle. I also follow him to think that, the real worry here is strictly a *petitio* – assuming something that has yet to be proved – rather than a circular argument (Allison 2011: 315); cf. IV: 453. 3-15. [↑](#footnote-ref-248)
249. He expands on this: “[…] and this natural necessity is no experiential concept either, precisely because it carries with it the concept of necessity, and hence of a priori cognition. But this concept of a nature is confirmed by experience and unavoidably must itself be presupposed if experience – i.e. cognition of objects of the sense cohering according to universal laws – is to be possible.” (IV: 455. 18-24). He later adds the following remark: “That is why freedom is only an idea of reason, the objective reality of which is in itself doubtful, whereas nature is a concept of the understanding that proves, and must necessarily prove, its reality in examples of experience” (IV: 461. 24-27). [↑](#footnote-ref-249)
250. Cf. Bxxix: “[…] that speculative reason had proved that freedom cannot be thought at all, then that presupposition, namely the moral one, would necessarily have to yield to the other one, whose opposite contains an obvious contradiction; consequently **freedom** and with it morality (for the latter would contain no contradiction if freedom were not already presupposed) would have to give way to the **mechanism of nature**. But then, since for morality I need nothing more than that freedom should not contradict itself, that it should at least be thinkable that it should place no hindrance in the way of the **mechanism of nature** in the same action (taken in another relation), without it being necessary for me to have any further insight into it.” Cf. Hill (1998: 265) and Tenenbaum (2012: 559-60). [↑](#footnote-ref-250)
251. As noted in the introduction, Gregor describes this as “something that belongs to no one” (1996: 103n), Timmermann as “a good that has no apparent owner” (2007: 146). [↑](#footnote-ref-251)
252. Cf. The third antinomy (A 444-51/B 472-9). [↑](#footnote-ref-252)
253. I want to circumvent the interpretative debate about how to best understand transcendental idealism, as I think that the problems I raise affect any account of transcendental idealism. I try to keep my discussion at a general level, to reflect this (although in the next section, I will consider a possible response from a ‘two-standpoint’ view). [↑](#footnote-ref-253)
254. Cf. Ameriks (2003: 227-8). [↑](#footnote-ref-254)
255. Ameriks (2003: 216) also makes this charge against Allison, who he worries comes close to attributing a kind of intellectual intuition to Kant, when he speaks of “the consciousness of our rational agency […where] we are directly aware of a capacity (to act on the basis of an ought [which includes epistemic and prudential oughts]) that, […] we cannot regard as empirically conditioned” (Allison 1990: 44-5). [↑](#footnote-ref-255)
256. Cf. IV: 455. 13-4; IV: 459. 5-7. [↑](#footnote-ref-256)
257. In the *Groundwork*, Kant does complain about the “common understanding” wanting to make the world of understanding into “an object of intuition” (IV: 454. 5), but at the same time, as we have seen, he makes a lot of the consciousness of the causality of our reason. Hence, Amerik’s claim that the arguments of *Groundwork* III are merely (2000: 214): “a slightly more self-conscious continuation of some relatively crude beliefs about freedom that Kant had held for some time and simply had not gotten around to submitting to a thorough critique.” [↑](#footnote-ref-257)
258. Cf. Bxxviii: “I cannot cognize freedom as a property of any being to which I ascribe effects in the world of sense, because then I would have to cognize such an existence as determined, and yet not as determined in time (which is impossible, since I cannot support myself with any intuition)”; “It is […] absolutely impossible to give anywhere in experience an example of it, since among the causes of thins as appearances no determination of causality that would be absolutely unconditioned can be found” (V: 48. 23-7). [↑](#footnote-ref-258)
259. Cf. Grenberg (2013: 5). [↑](#footnote-ref-259)
260. Cf. Korsgaard (1996a: xii): “Kant approaches moral philosophy in a very different way than the British Empiricists and their heirs in the analytic tradition do. The basic problem, set by the plight of rational agency, is “what should I do?” The approach is to raise practical questions as they are faced by the reflective moral agent herself. Moral philosophy is the extension and refinement of ordinary practical deliberation, the search for practical reasons. This makes Kant’s enterprise very different from that of philosophers who talk about morality and the moral agent from the outside, third-personally, as phenomena that are in need of explanation. Kant arguments are not *about* us; they are addressed *to* us.” Pippin puts forward a similar line (2009: 38): “Normative questions, I mean, are irreducibly “first-personal” questions, and these questions are practically unavoidable and necessarily linked to the social practice of giving and demanding reasons for what we do, especially when something someone does affects, changes, or limits what another would otherwise have been able to do. By irreducibly first-personal, I mean that whatever may be our “snap judgments” or immediate deeply intuitive reactions, whenever anyone faces a normative question (which is the stance from which normative issues are issues), no third-personal fact–why one as a matter of fact has come to prefer this or that, for example–can be relevant to what I must decide, unless I count it as a relevant *practical* reason in the justification of what I decide ought to be done or believed.” [↑](#footnote-ref-260)
261. Sticker (2014: §4.2.1) draws our attention to the Sulzer footnote in the *Groundwork* (IV: 411n), as well as the fact of reason itself. Grenberg points out that, in the second gallows case, Kant claims that an agent becomes aware of the moral law introspectively and “immediately” (V. 29. 30-5). Sticker notes that Kant does not present the case directly to his reader, but rather suggests that the reader present the case to a third party – “But ask him […]” (V: 29. 27) – and learn something from this third party’s verdict. Kant stresses that an agent when presented with the gallows must “admit” (V: 30. 33) that they can do something because they ought to. [↑](#footnote-ref-261)
262. Darwall (2006: 124) argues that, even if Kant’s first-person approach works, it fails to vindicate the distinctive normativity of moral obligations, which he sees as essentially involving a second-personal element. In this paper, I offer an analogous argument concerning the third-person. [↑](#footnote-ref-262)
263. This famously finds its expression in Lewis White Beck’s two-aspect view of human conduct (1960: 29-32; 1975). [↑](#footnote-ref-263)
264. Fichte is aware of this problem – see §3 of the *Foundation of Natural Rights*; cf. Beiser (2002: 334-7). [↑](#footnote-ref-264)
265. A key passage occurs at A 549-50/B 577-8. [↑](#footnote-ref-265)
266. Cf. A 359; Allison (1990: 30-4); Van Kirk (1986: 52), [↑](#footnote-ref-266)
267. Cf. V: 468. 24-30. [↑](#footnote-ref-267)
268. Bennett (1984: 105) picks up on a similar worry. [↑](#footnote-ref-268)
269. For another important objection to the ‘two-standpoint’ reading of Kant’s theory of freedom, see Nelkin (2000). [↑](#footnote-ref-269)
270. Cf. Schopenhauer (1819: §19). [↑](#footnote-ref-270)
271. Again, cf. Bxxviii. [↑](#footnote-ref-271)
272. Cf. Korsgaard (1996a: 162-3; 1996b: 94-7). [↑](#footnote-ref-272)
273. Cf. Grenberg (2013: 129; 132). [↑](#footnote-ref-273)
274. Grenberg herself notes that (2013: 129): “There is a certain gleeful indeterminacy to her agency.” I agree, but this is a third-personal judgement, and it is hard to see how she is entitled to it, given her exclusively first-personal approach. [↑](#footnote-ref-274)
275. Cf. Brown (1998: 90; 94); Wood (1990: 80). [↑](#footnote-ref-275)
276. For an excellent treatment of this problem, see Freyenhagen (2008). [↑](#footnote-ref-276)
277. On the malicious lie, Bennett notes that (1984: 106): “[…] Kant’s theory allows us to pass that judgment not only in this case but also in one where the natural causes of the lie involve a profound psychopathology in someone who is not vicious and is greatly given to shame.” [↑](#footnote-ref-277)
278. Cf. Bennett (1984: 106). [↑](#footnote-ref-278)
279. On this, Wood follows Hegel to complain that (1990: 153): “When we displace morality to a beyond, locating it on the struggle of motives in an intelligible world, we alienate our practical reason from our self-satisfaction. When we locate our moral worth in a noumenal self that is hidden from us, we devalue the only self-worth we can actually possess.” [↑](#footnote-ref-279)
280. Blӧser (2015) offers a two-level theory of imputation, which she contends allows Kant to make sense of degrees of responsibility (the first level), but not degrees of accountability (the second level). This might entitle Kant to some account of praise and blame, but his set-up still precludes a workable count of accountability and interaction. [↑](#footnote-ref-280)
281. Cf. V: 94. 2-7. Callanan (2014) offers an intriguing account of the importance of this claim. [↑](#footnote-ref-281)
282. See, for example, his revealing discussion of a tree at V: 371. 7 – 372. 11. [↑](#footnote-ref-282)
283. Cf. V: 374. 21-6 “An organized being is thus not a mere machine, for that has only a **motive** power, while the organized being possesses in itself a **formative** power, and indeed one that it communicates to the matter, which does not have it (it organizes the latter): thus it has a self-propagating formative power, which cannot be explained through he capacity for movement alone (that is, mechanism).”; V: 376. 13-4; V: 377. 1-3; V: 382. 30-4. [↑](#footnote-ref-283)
284. Cf. V: 360.21 – 361. 11; V: 365. 17-25; V: 379. 10-12; V: 383. 12-4. [↑](#footnote-ref-284)
285. Cf. V: 404. 30-6: “[…] the concept of the purposiveness of nature in its products is a concept that is necessary for the human power of judgment in regard to nature but does not pertain to the determination of the objects themselves, thus a subjective principle of reason for the power of judgment which, as regulative (not constitutive), is just as necessarily valid for our **human** **power of judgment** as if it were an objective principle.” [↑](#footnote-ref-285)
286. Cf. Schelling (1803: SWII: 42): “Certainly there are philosophers who have *one* universal answer to all these questions, which they repeat at every opportunity and cannot repeat enough. That which is form in the things, they say, we initially impose on the things. But I have long sought to know just how you could be acquainted with what the things are, without the form which you first impose on them, or what the form is, without the things on which you impose it. You would have to concede that, *here* at least, the form is absolutely inseparable from the matter, and the concept from the object. Or, if it rests with your choice whether or not to impose the idea of purposiveness on things outside you, how does it come about that you impose this idea only on *certain* things, and not on *all*, that further, in this representing of purposeful products, you feel yourself in no way *free*, but absolutely constrained? You could give no other reason for either than that this purposive form just belongs to certain *things* outside you, originally and without assistance from your choice.” [↑](#footnote-ref-286)
287. Cf. McDowell (1994: 83-6; 2000: 99). [↑](#footnote-ref-287)
288. Cf. McDowell (2006: 238): “I am quite happy to suppose there are two kinds of happenings in nature: those that are subsumable under natural law, and those that are not subsumable under natural law, because freedom is operative in them.” [↑](#footnote-ref-288)
289. It is worth noting that McDowell does not think that this is a problem for his approach (1994: 178): “If we take ourselves to be addressing [… the question “what constitutes the structure of the space of reasons?”], my invocation of second nature, sketchy and unsystematic as it is, will seem at best a promissory note towards a proper response. But that would miss my point. I think the response we should aim at being entitled to, if someone raises a question like [that….] is something like a shrug of the shoulders […] Their sheer traditional status cannot by itself oblige us to take such questions seriously. Rather, there is an assumed background that is supposed to make them urgent. When I invoke second nature, that is meant to dislodge the background that makes such questions look pressing, the dualism of reason and nature. It is not meant to be a move – which could be at best a first move – in constructing a response to that question.” [↑](#footnote-ref-289)
290. Over the course of this thesis, I have departed from Allison’s non-metaphysical reading of premise 1 (ch.1, §3), and his constructivist reading of premise 2 (ch.2, §1.5). [↑](#footnote-ref-290)
291. Allison (1990: 227-8; 1996: 109-14; 129-42; 2012: 115-6; 2013, 136-7). [↑](#footnote-ref-291)
292. Grenberg (2009: 348-52) and Reath (1993: 424-5) follow Allison fairly closely here. For other accounts of this objection, see: Ameriks (2000: 205; 2003: 173, 241-4), Darwall (2006: 216-7), Mieth and Rosenthal (2006: 281-2), and Tenenbaum (2012: 569-70). [↑](#footnote-ref-292)
293. Cf. Allison (1990 : 35-42; 1996: 130-4; 2011 : 114-20; 2012: 144-5). [↑](#footnote-ref-293)
294. Cf. Allison 1996: 109-14; 202n7. For an extended attempt to overcome the issue at hand in terms of practical and transcendental freedom, see Engstrom (1993) and Saunders (2014: 131-4). [↑](#footnote-ref-294)
295. Cf. Allison (1996: 111): “[…] an ineluctably heteronomous agent would require a sensuous inclination in order to have a *sufficient* reason to adopt an end (or to act at all)”; Allison (1996: 137): “[…] causal independence […] does not entail autonomy in the full Kantian sense […] the latter involves more than independence from determination by alien causes; it also involves a positive capacity to be motivated by reason that are totally independent of one’s needs as a sensuous being.” [↑](#footnote-ref-295)
296. Allison (1996: 205n12) acknowledges this similarity. [↑](#footnote-ref-296)
297. In a famous passage in the *Religion*, Kant himself appears to affirm this possibility. He distinguishes between humanity and personality, and claims (in a footnote) that (VI: 26n): “[…] from the fact that a being has reason does not at all follow that, simply by virtue of representing its maxims as suited to universal legislation, this reason contains a faculty of determining the power of choice unconditionally, and hence to be "practical" on its own; at least, not so far as we can see.” [↑](#footnote-ref-297)
298. Engstrom makes the case that Kant conceives of the faculty of desire *[Begehrungsvermögen]* in a broad sense (2009: 27): “Kant emphasizes that the faculty of desire should not be defined in a way that would build into its very concept the supposition that a feeling of pleasure must lie at the basis of its determination, so that desiring something always depends in the end on this capacity’s being or having been affected by some pleasing experience. To do so would be in effect to rule out from the start, by definitional fiat, the possibility of conceiving of the will as practical reason rather than merely as rational desire. It would be to build into the very concept of desire the Humean supposition that desire cannot arise from reason, even though it may be directed by it. A thoroughgoing empiricism would be introduced, by stipulation, into practical philosophy. Instead, the definition should be confined to what can be understood just from the general concept of desire.” (As we saw in chapter 2 (§1.5) however, Engstrom himself seems to overlook this in his critique of moral realism.) Cf. Hill (1998: 260-2). [↑](#footnote-ref-298)
299. Cf. The *Nicomachean Ethics*, bk. III, ch.3. [↑](#footnote-ref-299)
300. Once more, cf. Allison (2012: 118). [↑](#footnote-ref-300)
301. Cf. Timmermann (2006: 91n37). Timmermann thinks that the feeling of respect [*Achtung*] plays a unique role, in that it is a feeling (and thus can motivate us), but one caused by recognition of the moral law; see, for example, Timmermann (2007: 182). In resisting the thought that we need pre-existing desires to motivate us, I do not need the feeling of respect to play this role. [↑](#footnote-ref-301)
302. Cf. Redding (2007: 1-20). [↑](#footnote-ref-302)