

■ An Introduction to  
Kant's Moral  
Philosophy

Jennifer K. Uleman



CAMBRIDGE

CAMBRIDGE

[www.cambridge.org/9780521199629](http://www.cambridge.org/9780521199629)

This page intentionally left blank

# AN INTRODUCTION TO KANT'S MORAL PHILOSOPHY

Immanuel Kant's moral philosophy is one of the most distinctive achievements of the European Enlightenment. At its heart lies what Kant called the "strange thing": the free rational human will. This introduction explores the basis of Kant's anti naturalist, secular, moral vision. Moving from a sketch of the Kantian will, with all its component parts and attributes, to Kant's canonical arguments for his categorical imperative, it shows why Kant thought his moral law the best summary expression of both his own philosophical work on morality and his readers' deepest shared convictions about the good. Kant's central tenets, key arguments, and core values are presented in an accessible and engaging way, making this book ideal for anyone eager to explore the fundamentals of Kant's moral philosophy.

JENNIFER K. ULEMAN is Assistant Professor of Philosophy, Purchase College (State University of New York). She is the author of numerous articles and reviews.



AN INTRODUCTION  
TO KANT'S MORAL  
PHILOSOPHY

JENNIFER K. ULEMAN



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,  
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press  
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521199629](http://www.cambridge.org/9780521199629)

© Jennifer K. Uleman 2010

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2010

ISBN 13 978 0 511 67664 2 eBook (NetLibrary)

ISBN 13 978 0 521 19962 9 Hardback

ISBN 13 978 0 521 13644 0 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

*For Ethan James and Sahara Rose,  
with all the love in the world.*





# Contents

<i>Acknowledgements</i>	<i>page</i>	<i>ix</i>
1 Introduction: the strange thing		I
The strange thing		I
The free rational will		9
The value of free rational will		15
The importance of the strange thing for moral philosophy		19
2 A sketch of Kantian will: desire and the human subject		23
Desire, choice, will		25
Desire and the human subject		34
3 A sketch continued: the structure of practical reason		39
Will as practical reason: practical rules, laws, and principles		39
Maxims, or subjective practical principles		41
Grounds for action: the representation in a principle of something as good		48
Imperatives		51
Pure practical reason, or the possibility of a categorical imperative		56
4 A sketch completed: freedom		63
An overview of the free Kantian will		63
The free Kantian will in more detail		65
Rational freedom		72
5 Against nature: Kant's argumentative strategy		75
The problem		75
Kant's understanding of nature		79
Kant's common-sense case against a natural foundation for morality		80
Kant against nature		107

6	The categorical imperative: free will willing itself	111
	Kant's formalism	112
	Kant's categorical imperative: its form and its content	121
	Free will willing itself	140
7	What's so good about the good Kantian will? The appeals of the strange thing	145
	Introduction	145
	The good of free rational willing	149
8	Conclusion: Kant and the goodness of the good will	175
	<i>Bibliography</i>	180
	<i>Works by Kant</i>	180
	<i>Newer titles of note</i>	181
	<i>Works cited</i>	181
	<i>Index</i>	187

## *Acknowledgements*

This book has been with me a long time, and I am grateful to many who helped, in many different ways. Thanks to Rachel Zuckert, Yasmin Yildiz, Sophia Wong, Elizabeth Wells, Jami Weinstein, Helga Varden, Corina Vaida, Kate Uleman, Kara Uleman, Jim Uleman, Lucas Thorpe, Amie Thomasson, Aviva Taubenfeld, Michele Sviridoff, Karen Struening, Michelle Stewart, Matt Statler, Maggie Smith, Michael Slotte, Alison Simmons, Harvey Siegel, Sally Sedgwick, Carolina Sanin, Michael Rothberg, Marilyn Rifkin, Giovanna Pompele, Thomas Pogge, Cara Perlman, Peter Ohlin, Darrell Moore, Marjorie Miller, Eric Mendenhall, Elias Markolefas, Peter Lewis, Kostis Kourelis, Morris Kaplan, Andrew Janiak, Risto Hilpinen, Sally Haslanger, Casey Haskins, Sean Greenberg, Lanny Goldman, Rudi Gaudio, Tracy Fitzpatrick, Frank Farrell, Simon Evnine, Jo Dexter, Marilyn Cleveland, Taylor Carman, Stephanie Camp, Marc Brudzinski, John Beusterien, and Tessa Addison. Thanks also to those I have failed to thank by name, but who have nonetheless helped me think, talk, and write about Kant.

I am grateful for the material support this project has received. Thanks to several Max Orovitz research grants from the University of Miami, to Paolo Tuzzi for the house in Cossogno that summer, and to the Learning Center at Purchase College for help funding my brilliant undergraduate research assistant, Felix Pichardo.

A student's thanks to Paul Guyer for being an outstanding teacher and mentor. He will surely disagree with parts of this book, but it is also thanks to his teaching that I wrote it. A writer's thanks to Hilary Gaskin and two anonymous reviewers at Cambridge University Press for all their care and dispatch. Finally, a teacher's very special thanks to all my Kant students, whose willingness to talk, listen, read, and argue has, more than anything else, helped me figure out what I wanted this book to say.



## CHAPTER I

### *Introduction: the strange thing*

#### THE STRANGE THING

“The thing is strange enough and has no parallel in the remainder of practical knowledge” (KpV 5:31). So writes Kant about the activity of human will. According to Kant, human will authors an ultimate action-guiding principle – a moral law – that tells what matters most and how to act accordingly. It binds itself to this law, experiencing the law’s commands as absolute and expecting as reward neither happiness nor heaven, eschewing both sensuous and divine incentives. According to Kant, human will understands the moral law it has authored as holding not only for itself but universally. The strange activity of this strange thing is strange for many reasons. It is free in a determined world; it subjects itself to itself, despite the seeming paradox of this; in the end, and strangest of all, the will that authors and can bind itself to moral law is *itself* what matters most, is itself the aim of morality. The strange will is thus its own object: at the heart of Kant’s moral theory is, to use Hegel’s words, “the free will which wills the free will.”<sup>1</sup> The moral law that Kantian free will authors is, to put it another way, strangely and ingeniously self-serving. This book is about all these strange things, and especially about why, for Kant, the strange, free, law-giving will is its own ultimate aim.

This book is about these things in order to offer an introduction to, as well as an interpretation of, Kant’s moral theory. It therefore surveys the foundations of Kant’s moral thought, laying out basics and making clear what Kant values, why he values it, and why he thought his famous “categorical

<sup>1</sup> Or in Hegel’s German, “*der freie Wille, der den freien Willen will.*” Hegel is here describing ‘the abstract concept of the idea of will in general,’ and though he does not name Kant in the passage, Hegel makes clear elsewhere that he admires Kant for identifying and attending to the will so conceived, even though he thinks Kant’s final moral theory comes up short (G. W. F. Hegel, *Grundlinien der Philosophie des Rechts* [Werke 7] [1821] [Frankfurt: Suhrkamp Verlag, 1970], §27; translation: *Elements of the Philosophy of Right*, trans. H. B. Nisbet, ed. Allen Wood [Cambridge: Cambridge University Press, 1991], §27, p. 57).

imperative”<sup>2</sup> the best summary expression of both his philosophical work on morality and of his readers’ deepest shared convictions about morality and value. It aims to show that Kant’s moral theory is driven by respect and awe for the specifically human capacity to act in the world in ways that are, as Kant understood the terms, free and rational. It aims to show that the core of Kantian moral theory is indeed the free rational will that wills itself. And it aims to show what a theory driven by respect and awe for free rational will asks of us – what forms of life, what micro-commitments, what conceptions of self, what collective arrangements it requires us to embrace, and what it requires us to reject.

For Kant, the complex that is ‘free rational willing’ or ‘free rational activity of the will’ cannot really be taken apart and still make sense. Each of its terms – ‘free,’ ‘rational,’ and ‘will’ – is ultimately defined in ways that implicate and depend on each other. Intuitions and ordinary usage thus cannot provide the guidance we need to understand the thing I claim is at the heart of Kant’s moral theory, since intuitions and ordinary usage would let us separate these three. Indeed, intuitions and ordinary usage would sometimes oppose them.

Is there a term that names the Kantian complex? In early work on this book, I found myself using ‘autonomy’ to describe free rational willing. Autonomous activity *is* more or less the same as free rational activity of the will; ‘autonomy’ is characteristic of a will that (freely) gives itself a (rational) action-guiding law. ‘Autonomy’ thus has the advantage that it encompasses and inextricably relates, in one word, Kantian freedom, rationality, and will. But I have decided not to use the term here, at least not very often, despite its being, in some contexts, a key term for Kant himself. Not unlike ‘freedom,’ ‘rationality,’ and ‘will,’ the term ‘autonomy’ is so freighted, its accreted connotations so thick, its post-Kantian adventures so various and storied, that I prefer less felicitous terms and phrases, like ‘free rational practical activity,’ ‘free rational willing,’ and ‘free rational activity of the will.’ Besides triggering fewer associations for readers, these also have the advantage, when they come as phrases, of reminding us just what Kant *is* seeking to encompass and inextricably relate. They may thus be worthwhile in helping to keep Kant’s conceptions strange and interesting.

<sup>2</sup> Kant’s ‘categorical imperative’ is formulated in several ways. The most familiar are these: (1) act only on maxims that you can at the same time will as universal laws, and (2) treat others never merely as means but always also as ends in themselves. See G 4:421 and 429. Much more will be said about Kant’s categorical imperative in subsequent chapters.

The interpretation of Kant's moral theory that I offer cuts against the grain of interpretations that emphasize Kant's commitments to formal rules and rationalism. Such interpretations have deservedly influential proponents<sup>3</sup> and, despite recent scholarship that pulls toward feeling and content,<sup>4</sup> formalist, rationalist views of Kant still circulate widely in lecture halls and college corridors and in the collective intellectual imagination more generally.<sup>5</sup> Not without reason: Kant *was* deeply committed to a kind of formalism, and was deeply committed to rationality. But if these commitments are overemphasized, or emphasized in the wrong ways, we are left with a view that is less engaging and more academic than Kant's. Kant's Aristotelian, Humean, Hegelian, and other foes have taken note: overly formalist, rationalist interpretations have the capacity to drain the life out of Kant's views, and accordingly have been offered as often by Kant's enemies as by his friends.<sup>6</sup>

<sup>3</sup> Important work by Christine Korsgaard, Onora O'Neill, and John Rawls pulls Kant in what I think of as formalist, rationalist directions. See Christine Korsgaard, *The Sources of Normativity*, with G.A. Cohen, Raymond Geuss, Thomas Nagel, and Bernard Williams, ed. Onora O'Neill (Cambridge and New York: Cambridge University Press, 1996) and the essays collected in Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge and New York: Cambridge University Press, 1996), esp. "Kant's Formula of Universal Law," 77–105. See Onora O'Neill, *Constructions of Reason: Explorations of Kant's Practical Philosophy* (Cambridge and New York: Cambridge University Press, 1989), esp. "Consistency in Action," 81–104. See John Rawls, "Themes in Kant's Moral Philosophy," in *Kant's Transcendental Deductions*, ed. Eckart Förster (Stanford: Stanford University Press, 1989), 81–113.

<sup>4</sup> I am thinking especially of work by Paul Guyer, Barbara Herman, and Allen Wood. See the essays in Paul Guyer, *Kant and the Experience of Freedom* (Cambridge and New York: Cambridge University Press, 1993), esp. "Duty and Inclination," 335–93, and the essays in Paul Guyer, *Kant on Freedom, Law, and Happiness* (Cambridge and New York: Cambridge University Press, 2000), esp. "Freedom As the Inner Value of the World," 96–125, and "Kant's Morality of Law and Morality of Freedom," 129–71; the essays in Barbara Herman, *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), esp. "Leaving Deontology Behind," 208–40; and Allen Wood, *Kant's Ethical Thought* (Cambridge and New York: Cambridge University Press, 1999). Marcia Baron's *Kantian Ethics (Almost) without Apology* (Ithaca: Cornell University Press, 1995) may also be seen to pull in this direction.

<sup>5</sup> A representative textbook account of Kant as invested in reason and form can be found in Samuel Enoch Stumpf and James Fieser, *Socrates to Sartre and Beyond: A History of Philosophy*, 8th edn. (New York: McGraw-Hill, 2008). The account, which is admirably clear and fair, ends by observing that Kant's categorical imperative "speaks of the universality of the moral law, affirms the supreme worth of each rational person, and assigns freedom or autonomy to the will," but does not try to explain how these are connected to each other, or why any of them is morally attractive (Stumpf and Fieser, *Socrates to Sartre*, p. 289). A textbook account that is congenial to the view I am advocating can be found in James and Stuart Rachels' widely used *The Elements of Moral Philosophy*, 5th edn. (New York: McGraw-Hill, 2006), 130–40.

<sup>6</sup> Annette Baier and Bernard Williams both portray Kant in a rational formalist light, and critique him for over-reliance on reason and formal procedures. See Annette Baier, *Moral Prejudices: Essays on Ethics* (Cambridge, Mass.: Harvard University Press, 1994); and Bernard Williams, *Moral Luck: Philosophical Papers 1973–1980* (Cambridge and New York: Cambridge University Press, 1981) and *Ethics and the Limits of Philosophy* (Cambridge, Mass.: Harvard University Press, 1985).

What do I mean by ‘formalism,’ and how can it be overemphasized? Kant does insist that diverse human aims and ends, to be morally acceptable, must conform to the ‘form of universalizability’ – that is, that they may be pursued only if they could be universally endorsed. Described thus, Kantian morality does not advance *any particular* aim or end – it just insists that we pursue our aims and ends only if they pass a formal test. Because it does not dictate particular aims or ends, Kantian morality seems able to accommodate good human lives lived across circumstance and historical time and place: it seems pluralistic and inclusive. Because it insists on universal acceptability, it seems to respect the value of hearing from everyone, or at least of imaginatively trying to, by putting yourself in other people’s shoes (would it be OK with everyone?): it seems deeply democratic. Kant’s view thus comes across as a natural predecessor to the sort of contemporary procedural liberalism advocated by thinkers like John Rawls, which claims a strong commitment to neutrality between competing conceptions of the good.<sup>7</sup>

However, as Rawls himself knew, and as I argue here, Kant is not so neutral, and demands much more than accord between ends and aims and a certain form: he demands that we embrace, as intrinsically and ultimately good, the free rational human will itself.<sup>8</sup> Embracing the free rational will as good means organizing our individual and collective lives in ways that actively honor this good. As a consequence, Kantian morality rejects moral projects the ultimate object of which is to serve God, or to alleviate material suffering; these projects, for Kant, unacceptably subjugate free rational will to other ends. As we will see, Kantian morality also rejects projects, such as Nietzsche’s, of radical self-invention, congenial as self-invention may sound to a project that values free will (especially construed as ‘autonomy’). The free will Kant values is one that is fundamentally legible to others, and committed to a radically shared rationality. In emphasizing the aim or end – free rational willing – that Kant requires us to embrace, my presentation seeks to draw out the specific shape of the moral life Kant demands we live.

<sup>7</sup> Rawls’ basic thought is that human beings can arrive at ground rules (‘principles of justice’) capable of fairly governing social and political institutions without prejudice between particular cultural, religious, or other conceptions of the good (John Rawls, *A Theory of Justice* [Cambridge, Mass.: Harvard University Press, 1971], e.g., 11–16 and 446–52). Many see Rawls as ‘softening’ this view in his later work, via acknowledgement that his principles of justice *are* hostile to some historical and contemporary ways of life (whether ‘traditional,’ religious, or strongly communitarian) (John Rawls, *Political Liberalism* [New York: Columbia University Press, 1996] esp. xv–xxxii, 174–6, and 243–4).

<sup>8</sup> See John Rawls, “Justice as Fairness: Political, Not Metaphysical,” *Philosophy and Public Affairs* 14:3 (1985), 223–51. In this essay, Rawls distances his own ‘Kantian’ view from the ‘metaphysical’ commitments about the nature of the soul and its vocation that he finds in Kant.



How does my presentation cut against overly ‘rationalist’ interpretations? Such interpretations emphasize the moral law’s origins in Kantian pure reason. Now, the moral law *does* have such origins for Kant. And such origins *do*, as Kant intended, ensure that morality is not subject to local emotional whims or physical exigencies; such origins also ensure that the moral law holds universally for ‘all rational creatures’ (including all human beings).<sup>9</sup> But because rationalist interpretations too often employ (or at least let stand) a soulless, calculative conception of reason, they can fail to make palpable reason’s own strong commitments, including its interests in and reverential respect for its own strivings. If the very idea of reason *having* commitments and interests of its own seems strange, that is because we have become accustomed to thinking of reason in precisely this soulless, calculative way. But we need not – think, for instance, of the commitments to and interests in things like accuracy, consistency, scope, simplicity, and fruitfulness that Thomas Kuhn has persuasively shown us are part and parcel of rational scientific investigation.<sup>10</sup> These are interests internal to reason itself. The interests just named are internal to, in Kant’s language, ‘theoretical’ – or theory-building – reason. But there are also interests internal to Kantian ‘practical’ – or action-guiding – reason, chief among them an interest in free rational striving itself (or so I will argue here). Overly rationalist interpretations thus also often do violence to the emotional investments Kant thinks we have in freedom, rationality, and willing. ‘Awe’ and ‘respect’ are just two of the key terms Kant frequently uses to describe our reactions to free rational willing, whether our own or other people’s. In emphasizing the interest and reverential respect that animates Kant’s reason-grounded morality, I break with the calculative, emotionally neutral spirit of much ‘rationalist’ interpretation of Kant.

I said above that some of Kant’s foes have favored formalist and rationalist interpretations. And indeed, formalism and rationalism, especially taken together, can be harnessed to cast Kant in a very unappealing light. The reader has perhaps been introduced to this Kant: he cares more about rules than about ends, he is wedded to impersonal calculation, he is unwilling to acknowledge his own particularity, he eschews all feeling, even (if not

<sup>9</sup> Barring only those who for some reason lack the capacities that constitute reason – for example, infants, young children, and those with severe mental impairments. Lest this seem to exclude too many, notice that, for Kant, even “children of moderate age,” who presumably lack fully developed reason, nonetheless have sound moral intuitions and feel respect for duty (G 4:411n; see also KpV 6:155–7).

<sup>10</sup> Thomas Kuhn, “Objectivity, Value Judgment, and Theory Choice,” in his *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press, 1977) 320–39.

especially) kind and warm feelings.<sup>11</sup> All the while, he insists on ‘duty’ for its own (incomprehensible) sake, and he generally comes off as a very cold fish. A significant part of my aim here is to move away from interpretations that focus on this Kant and emphasize instead the ends, interests, and feelings (such as reverential respect), that drive Kant’s theory. The effect, I hope, will be to make his view less vulnerable to charges of motivational opacity, of emotional coldness, and of a stultifying obsession with rules. If the aim of moral thought and action is free rational activity itself, if we understand ourselves as deeply committed to this activity, and if rules are just the best way to express strategies for honoring this activity, Kant’s view seems sturdier. It gains a kind of lived-life plausibility, feels more ennobling and less punishing, less rigid and more productive than critics have often charged. Moreover, when the interests, feelings, and ends internal to Kant’s project are made more apparent, the temptation to read Kant as imperiously and hubristically insisting that ‘all rational creatures’ must embrace his view (on pain of being banished from the realm of the rational) is diminished. If we read Kant’s as a view that must court our allegiance by identifying interests, feelings, and ends we share, the view becomes both more interesting and more satisfying to entertain. Charges of false and condescending universalism must give way to argument about the substance of Kant’s view, and the value of the interests, feelings, and ends he identifies.

The price of inviting argument about the value of Kantian interests, feelings, and ends is, of course, that argument will be offered. Kantian

<sup>11</sup> As in Friedrich Schiller’s satirical verse, meant to mock Kant’s commitment to duty over feeling:

Gewissensskrupel  
 Gerne dien’ ich den Freunden, doch tu’ ich es leider mit Neigung.  
 Und so wurmt es mir oft, daß ich nicht tugendhaft bin.  
 Decisum  
 Da ist kein anderer Rat, du mußt suchen, sie zu verachten  
 Und mit Abscheu alsdann tun, wie die Pflicht dir gebet  
 [Scruple of Conscience]  
 Gladly I serve my friends, but alas I do it with pleasure.  
 Hence I am plagued with doubt that I am not a virtuous person.  
 [Ruling]  
 Surely, your only resource is to try to despise them entirely,  
 And then with aversion to do what your duty enjoins you.

From Friedrich Schiller, *Xenien* [1797], collected in Goethe, *Werke I*, ed. Erich Trunz (Hamburg: Christian Wegner Verlag, 1949), 221. This translation (apart from headings, which I’ve added) appears in H. J. Paton, *The Categorical Imperative: A Study in Kant's Moral Philosophy* (London: Hutchinson & Co., 1947), 48; Paton notes, “the translation, which I take from Rashdall, *The Theory of Good and Evil*, Vol. I, p. 120, is by A. B. Bullock.” (Thanks to Anne Margaret Baxley for helping me track down these sources.)

morality *does*, as I have just suggested, reject some historically real human moral projects, and this opens it to direct challenge. Without the pretense that it is neutral between competing conceptions of the good, its own conception of the good emerges more clearly and becomes more vulnerable to attack. Kant has to show what is wrong with following God's will, or minimizing pain, at least as ultimate aims. Charges that Kantian views overstate human independence and self-sufficiency, or that they unjustly denigrate the body and nature, must also be met head-on. The task left for the Kantian is to defend the Kantian conception of what matters, not as incorporating or accommodating all other plausible conceptions of the good, but as in fact superior to them.

In fact, the ultimate aim of this book is to show both *that* and *why* Kant thought his conception of the good superior – to show that Kant thought the strange, moral-law-authoring, free human will more valuable than anything else, and to show why. In the remainder of this Introduction, I will say more about this strange thing, about the general thought that Kant's moral theory is set up to honor and revere it, and about my strategy in pursuing this thought.

The three short chapters that immediately follow this Introduction – “A sketch of the Kantian will: desire and the human subject” (Chapter 2), “A sketch continued: the structure of practical reason” (Chapter 3), and “A sketch completed: Freedom” (Chapter 4) – are intended to provide readers with a portrait of the Kantian will, which is, as must be clear, the book's central character. The process of sketching this portrait allows me to establish the claims that the Kantian will cannot act without an end, and to show how Kant understands both the will's rationality and its freedom. Chapter 5 (“Against nature: Kant's argumentative strategy”) argues that Kant's preference for formal principles issues not from an in-principle desire to deprive morality of a substantive end, but instead from Kant's low view of nature, and so also argues for the impossibility of meaningfully ‘naturalizing’ Kant's view. This chapter also introduces readers to the basic intuitions Kant thinks his readers share about morality, including the key thoughts that moral value lies in the quality of an agent's intention, and that moral praise is never merited by action undertaken on ‘ulterior motives’ – intuitions that will ultimately be satisfied by a moral theory based on the value of a certain kind of will. Chapter 6 (“The categorical imperative: free will willing itself”) makes the case for free rational willing as the ultimate Kantian value via a close reading of canonical texts (the *Groundwork of the Metaphysics of Morals* and the *Critique of Pure Reason*). I show how my reading makes sense of a host of Kantian moves, including the claim,

perplexing on its face, that three central formulations of the famous Kantian categorical imperative are 'at bottom' the same.

The book then turns away from arguments *that* Kantian moral theory aims at free rational will, and asks *why*, for Kant, free rational will is so valuable. What qualifies free Kantian will for the reverential respect Kant thinks it occasions? What makes it intrinsically good, an end in itself, indeed the ultimate end of moral thought and action? [Chapter 7](#), "What's so good about the good Kantian will? The appeals of the strange thing," takes these questions up, drawing on Kant's many scattered comments about the merits of free rational willing. The task of this chapter is delicate. The question of why Kant thought that free rational will is intrinsically and ultimately valuable cannot be addressed by showing what he might have thought it good for, since the claim is that it is *good in itself*, not for some other reason or purpose; comparing it to other goods is also not so helpful, since it is meant to be better than anything else. The task is thus one of unpacking or making explicit the 'goods' carried by something intrinsically and ultimately valuable. The goods to be discovered are intellectual, psychological, emotional, interpersonal, social, political, and arguably even physical and spiritual. To ask why free rational willing is valuable, for Kant, is to ask for more information about the package we adopt, in terms of self-conceptions and the hoped-for overall shape of our individual and collective lives, if we endorse free rational willing as the ultimate end of our own wills. For this reason, [Chapter 7](#) is offered as an account of the lived self-conceptions and experiences of Kantian subjects committed to and acting in accord with free rational will. I defend such an approach in more detail below, but the reasons for wanting an account of lived Kantian free rational willing should be clear: I want to make the fundamental motivations for Kant's view apparent, and saying he values free rational will, without saying more about what this means, about how this looks on the ground, and hence about why it might appeal, leaves too much unspoken. [Chapter 8](#), "Kant and the goodness of the good will," reviews the argument of the whole, revisiting the strangeness and the accomplishment that is a moral system based on the value of free rational will willing itself.

Before moving to a more extended overview of the basic terms of the project, a few remarks about things I will *not* do here, and a note about interpretation. First, I make no effort to survey the extensive and very excellent literature on Kant's moral theory, though I try to acknowledge debts and conscious disagreements when I can. That literature offers an embarrassment of riches, and my aim here is more introductory than comprehensive. Second, although they are very interesting, I do not worry

deeply here about the metaphysical issues – chiefly about freedom – raised by my account. They have been thoughtfully addressed by others;<sup>12</sup> some I try to address elsewhere;<sup>13</sup> some I would like to think about another time. Finally, this about my approach to interpreting Kant: interpretations are always interpretations, and the many complicated factors that make different people interpret the same text differently are well known to all who have taken hermeneutics seriously.<sup>14</sup> I was initially drawn to Kant because of an interest in the devotion so many people, including me, seem to have to the moral value of freedom per se. Kant has been a compelling interlocutor in trying to understand (at least one version of) this devotion. In arguing, implicitly and explicitly, that the interpretation of Kant's moral philosophy I offer is a good one, and is better than some others, I mean to argue that, in hard-to-quantify proportions, it does a good job accommodating and elucidating a range of central texts and delivers up a reading that is philosophically, morally, and psychologically plausible and powerful. This is, of course, what most interpreters try to argue; whether I succeed is for the reader to decide.

What follows here, as promised above, is a more extended overview of the basic terms of the project as it will unfold in the rest of the book.

#### THE FREE RATIONAL WILL

To begin, what makes a will *free*? A will is free, for Kant, if it determines itself and is not determined by anything else. A will is free, in other words, if it chooses ends, and pursues courses of action aimed at realizing those ends, on grounds that are its own, and not on grounds given to it by something or someone external to it. There are, of course, high philosophical stakes in any account of free will, and there are many theorists who would gloss 'free will'

<sup>12</sup> I am thinking especially of Henry Allison's work on Kant's theory of freedom, and of the many responses to Allison's work. See Henry Allison, *Kant's Transcendental Idealism: An Interpretation and Defense* (New Haven: Yale University Press, 1983/2004) and *Kant's Theory of Freedom* (Cambridge and New York: Cambridge University Press, 1990). For replies and comments, see Karl Ameriks, "Kant and Hegel on Freedom: Two New Interpretations," *Inquiry: An Interdisciplinary Journal of Philosophy* 35:2 (1992), 219–32.; Stephen Engstrom, "Allison on Rational Agency," *Inquiry: An Interdisciplinary Journal of Philosophy* 36:4 (1993), 405–18; Paul Guyer, "Review of Allison's *Kant's Theory of Freedom*," *The Journal of Philosophy* 89:2 (Feb. 1992), 99–110; and Andrews Reath, "Intelligible Character and the Reciprocity Thesis," *Inquiry: An Interdisciplinary Journal of Philosophy* 36:4 (1993), 419–29.

<sup>13</sup> Jennifer Uleman, "External Freedom in Kant's *Rechtslehre*: Political, Metaphysical," *Philosophy and Phenomenological Research* 68:3 (May 2004), 578–601.

<sup>14</sup> Like so many others, I have been guided in thinking about interpretation by Hans-Georg Gadamer, *Truth and Method* [1960], trans. and rev. Joel Weinsheimer and Donald G. Marshall (New York: Crossroad Publishing, 1989) and by Charles Taylor's essays, particularly those in *Human Agency and Language: Philosophical Papers*, 1 (Cambridge and New York: Cambridge University Press, 1985).

differently, even at this level of generality. Descartes, for instance, understands freedom of the will as a mental ability to endorse and set oneself on a course of action or not (or, more simply, to *assent* to something or not), rather than as a function of the grounds determining action (or assent). Importantly for Descartes, nothing can limit this ability (though our actual efforts may be thwarted): our freedom (as mental endorsement) is 'infinite.'<sup>15</sup> Hobbes, in stark contrast, denies that wills can be free at all, arguing that only *bodies* can be free or unfree, where 'free' just means 'unimpeded.'<sup>16</sup> For our present purposes, the important point is this: for Kant, unlike for some others, the idea of choosing on grounds that are our own is foundational to his account of freedom and the free will.

We can bring this foundation into sharper focus by noticing two challenges faced by Kant's conception of free will. The first challenge is posed by nature, as Kant calls the physical world; the second is posed by reason. The challenge posed by nature stems from the fact that we, who have wills, are embodied finite physical beings. As long as Newtonian laws of physics are operating – universal and necessary laws of mechanical cause and effect which, for Kant, govern everything in the natural world – it seems that everything we do must, if traced just a little way, have causal roots in forces and events outside ourselves. But if this is so, then our wills are merely conduits for external causes, particular kinds of locations in a causal net that stretches out infinitely in all directions. Under these conditions, we could not meet Kant's criterion for being free: we would not determine ourselves, but would be externally determined. The challenge posed by reason is analogous. Just as a will determined by external mechanical causes is not free, neither is a will determined by the demands of reason. To the extent, for instance, that I cannot reject a conclusion (of, say, a chain of mathematical reasoning), I am not free and my acceptance of the conclusion is not properly a choice. It must be *up to me* what I chose, in some ultimate sense – choices, if they are to be real choices, cannot be dictated by external rules or standards. This point was made often enough by scholastic and early modern voluntarists, who insisted that in order to be truly free, God's will (or 'power of volition') had to be free from answerability to reason or, for

<sup>15</sup> This is at least an important piece of Descartes' view. See, for example, René Descartes, *Meditations on First Philosophy with Selections from the Objections and Replies* [1641], trans. John Cottingham (Cambridge and New York: Cambridge University Press, 1986), 39–42 (AT 56–61). For a nice discussion of the complexities of Descartes' view, see Gary Hatfield, *Routledge Philosophy Guidebook to Descartes and the Meditations* (New York: Routledge, 2003), 183–202.

<sup>16</sup> Hobbes writes, "the Liberty of the man ... consisteth in this, that he finds no stop." Quotation and discussion both Thomas Hobbes, *Leviathan* [1651], ed. Richard Tuck (Cambridge and New York: Cambridge University Press, 1991), 146 (Ch. 21).

that matter, to any set of independent standards, whether of truth or goodness or beauty or whatever. If reason constrains me, forcing me to recognize six as the sum of three and three, I do not determine myself.

Faced with both these challenges, Kant's predecessor Locke concluded that human will is never free, but is always determined, by either nature or reason: for Locke, 'free will' is a nonsensical phrase. *Human beings*, Locke thought, could properly speaking be free or not, depending on whether they, as whole beings, were free to pursue what they willed.<sup>17</sup> But wills are always determined by something other than will itself, and so are never *themselves* free. But for a theorist like Kant, who wants to keep a meaningful sense of 'free will,' another route must be found. There must, for Kant, be a way to insist on the possibility of the will's *own self*-determination, that is, on will's being itself definitively pushed neither by laws of nature nor by laws of reason.

Kant's solution to the problem of determination by natural law is well known, if not widely accepted. Kant argues that there are two sorts of things, one of which is subject to natural causality, one of which is not. Spatio-temporal objects and events of the sorts available to the senses – *phenomena*, things that *appear* (φαίνω or *pheno*) – are subject to natural causal law; things that are not spatio-temporal or available to the senses, things like God and the immortal soul – *noumena*, which can only be grasped through the intellect or mind (νοῦς or *nous*) – are exempt from the natural causal order. Human wills, Kant argues, are, in significant if limited ways, non-spatio-temporal non-sensible things, and so are exempt (in significant if limited ways) from natural determination. Will thus joins God, the immortal soul, and possibly many other things, in 'the noumenal realm,' a realm [*Reich*] of objects and concepts related to each other not by mechanical cause and effect but by a different set of logical and conceptual relations and a different set of laws.<sup>18</sup> How to conceive the relationship between the noumenal and the phenomenal – a relationship there must be since our noumenal wills (like God's) produce results in the spatio-temporal sensory world – remains vexed. But what matters here is that Kant, in seeking to preserve will's freedom, insists on envisioning will as something other than a mechanism determined from without and instead as another sort of thing, belonging to an order outside nature, outside space and time, which can begin its own unconditioned beginnings.

<sup>17</sup> John Locke, *An Essay Concerning Human Understanding* [1689], ed. Peter H. Nidditch (Oxford: Oxford University Press, 1975), Book II, Ch. 21.

<sup>18</sup> At least some of which Kant will, not insignificantly, call 'laws of freedom.' See, for example, G 4:387, or KpV 5:65, or MS 6:214.

What about the challenge from reason? Reason threatens freedom by placing *intellectual* demands. This threat may be less familiar to us, but was well known to Kant. The threat as Kant faced it can be expressed thus: "If I will what I will for reasons, is it not fair to say that reasons determine my will?" Or, more specifically, "If it is inevitable that I will whatever I judge best, and if judgment is a process of applying standards and rules, am I not therefore bound to will according to rules and standards?" Kant's solution to the challenge posed by reason in fact echoes Leibniz's solution to the scholastic and early modern problem of voluntarism. For the strict voluntarist, as noted above, God's will is absolutely free, unconstrained by any independent nature of things or by rules or standards. Neither the nature of space, time, and matter, nor the rules of math and logic, nor standards of beauty and goodness can constrain God; on the contrary, all must be at his pleasure. Otherwise, God's freedom would not be complete or perfect (which, like God's goodness and justice and knowledge, it must be). God must be able to choose and make real or true whatever he wills, even if it is that two plus three equals six, or that something be and not be at the same time.

Now one of the deepest problems facing those with a strong commitment to voluntarism is that it renders God both arbitrary and incomprehensible. If we have before us the thought that God can change the natures of things and the rules that govern them, willy-nilly, we cannot hope to understand God or his creation, or to make sense of his aims and purposes for us and for the world he created. God becomes a fearsome and arbitrary despot, not a just and wise creator and sustainer. And as Leibniz argued, the praise and worship God expects, and which we offer, make no sense unless we have good independent grounds for judging him praiseworthy.<sup>19</sup>

Leibniz offers an ingenious solution to the dilemma presented by voluntarism, maintaining God's freedom to follow rules or not, and to make rules as he chooses, without making him arbitrary or incomprehensible. For Leibniz, God creates rules and laws and standards according to what in some all-told, general sense seems best to him, though he doesn't have to; he could author sub-optimal rules. Even having authored optimal rules, God does not have to follow them. Instead, God *chooses* to follow those rules, and, as it turns out, does so over and over again. But it is still true that he doesn't have to author the best rules, or to follow them: we worship him because he reliably does so, because he chooses to do what is best, given his

<sup>19</sup> G. W. Leibniz, *Discourse on Metaphysics* (1686), in *Philosophical Essays*, trans. Roger Ariew and Daniel Garber (Indianapolis: Hackett Publishing, 1989), 36–7, §§2–3.



own overall background all-told sense of ‘best.’ The key, of course, lies in there being some such overall background all-told sense, one that belongs to God as God, and that is also comprehensible to us. Leibniz nominates God’s interest in maximal harmony with maximal diversity as providing such an overall background all-told global criterion of ‘bestness’: let there be as many things as possible organized as elegantly as possible, thinks God. (It’s an appropriate most general interest for a creator and sustainer, isn’t it?) God’s very general interest in maximal diversity joined with maximal harmony cannot be properly said to *constrain* his will, but instead represents his will’s most basic contents – and he could, in any case, choose to ignore it.

Kant’s effort to conceive the human will and its freedom bears deep similarities to the solution Leibniz offers to the dilemma presented by voluntarism. How so? God’s will, for Leibniz, and human will, for Kant, are both powers to choose things and to initiate action aimed at making those things real or true.<sup>20</sup> God’s will for Leibniz, and human will for Kant, are both free when this power to choose is not determined by external forces, but instead by the will itself, on grounds that are its own.<sup>21</sup> Just as Leibniz’s God has a fundamental desire that befits him, *qua* God, so free human will, for Kant, has a fitting interest, namely an interest in its own continued free agency. For Kant, this interest is both fully mine and fully rational – it is not fundamentally external to me. In contrast, grounds of action that stem ultimately from my physical body, with its physically determined needs and desires, *are* seen by Kant as foreign or external to me; they are the equivalent of rules or standards not my own. So, for Kant, are determining grounds attributable to ‘the will of God’ – for Kant, a person who puts God’s will before her own forsakes her own freedom.<sup>22</sup>

Like God’s will for Leibniz, Kantian free will *can* depart from reason, perversely deciding against what in some most general sense seems best to it: for both, will remains radically free.<sup>23</sup> But when it is responsive to reason, this responsiveness precisely honors its own deepest internal structure and commitments. For Kant, the will’s responsiveness to reason honors will’s own deep interest in free rational agency, that is, in itself. (God, being

<sup>20</sup> See Kant, KpV 5:9n and MS 6:211–13; see G. W. Leibniz, “On Freedom and Possibility” [1680–82?], in *Philosophical Essays*, trans. Roger Ariew and Daniel Garber (Indianapolis: Hackett Publishing, 1989), 19–23.

<sup>21</sup> See Kant, G 4:446, for example; see Leibniz, *Discourse on Metaphysics*, §30, pp. 60–1.

<sup>22</sup> If, however, an agent decides, for reasons of her own, that God’s will is good, and that she therefore endorses its demands, she maintains her freedom. At stake here is whether the ultimate standard of goodness, the standard guiding judgment, is understood to lie within (maintaining freedom) or without (impairing it). See REL 6:177.

<sup>23</sup> See, for example, Kant, KpV 5:20; see Leibniz, “On Freedom and Possibility,” 19–23.

perfect, will always honor his deepest internal structure and commitments – we, on the other hand, often mess this one up.) A will, on this ‘Leibnizo-Kantian’ view, does not therefore have to act in ways that are arbitrary or indifferent to be free. A free will can be responsive to (though not wholly determined by) reasons for acting, without any diminution of freedom, given that the choice to be responsive was its own. Moreover, insofar as the will chooses to be responsive to grounds that are truly its own, that is, insofar as the will chooses in ways that ‘make sense’ (are rational) given its own interests, it maintains its full freedom at the same time that it completely *self-determines*.

We can summarize these thoughts by noticing how they fit with Kant’s helpful distinction between *Willkür* – the capacity to choose – and *Wille* – the faculty of practical reason as a whole. Both, as the words themselves suggest, are part of will.<sup>24</sup> And both contribute distinctive elements or components to human freedom. The first, *Willkür*, is the capacity for free choice itself, the capacity to choose ‘at will’ between alternatives – alternative ends, alternative courses of action, alternative guiding principles of action. For Kant, *Willkür* is metaphysically necessary for morality since without it praise and blame and responsibility-holding would not make sense: to be held responsible, to be considered the author of an action, an agent must be the ultimate source of her choices.

*Wille*, the second term, is the capacity to formulate ends, and to formulate action-guiding principles aimed at serving those ends. Thus does Kant call *Wille* ‘practical reason itself’:<sup>25</sup> *Wille* conceptualizes and formulates in ways that actually guide practice, or intentional action. For Kant, ends and action-guiding principles formulated by *Wille* insofar as it seeks grounds within itself and not in external sources, that is, ends and action-guiding principles formulated by *pure* practical reason, count, not surprisingly, as ends and action-guiding principles that are deeply *mine*. Such ends and principles are grounded in interests internal, for Kant, to my deepest self, my free rational self. And by choosing to act in accordance with such purely rational ends and principles, I choose action that is given aim and shape by this self. Of course, once I choose a course of action, I am determined – I am no longer exercising a capacity to go this way or that – but if I have chosen to act toward ends and on principles that are truly my

<sup>24</sup> As is noted below, and as will be discussed at greater length in Chapter 4, ‘*Wille*’ is both the Kantian term for that part of will that is specifically rational, and for will as a whole – just as ‘reason’ for Kant is sometimes the term for pure reason (in contrast to sensibility or understanding) and sometimes for the whole faculty (including sensibility and understanding).

<sup>25</sup> MS 6:213. See also MS 6:226; G 4:412.









































































































































































































































































































































































































