

## KANT'S *METAPHYSICS OF MORALS*

Immanuel Kant's *Metaphysics of Morals* (1797), containing the *Doctrine of Right* and *Doctrine of Virtue*, is his final major work of practical philosophy. Its focus is not rational beings in general but human beings in particular, and it presupposes and deepens Kant's earlier accounts of morality, freedom, and moral psychology. In this volume of newly commissioned essays, a distinguished team of contributors explores the *Metaphysics of Morals* in relation to Kant's earlier works, as well as examining themes which emerge from the text itself. Their topics include the relation between right and virtue, property, punishment, and moral feeling. Their diversity of questions, perspectives, and approaches will provide new insights into the work for scholars in Kant's moral and political theory.

LARA DENIS is Professor of Philosophy at Agnes Scott College in Decatur, Georgia, USA. She is the author of *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (2001) and the editor of a supplemented edition of Kant's *Groundwork of the Metaphysics of Morals* (2005).

CAMBRIDGE CRITICAL GUIDES

Titles published in this series:

Hegel's *Phenomenology of Spirit*

EDITED BY DEAN MOYAR AND MICHAEL QUANTE

Mill's *On Liberty*

EDITED BY C. L. TEN

Kant's *Idea for a Universal History with a Cosmopolitan Aim*

EDITED BY AMÉLIE OKSENBERG RORTY AND JAMES SCHMIDT

Kant's *Groundwork of the Metaphysics of Morals*

EDITED BY JENS TIMMERMANN

Kant's *Critique of Practical Reason*

EDITED BY ANDREWS REATH AND JENS TIMMERMANN

Wittgenstein's *Philosophical Investigations*

EDITED BY ARIF AHMED

Kierkegaard's *Concluding Unscientific Postscript*

EDITED BY RICK ANTHONY FURTAK

Plato's *Republic*

EDITED BY MARK L. MCPHERRAN

Plato's *Laws*

EDITED BY CHRISTOPHER BOBONICH

Kant's *Metaphysics of Morals*

EDITED BY LARA DENIS

Spinoza's *Theological Political Treatise*

EDITED BY YITZHAK Y. MELAMED AND MICHAEL ROSENTHAL

KANT'S  
*Metaphysics of Morals*  
A Critical Guide

EDITED BY  
LARA DENIS  
*Agnes Scott College*

 CAMBRIDGE  
UNIVERSITY PRESS

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

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  
Information on this title: www.cambridge.org/9780521513937

© Cambridge University Press 2010

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

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this publication is available from the British Library*

*Library of Congress Cataloguing in Publication data*  
Kant's Metaphysics of morals : a critical guide / [edited by] Lara Denis.  
p. cm. – (Cambridge critical guides)

Includes bibliographical references (p. ) and index.

ISBN 978-0-521-51393-7 (hardback)

1. Kant, Immanuel, 1724–1804. Metaphysik der Sitten. 2. Ethics.  
I. Denis, Lara, 1969– II. Title. III. Series.

B2766.Z7K375 2010

170–dc22

2010024610

ISBN 978-0-521-51393-7 Hardback

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.

## Contents

<i>Notes on contributors</i>	page vii
<i>Acknowledgements</i>	ix
<i>List of translations and abbreviations</i>	x
Introduction	I
LARA DENIS	
1 Kant's <i>Metaphysics of Morals</i> : the history and significance of its deferral	9
MANFRED KUEHN	
2 Reason, desire, and the will	28
STEPHEN ENGSTROM	
3 Justice without virtue	51
KATRIN FLIKSCHUH	
4 Kant's innate right as a rational criterion for human rights	71
OTFRIED HÖFFE	
5 Intelligible possession of objects of choice	93
B. SHARON BYRD	
6 Punishment, retribution, and the coercive enforcement of right	111
ALLEN W. WOOD	
7 Moral feelings in the <i>Metaphysics of Morals</i>	130
PAUL GUYER	
8 What is the enemy of virtue?	152
JEANINE GRENBERG	

9	Freedom, primacy, and perfect duties to oneself LARA DENIS	170
10	Duties to and regarding others ROBERT N. JOHNSON	192
11	Duties regarding animals PATRICK KAIN	210
12	Kant's <i>Tugendlehre</i> as normative ethics THOMAS E. HILL, JR.	234
	<i>Bibliography</i>	256
	<i>Index</i>	264

## Contributors

B. SHARON BYRD is Professor of Anglo-American Law and Jurisprudence at the Friedrich Schiller University School of Law in Jena, Germany. She recently co-authored *Kant's Doctrine of Right: A Commentary* with Joachim Hruschka (2010). She is a co-editor of the *Jahrbuch für Recht und Ethik*.

LARA DENIS is Professor of Philosophy at Agnes Scott College in Decatur, Georgia. Among her publications are *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (2001) and articles in the *Kantian Review*, *Philosophy and Phenomenological Research*, *History of Philosophy Quarterly*, *Kant-Studien*, and elsewhere.

STEPHEN ENGSTROM is Professor of Philosophy at the University of Pittsburgh. He is a co-editor of *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty* (1996) and the author of *The Form of Practical Knowledge: A Study of the Categorical Imperative* (2009).

KATRIN FLIKSCHUH is Reader in Political Theory in the Department of Government at the London School of Economics. In addition to numerous articles in journals and edited volumes, she is author of *Kant and Modern Political Philosophy* (2000) and *Freedom: Contemporary Liberal Perspectives* (2007).

JEANINE GRENBERG is Associate Professor of Philosophy at St. Olaf College in Northfield, Minnesota. She has published *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue* (2005) as well as articles in *Philosophy and Phenomenological Research*, *Journal of the History of Philosophy*, *Kant-Studien*, and *Social Theory and Practice*.

PAUL GUYER is Professor of Philosophy and the Florence R.C. Murray Professor in the Humanities at the University of Pennsylvania. Among his recent books are *Values of Beauty: Historical Essays in Aesthetics* (2005),

*Kant's System of Nature and Freedom* (2005), *Kant* (2006), and *Knowledge, Reason, and Taste: Kant's Response to Hume* (2007).

THOMAS E. HILL JR. is Kenan Professor of Philosophy at the University of North Carolina, Chapel Hill. His essays on moral and political philosophy are collected in *Autonomy and Self-Respect* (1991), *Dignity and Practical Reason in Kant's Moral Theory* (1992), *Respect, Pluralism, and Justice: Kantian Perspectives* (2000), and *Human Welfare and Moral Worth: Kantian Perspectives* (2002).

OTFRIED HÖFFE is Head of the Research Center for Political Philosophy and Professor of Philosophy at the University of Tübingen. His books include *Lebenskraft und Moral. Oder: Macht Tugend glücklich?* (2007), *Kants Kritik der reinen Vernunft. Die Grundlegung der Modernen Philosophie* (2003, 2004), *Immanuel Kant. Leben – Werk – Wirkung* (1983), and *Praktische Philosophie – Das Modell des Aristoteles* (1971).

ROBERT N. JOHNSON is Professor of Philosophy at the University of Missouri. He has published articles in journals such as *Ethics*, *Philosophical Quarterly*, *Philosophical Studies*, and *Pacific Philosophical Quarterly*, and wrote the entry “Kant's Moral Philosophy” in the *Stanford Encyclopedia of Philosophy*.

PATRICK KAIN is Associate Professor of Philosophy at Purdue University. He is a co-editor and contributor to *Essays on Kant's Anthropology* (2003), and has published articles in journals such as *Journal of the History of Philosophy*, *Archive für Geschichte der Philosophie*, *Kantian Review*, and *Philosophy Compass*.

MANFRED KUEHN is Professor of Philosophy at Boston University. He is author of *Scottish Common Sense in Germany* (1988) and *Immanuel Kant: A Biography* (2001) as well as of many papers on Hume, Kant, Mendelssohn, Reid, and the Enlightenment.

ALLEN W. WOOD is Ward W. and Priscilla B. Woods Professor at Stanford University. His books include *Kant's Moral Religion* (1970), *Kant's Rational Theology* (1978), *Hegel's Ethical Thought* (1990), *Kant's Ethical Thought* (1999), and *Kantian Ethics* (2008).



## *Acknowledgements*

My thanks to Hilary Gaskin of Cambridge University Press for inviting my proposal for a Cambridge Critical Guide to the *Metaphysics of Morals*, and to others at the Press, including Joanne Tunncliffe, Gillian Dadd, and Aline Guillermet, and to the anonymous external readers, for helping me bring the project to completion. Thanks also to this collection's contributors, whose excellent essays I have already benefited from reading. Allen Wood merits special appreciation for providing counsel and encouragement during the proposal preparation stage. I also appreciated Patrick Kain's correspondence with me about the provenance of Kant's lecture notes. Roger Wertheimer has, as ever, been a steady source of love and support, for which I am grateful.

## *Translations and abbreviations*

Unless footnoted otherwise in their essays, authors use translations from the Cambridge Edition of the Works of Immanuel Kant, series editors Paul Guyer and Allen W. Wood (New York: Cambridge University Press, 1992–). Page citations are to *Kants gesammelte Schriften*, Ausgabe der Königlich Preussischen Akademie der Wissenschaften (Berlin: Walter de Gruyter, 1902–) using the volume:page number format or the volume:page number:line format; exceptions are noted. In some cases section numbers or headings are used in addition to or instead of page numbers. Abbreviations authors use within citations include the following:

- AA      Akademie Ausgabe  
CE      Cambridge Edition  
ApH    *Anthropologie in pragmatischer Hinsicht* [1798] (AA 7)  
          *Anthropology from a Pragmatic Point of View*, Robert B. Louden (trans.) (CE *Anthropology, History, and Education*, Robert B. Louden and Günter Zöller [eds.], 2007)  
BBM    “Bestimmung des Begriffs einer Menschenrasse” [1785] (AA 8)  
          “Determination of the Concept of a Human Race,” Holly Wilson and Günter Zöller (trans.) (CE *Anthropology, History, and Education*)  
Br.     Briefe (AA 10–13)  
          Kant’s letters (CE *Correspondence*, Arnulf Zweig [ed. and trans.], 1999)  
C      Moralphilosophie Collins [reflecting notes originally from 1774 to 1777; transcribed for or by Collins, 1784–1785] (AA 27)  
          Collins notes on Kant’s moral philosophy lectures (CE *Lectures on Ethics*, Peter Heath [ed. and trans.] and J.B. Schneewind [ed.], 1997)

- ED *Das Ende aller Dinge* [1794] (AA 8)  
*The End of All Things*, Allen W. Wood (trans.) (CE  
*Religion and Rational Theology*, Allen W. Wood and George  
di Giovanni [eds.], 1996)
- F Naturrecht Feyerabend [1784] (AA 27)  
Kant's lectures on natural right (to appear, translated by  
Frederick Rauscher, in CE *Lectures and Drafts on Political  
Philosophy*, Frederick Rauscher and Kenneth Westphal  
[eds.], forthcoming)
- FS "Die falsche Spitzfindigkeit der vier syllogistischen Figuren  
erwiesen" [1762] (AA 2)  
"The False Subtlety of the Four Syllogistic Figures" (CE  
*Theoretical Philosophy, 1755–1770*, David Walford and Ralf  
Meerbote [eds.], 2003)
- G *Grundlegung zur Metaphysik der Sitten* [1785] (AA 4)  
*Groundwork of the Metaphysics of Morals* (CE *Practical  
Philosophy*, Mary J. Gregor [ed. and trans.], 1996)
- GtP "Über den Gebrauch teleologischer Principien in der  
Philosophie" [1788] (AA 8)  
"On the Use of Teleological Principles in Philosophy,"  
Günter Zöller (trans.) (CE *Anthropology, History, and  
Education*)
- H Praktische Philosophie Herder [1762–64] (AA 27)  
Herder's notes on Kant's lectures on practical philosophy.  
Translated selections appear in CE *Lectures on Ethics*.
- HN Handschriftlicher Nachlass (AA 14–23)  
Handwritten remains (selections in CE *Notes and  
Fragments*, Paul Guyer [ed.], Curtis Bowman, Paul Guyer,  
Frederick Rauscher [trans.], 2000, and in CE *Lectures and  
Drafts on Political Philosophy*)
- KpV *Kritik der praktischen Vernunft* [1788] (AA 5)  
*Critique of Practical Reason* (CE *Practical Philosophy*)
- KrV A/B *Kritik der reinen Vernunft* [1781, 1787], cited by A (first  
edition) / B (second edition)  
*Critique of Pure Reason*, Paul Guyer and Allen Wood  
(eds. and trans.) (CE, 1998)
- KU *Kritik der Urteilskraft* [1790] (AA 5)  
*Critique of the Power of Judgment*, Paul Guyer (ed. and  
trans.) (CE, 2000)

- LJ *Immanuel Kants Logik, ein Handbuch zu Vorlesungen*, G.B. Jäsche (ed.) [originally published Königsberg: Friedrich Nicolovius, 1800] (AA 9)  
     Jäsche Logic (Part IV of CE *Lectures on Logic*)
- M I Moral Mrongovius I [1774–77] (AA 27)  
     Mrongovius notes on Kant’s moral philosophy (brief passages used to supplement Collins notes in CE *Lectures on Ethics*)
- M II Moral Mrongovius II [1784–85] (AA 29)  
     Mrongovius notes on Kant’s moral philosophy lectures (selections in CE *Lectures on Ethics*)
- MS *Die Metaphysik der Sitten* [1797–98] (AA 6)  
     *Metaphysics of Morals* (CE *Practical Philosophy*)
- P Praktische Philosophie Powalski [1782–83] (AA 27)  
     Powalski notes on Kant’s practical philosophy lectures
- PG *Immanuel Kants physische Geographie*, F.T. Rink (ed.) [originally published Königsberg: Göbbels and Unzer, 1802] (AA 9)  
     *Physical Geography* (forthcoming, Davod Oldroyd and Olaf Reinhardt (trans.), in CE *Natural Science*, H.B. Nisbet and E. Watkins [eds.], forthcoming)
- Pr. *Prolegomena zu einer jeden künftigen Metaphysik* [1783] (AA 4)  
     *Prolegomena to Any Future Metaphysics*, Gary Hatfield (trans.) (CE *Theoretical Philosophy after 1781*, Henry Allison and Peter Heath [eds.], 2002)
- PS Preisschrift – “Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral” [1764] (AA 2)  
     “*Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality*” (CE *Theoretical Philosophy, 1755–1770*)
- R Reflexionen (AA 14–19)  
     Selections appear in CE *Notes and Fragments* and CE *Lectures and Drafts on Political Philosophy*. Citations include Reflexionen number.
- RGV *Die Religion innerhalb der Grenzen der blossen Vernunft* [1793] (AA 6)  
     *Religion within the Boundaries of Mere Reason* (CE *Religion and Rational Theology*)
- RL *Metaphysische Anfangsgründe der Rechtslehre* [1797] (AA 6)  
     *Metaphysical Foundations of the Doctrine of Right*, part one and the first published part of the *Metaphysics of Morals* (CE *Practical Philosophy*)

- TL *Metaphysische Anfangsgründe der Tugendlehre* [1797] (AA 6)  
*Doctrine of Virtue*, part two and the second published part of  
the *Metaphysics of Morals* (CE *Practical Philosophy*)
- TP “Über den Gemeinspruch: Das mag in der Theorie richtig sein,  
taugt aber nicht für die Praxis” [1793] (AA 8)  
“On the Common Saying: That May be Correct in Theory,  
but it is of No Use in Practice” (CE *Practical Philosophy*)
- V Die *Metaphysik der Sitten* Vigilantius [1793–94] (AA 27)  
Vigilantius notes on Kant’s lectures on the metaphysics of  
morals (CE *Lectures on Ethics*)
- VA Vorlesungen über Anthropologie (AA 25)  
Lectures on Anthropology (selections to appear in CE  
*Lectures on Anthropology*, Allen Wood and Robert Louden  
[eds.], forthcoming)
- VM Vorlesungen über Metaphysik (AA 28, 29)  
Lectures on Metaphysics (selections in CE *Lectures on*  
*Metaphysics*, Karl Ameriks and Steve Naragon [eds. and trans.],  
1997)
- VpR Vorlesungen über die philosophische Religionslehre [1783–  
86?] (AA 28)  
Lectures on the Philosophical Doctrine of Religion, Allen  
W. Wood (trans.) (CE *Religion and Rational Theology*)
- VR “Von den verschiedenen Racen der Menschen, zur  
Ankündigung der Vorlesungen der physischen Geographie im  
Sommerhalbjahr” [1775] (AA 2)  
“Of the Different Races of Human Beings,” Holly Wilson  
and Günter Zöller (trans.) (CE *Anthropology, History, and*  
*Education*)
- VRL Vorarbeiten zu *Die Metaphysik der Sitten*, Erster  
Teil: *Metaphysische Anfangsgründe der Rechtslehre* (AA 23)  
Preparatory Work on the *Metaphysics of Morals*, First  
Part: *Metaphysical Principles of the Doctrine of Right*
- ZeF “Zum ewigen Frieden” [1795] (AA 8)  
“Toward Perpetual Peace” (CE *Practical Philosophy*)



# *Introduction*

*Lara Denis*

The *Metaphysics of Morals*, Kant's final major work in moral philosophy, is wide-ranging, complex, and often provocative. Its focus is not rational beings as such, but human beings in particular, and the duties, rights, and morally practical relations that obtain among us. As a whole, the *Metaphysics of Morals* deepens Kant's accounts of morality, moral psychology, and the spheres of right and virtue. Its *Doctrine of Right* sets forth not only fundamental, coercively enforceable principles of interpersonal conduct, but also a rational idea of a peaceful community of all nations. Its *Doctrine of Virtue* develops Kant's conception of virtue and accounts of particular ethical duties. Aspects of the *Metaphysics of Morals*, such as its overall lack of abstraction relative to the *Groundwork* or second *Critique* and its richness of anthropological detail, recall Kant's lectures on ethics. Other features, such as its preliminary discussion of the relation of the faculty of desire to practical principles, clearly exhibit the commitments of his critical moral philosophy.

Some of the positions that Kant articulates here – for example, about the relationship between the fundamental principles of right and the categorical imperative, or the role of feeling in moral motivation – have long been sources of dispute among Kant's interpreters. Some of his arguments – for instance, those concerning capital punishment, or duties regarding non-human animals – have ignited debate among ethicists more broadly. The *Metaphysics of Morals* invites questions about where and how Kant augments, refines, revises, withdraws, or supplants arguments and positions set forth in earlier works, as well as about how best to understand the arguments and positions he provides here, and how plausible, defensible, or distinctive they are.

The dozen essays collected here address questions both interpretive and philosophical. They focus on passages, positions, or arguments in the *Metaphysics of Morals* that strike us as particularly interesting and important – and that we hope will engage the interest of colleagues specializing

in ethics, political philosophy, Kant interpretation, and the history of philosophy. Many of the essays in this collection draw heavily on other (often less-studied) works by Kant, thereby enriching our understanding of their topic and the relevant arguments in the *Metaphysics of Morals*. There is much here both for readers interested in the development of Kant's own thought or the history of ethics and for those curious about how best to construe and evaluate various facets of Kant's mature moral and political philosophy.

A natural place to begin our exploration is by asking about the nature of Kant's project in the *Metaphysics of Morals*. How did he conceive of this work? Why did it take so long for him to write it? How does it relate to his more overtly "critical" works in ethics? In the opening chapter, Manfred Kuehn situates the *Metaphysics of Morals* within the broad contexts of Kant's philosophical development and ethical thought. He argues that Kant's *Metaphysics of Morals* should not be disqualified as precritical, for although it contains precritical elements, they are not what is most important about it. Furthermore, it would be wrong simply to judge the *Groundwork* and the second *Critique* from the point of view of "the final form of Kant's practical philosophy," for the concerns of the *Metaphysics of Morals* are different from those of his critical philosophy. In order to establish the significance and place of Kant's late work on moral philosophy in relation to his philosophy as a whole, Kuehn first explores the history – or prehistory – of its composition, which is primarily a history of postponements. The projected *Metaphysics of Morals* radically changed several times in these deferrals, and the ways in which the project changed are not insignificant for understanding the place it ultimately assumed. Second, Kuehn briefly indicates what Kant's lectures reveal about his project of a "Metaphysics of Morals," suggesting that the relevance of these lectures has been largely underestimated in discussions of the "final form" of Kant's practical philosophy. Third, Kuehn shows how the *Metaphysics of Morals* is related to the *Groundwork* and the second *Critique*. He argues that the later work both responds to concerns different from those of the earlier two, and yet also reveals an important, new perspective on issues central to those earlier works, such as the categorical imperative.

The rest of the essays in the first half of the book focus on questions raised primarily by Part I of the *Metaphysics of Morals*, that is, the Introduction to the *Metaphysics of Morals* and the *Doctrine of Right* proper. In chapter 2, Stephen Engstrom turns our attention to a topic within the Introduction to the *Metaphysics of Morals* that goes to the heart of Kant's moral philosophy: Kant's conception of the will as it relates to reason and



desire. While much attention has been devoted to Kant's doctrine of the will's autonomy, Engstrom points out that Kant advances another striking proposition about the will: it is nothing but practical reason. Engstrom contends that this less-investigated idea is vitally important in its own right as well as for the light it throws on other aspects of his ethical thought, including autonomy. According to Engstrom, Kant *transforms* the traditional understanding of the will as rational desire. Kant's way of combining the notions of the will, reason, and desire in his conception of a practical or efficacious employment of the cognitive capacity of reason is responsible for much of what is distinctive in his moral philosophy. After outlining Kant's conception of reason as a cognitive capacity, Engstrom takes up Kant's conception of desire as a form of causality peculiar to living beings, which provides the broad heading under which he situates the will and relates it to the power of choice. He then draws on the preceding accounts of reason and desire to elucidate Kant's account of the will. Engstrom's interpretation places the will in the desiderative economy of human life, while underscoring the cognitivist character of Kant's conception of the will and of his practical philosophy more generally.

Chapter 3 offers a critical discussion of ongoing debates concerning the moral status of Kant's philosophy of right. Katrin Flikschuh defends an account of right as a public morality and, as such, as systematically distinct from the personal morality of Kant's ethics. It follows that the principle of autonomy, as a principle of *ethical* self-obligation, has no place within the philosophy of right which, as public, concerns the morality of *external* legislation. From the irrelevance of the principle of autonomy the non-moral status of right does not, however, follow. Flikschuh employs Kant's *Wille/Willkür* distinction to show that, within the domain of right, the a priori idea of the general united will replaces the principle of autonomy as the ground of moral obligation. As *Wille* in its juridical conception, the idea of the general united will locates the grounds of juridical moral obligation outside the subjective will of the individual agent, ensuring conformity of action of the subjective *Willkür* of each with universal laws of right. The externally free agent is, as such, *non-autonomously* free. One implication of this view is that the presumed centrality of the principle of autonomy to Kant's practical philosophy in general must be reassessed; the principle of autonomy is derivative of the philosophically deeper idea of freedom itself. Flikschuh resists attempts to close the gap between ethical and political judgment. Instead of seeking to align political with ethical judgment, we should, she says, acknowledge the political as a distinct mode of public moral judgment.

Chapter 4 considers questions concerning Kant's conception of human rights (or of the *one* human right) and how it operates within Kant's political philosophy – especially according to the *Rechtslehre*. Here, Otfried Höffe explicates the innate right to freedom, which, Kant says, belongs to every human being “by virtue of his humanity.” He aims to show how this single innate right functions as a rational criterion for multiple human rights. To this end, Höffe clarifies Kant's distinction between moral (natural) right and positive (empirical) right; the relation between the moral concept of right, the universal principle of Right, and the universal law of Right; and Kant's justification of coercive enforcement of legal rights. Provocatively, Höffe argues that Kant views self-recognition – specifically, the practical recognition of humanity in one's own person by upholding one's rightful honor and refusing to submit to legal degradation – as a primary condition for establishing oneself as a legal entity. Höffe then explores the derivation of the four human rights Kant regards as implicit within the innate right to freedom. Finally, he suggests that while the rights to own property and to live in a public legal order are not, for Kant, human rights in the strict sense, they are grounded in such a way as to be considered “quasi-human rights.”

Chapter 5 addresses a question fundamental to Kant's doctrine of private right: how to have something external as one's own. Sharon Byrd traces Kant's arguments and shows that they turn on his concepts of possession. These concepts move from an *empirical concept* of possession as having something in one's hand to an *intelligible concept* of possession as having something as one's own based on a duty all others have not to interfere with what one intelligibly possesses. His arguments depend on the postulate of practical reason. This postulate has been interpreted to provide a justification. A justification, however, suggests that what would otherwise be wrongful or prohibited conduct is rightful conduct because of the situation. Byrd's position is that there is nothing wrongful about taking something external to oneself and calling it one's own. The taker thus does not need any justification for doing so. Byrd here relies on an alternative interpretation of the postulate as a *power-conferring norm*. On this reading, the postulate empowers us to have external objects of our choice as our own. Although we may unilaterally impose an obligation on all others to respect what we have declared to be ours, this power flows from our will's compatibility with the universal united will. Nothing in Kant's arguments for individual rights to have objects of choice as our own depends on the existence of a state. Indeed Kant notes that without a right to property and other objects of our choice there would be no duty

to move to the civil social order. Property rights therefore are rights we have in the state of nature. They do not depend on social approval any more than our right to freedom of choice in general depends on social approval and recognition. The sole purpose of the state for Kant is securing rights we already have before leaving the state of nature and moving to the civil state. The state secures our right to freedom and our rights to external objects of our choice.

Chapter 6 tackles questions of the substance and justification of Kant's theory of punishment. Regarding issues about legal punishment, Kant is best known as a defender of an extreme retributivist position on the justification of punishment. Allen Wood argues, however, that the deeper truth about his views on this topic is far more complex and even troubling. According to Wood, although Kant is undoubtedly a retributivist, the justification of punishment Kant provides that is best rooted in his theory of right is *not* a retributivist one. Furthermore, Kant's retributivism is apparently *inconsistent* with some fundamental tenets of his practical philosophy.

The remaining chapters focus primarily on the *Doctrine of Virtue*. In chapter 7, Paul Guyer considers the relation of feelings to moral motivation, and traces the development of Kant's view of this matter. Kant is commonly supposed to have excluded all feeling from the incentives of morally worthy action, and accepted only the determination of the will by the moral law itself as a morally worthy motivation. Guyer shows that this view ignores Kant's increasingly sophisticated moral psychology, which reaches its zenith in the *Metaphysics of Morals*. In the *Groundwork*, Kant recognizes a feeling of respect as the *effect* of the moral law, but does not assign it any clear role in the etiology of moral action. By the *Critique of Practical Reason*, however, Kant clearly holds that the feeling of respect plays a *causal role* in the production of morally worthy action at the phenomenal level, even though he is unclear what this role is. Finally, in the Introduction to the *Doctrine of Virtue*, Kant refines this recognition into a sophisticated theory of the "aesthetic preconditions" of receptivity to duty, or complex causes of moral action at the phenomenal level, and argues that the cultivation of these predispositions is a fundamental feature of what he called, much earlier, "moral praxis."

Chapter 8 confronts a fundamental question for readers of the *Doctrine of Virtue*, namely, what is Kant's conception of virtue? Jeanine Grenberg seeks to understand Kantian virtue indirectly, by asking: what is the *enemy* of virtue? What explains the empirically undeniable fact that becoming virtuous is a struggle, something accomplished in the face of some

opposing force? As Grenberg sees it, if we do not understand what virtue has overcome, we do not really understand the state that results from the struggle. Kant, however, appears contradictory, or at least ambiguous, in identifying the enemy of virtue. He sometimes suggests that we must struggle against our inclinations; yet at other times he suggests that it is a corruption of reason itself that is the true enemy. Grenberg investigates both lines of thought, eventually showing that Kant's apparently contradictory claims in the *Metaphysics of Morals* and *Religion within the Limits of Mere Reason* can be reconciled. Ultimately, she argues, the central connection Kant makes between virtue as strength and inner freedom in the *Metaphysics of Morals* can make sense only if we reject any natural opponents of virtue and admit that the battle for virtue takes place on the territory of reason and freedom. Finally, Grenberg argues that by appeal to Kant's notion of an internal, rational enemy of virtue, we can more clearly distinguish Kantian and Aristotelian virtue.

Chapter 9 considers the primacy of perfect duties to oneself within Kant's moral theory. Kant makes a variety of striking pronouncements about the significance of perfect duties to oneself. But what exactly is the nature of their primacy, and why do they have it? To answer these questions, I explore the *Doctrine of Virtue* account of these duties (as concerned with one's moral self-preservation and moral health), along with two of Kant's earlier accounts from his lectures. In the Collins lecture notes, Kant explicates perfect duties to oneself as prohibiting acting against the necessary conditions of one's greatest, self-consistent use of freedom ("the essential ends of humanity"); in the Vigilantius lecture notes, as duties to which we are directly constrained by humanity in one's person (and thus as immediately grounded in "the right of humanity in our own person"). I show that on all three accounts, perfect duties to oneself bear especially fundamental, vital, and direct relations to freedom, and that these relations generate multiple, interrelated sorts of primacy for these duties.

In chapter 10, Robert Johnson raises and resolves an apparent puzzle about the duty to adopt others' happiness as our end. Because this is a wide and imperfect obligation, no one has a claim on our assistance in advancing her happiness in particular. However, in general, that we have an obligation *to* someone, as opposed to merely *regarding* her, is best understood as her having some claim on us. This apparently generates a puzzle: if we have a duty to others to adopt their end, then it seems that others have a claim on our so doing; but if our duty is wide and imperfect, then no one has a claim on our having her happiness as our end. Johnson shows that the puzzle arises only if we assume that there cannot be a collective

right – the right of collective “others” – and argues that Kant’s followers should not assume this. Johnson begins by exploring Kant’s views on the nature of beings to whom we can owe duties and about what it means to have a duty to a person rather than merely regarding her. Johnson then turns to the relationship between our having duties to a person and the claims she may have on us as a result of such duties. Johnson argues that duties and rights are reciprocal, that there are “collective” rights, and that the reciprocity between rights and duties allows us to distinguish those *to whom* we have duties from those only *with regard to whom* we have duties in the context of Kant’s duties to others. We are then in a position to understand how the wide, imperfect duty to promote others’ happiness is nevertheless genuinely a duty *to* others.

The theme of duties *to* and duties *regarding* others continues in chapter II, which sheds new light on Kant’s provocative insistence that we have duties *regarding*, but not *to*, non-human animals. Patrick Kain confronts questions both interpretive and philosophical about Kant’s account of moral status. Kain shows that a better appreciation of Kant’s commitments in a variety of disciplines reveals that Kant had a deeper understanding of human and non-human animals than commentators generally recognize. This sheds new light on Kant’s claims about the nature and scope of moral status, and helps to address, at least from Kant’s perspective, many of the familiar objections to his notorious account of “duties regarding animals.” According to Kain, Kant’s core principles about the nature of moral obligation structure his thoughts about the moral status of human beings and non-human animals. Through an examination of a broad array of little-studied sources, Kain shows that Kant’s commitments in biology, psychology, anthropology, and physical geography support his account of the nature of and distinction between humans and non-human animals. Kain argues that this account supports Kant’s judgment that we have duties to every human being and significant duties regarding non-human animals, duties which involve direct concern for animals because of their nature. Finally, by comparing Kant’s account with some recently proposed Kantian alternatives, Kain offers us additional perspective on some of the distinctive features, and strengths and weaknesses, of Kant’s approach.

The final chapter of the collection reviews, highlights, and raises questions about themes in Kant’s *Metaphysics of Morals*, especially the *Doctrine of Virtue*. In a wide-ranging discussion, Thomas E. Hill, Jr. comments briefly on how Kant’s normative ethics relates to science, metaphysics, metaethics, and philosophy of law and justice; the relation of

Kant's first principles to more specific moral principles and judgments; the contrast between contemporary and Kantian conceptions of duties to oneself; problems regarding moral negligence, self-deception, and weakness of will; and moral motivation. Hill emphasizes the constraints of law and justice on virtue, the moral (if not epistemological) priority of the first principles of ethics, and the irrelevance of many contemporary objections to Kant's conception of duties to oneself. Hill also highlights Kant's important second-order principles regarding due care in deliberation, self-scrutiny to expose excuses, and strength of will to resist temptations. The morally necessary motive of duty is interpreted, not as an extra duty added onto each particular duty, but as the basic choice to maintain a pervasive attitude that places moral responsibility before self-interest. Hill's exploration provides a fresh, broad perspective on Kant's mature normative ethics. This essay is a fitting one with which to conclude the collection. Partly this is because it revisits – from a different, illuminating angle – a variety of topics touched on in previous chapters. Equally, however, it is because it treats Kant's *Doctrine of Virtue* not simply as a rich, complicated work of practical philosophy, but also as a vibrant, even viable, normative ethics. By doing this, it encourages readers, whether primarily ethicists or Kant scholars, to plunge still more deeply into Kant's *Doctrine of Virtue*, *Metaphysics of Morals*, and moral philosophy as a whole, to discover all they have to offer.

## CHAPTER I

# *Kant's Metaphysics of Morals: the history and significance of its deferral*

*Manfred Kuehn*

### I INTRODUCTION

Kant's *Metaphysics of Morals* appeared in 1797.<sup>1</sup> It was one of Kant's last works. Only two other books appeared later: *The Dispute of the Faculties* and the *Anthropology from a Pragmatic Point of View*, both published in 1798. The *Logic* of 1801 and the *Physical Geography* of 1802 were edited by others, namely Benjamin Jäsche and Friedrich Theodor Rink, on the basis of Kant's lecture notes. It is tempting to view the *Metaphysics of Morals* and the *Anthropology from a Pragmatic Point of View* also as editions of lecture notes. The difference is just that Kant did the editing himself, though his age and the ever-increasing weakness of his mental faculties made this task very difficult. Some have argued that it might almost have been better if someone else had taken over this task for Kant in the case of the *Metaphysics of Morals* and *Anthropology* as well.<sup>2</sup>

Many of Kant's contemporaries felt this way, in any case. Friedrich Schleiermacher wrote a very negative review of the *Anthropology*, finding that "a summary of this book could not be much more than a collection of trivial matters. If, on the other hand, it were intended to give a sketch of the plan and its execution ... it would necessarily give a distinct picture of the most peculiar confusion."<sup>3</sup> Arthur Schopenhauer found that in the

<sup>1</sup> The title was *Die Metaphysik der Sitten in zwey Theilen. Abgefaßt von Immanuel Kant*. Königsberg, by Friedrich Nicolovius, 1797. "Erster Teil: Metaphysische Anfangsgründe der Rechtslehre ... Zweiter Teil: Metaphysische Anfangsgründe der Tugendlehre." One year later, it appeared again, with the title now indicating a second edition of the *Metaphysical First Principles of Right*, "with an Appendix of Explanatory Remarks and Additions." It also appeared separately as *Erläuternde Anmerkungen zu der Rechtslehre von Immanuel Kant* (Königsberg: Friedrich Nicolovius, 1798). A second edition of the *Metaphysics of Morals* was published in 1803.

<sup>2</sup> In a certain sense this has happened: Bernd Ludwig's edition of the *Metaphysics of Morals* (Hamburg: Meiner Verlag, 1986, 1990). Though Ludwig maintains that he is restoring Kant's original text, eliminating corruptions introduced by an incompetent copyist, there is absolutely no evidence that could establish that it was not Kant himself who introduced the mistakes.

<sup>3</sup> See Friedrich Schleiermacher, *Kritische Gesamtausgabe* (Berlin: Walter de Gruyter, 1984), vol. v.i, 365–69. All other quotations from Schleiermacher are from this review.

*Tugendlehre*, “this counterpart of his deplorable *Rechtslehre*, the effects of his weakness brought on by old age are predominant.”<sup>4</sup> Judgments like these could be multiplied. Thus it has been argued that Kant’s practical philosophy ultimately constituted a relapse into precritical dogmatism, which is not significantly different from the theories of his predecessors and contemporaries.<sup>5</sup> Nor are such verdicts without justification. The *Metaphysics of Morals*, like the *Anthropology*, *Logic*, and *Physical Geography* pales in comparison with the three *Critiques*, and it seems to be less critical than it should be. The centrality of the duty of one’s own perfection, for instance, seems to be a throwback to Wolffian ethics, just as the first legal duty of not harming anyone does not significantly depart from Pufendorf’s natural law theory.

<sup>4</sup> Arthur Schopenhauer, *Werke in fünf Bänden*, Ludger Lüdtkehaus (ed.) (Ulm: Haffmanns Verlag, 1988), vol. III, 475. In the “Critique of Kantian Philosophy” (“Kritik der kantischen Philosophie”) which appeared as an Appendix to the *World as Will and Representation*, he found that the *Rechtslehre*, “one of the latest works by Kant ... is so weak ... that it seems to be not the work of this great man, but the product of an ordinary human being and has to die of its own weakness” (*Werke*, vol. 1, 529–626, 607–08). It seemed to him in many places like a “satirical parody” of Kant.

I shall use “*Rechtslehre*” for the first part and “*Tugendlehre*” for the second part of the work. I will also use these terms to refer to Kant’s concerns with law and virtue throughout his various works. The main reason is that I consider the translation “Metaphysical First Principles of Right” (and, in general, the translation of “right” for “*Recht*”) as seriously misleading. The German word “*Recht*” does not mean what “right” means in English. “*Recht*” is much closer to “law” in English. “Natural law” in German means, for instance, “*Naturrecht*,” and a lawyer is a *Rechtsanwalt*, etc. Since the doctrine of rights is only a part of the doctrine of law in English, this way of translating “*Recht*” tends to identify Kant’s doctrine with only a part of law, and thus to confuse the reader. Mary Gregor argues in her “Translator’s Note on the Text of the *Metaphysics of Morals*,” in Immanuel Kant, *Practical Philosophy*, M.J. Gregor (ed. and trans.) (Cambridge University Press, 1996), 358, that “law” would obscure the conceptual ties of “*das Recht*” and “*ein Recht*.” I am not sure that there are any real or deep *conceptual* ties that go beyond the surface meaning of the German. But, however that may be, since these conceptual ties certainly do not exist in English, an English translation should not try to “preserve” them.

<sup>5</sup> This is most often argued with regard to the *Rechtslehre*, but it also concerns the *Tugendlehre*. Thus Christian Ritter, *Der Rechtsgedanke nach den frühen Quellen Kants* (Frankfurt am Main: Vittorio Klostermann, 1971) argued that Kant’s *Rechtslehre* remains essentially precritical and is not part of his transcendental philosophy. But see also Hariolf Oberer, “Zur Frühgeschichte der Kantischen Rechtslehre,” *Kant-Studien* 64 (1973), 88–102, and Werner Busch, *Die Entstehung der kritischen Rechtsphilosophie Kants, 1762–1780* (Berlin: Walter de Gruyter, 1979). Also relevant in this context are Josef Schmucker, *Die Ursprünge der Ethik Kants in seinen vorkritischen Schriften und Reflexionen* (Meisenheim am Glan: A. Hain, 1961), and Karl-Heinz Ilting, “Gibt es eine kritische Ethik und Rechtsphilosophie Kants?” *Archiv für Geschichte der Philosophie* 63 (1983), 325–45. But it is not always clear what the question amounts to, as Ilting, for instance, thinks there is no critical moral philosophy at all, and that even the *Groundwork* and the second *Critique* are non-critical, which seems to me absurd. But it appears to me that he makes a similar mistake when he tries to show that Kant’s categorical imperative depends upon the “principle of law,” as developed by Kant in his precritical work.

I have used the translations of the Cambridge Edition of the Works of Immanuel Kant (CE). In cases where I have found it necessary to change the translations, I have indicated the changes. Where there is no CE translation yet in print, translations are my own.



Still, there can be no doubt that the *Metaphysics of Morals* is the most important of these late works. There are at least three differences between it and the other late works. First, the others are much more closely related to the lectures on which they are based. The *Metaphysics of Morals* has more of a structure of its own. It follows a logic that is different from that of his lectures. While many of its particular doctrines can also be found in the lectures, the purpose is different. Second, it is a work that he had planned for a long time, not an idea that occurred to him relatively late. Third, it is a work that sums up one of the most important concerns of Kant's philosophical thinking as a whole, for morality or, as Kant would have put it, "the moral law within us," is concerned with what for Kant is most important about human beings; without morality we would be just like any other animal. So, the view that the *Metaphysics of Morals* should actually inform our view of Kant's ethics as it is expressed in the *Groundwork* and the second *Critique* and not the other way around is not entirely implausible either.<sup>6</sup>

## 2 THE PROTO-HISTORY OF THE *METAPHYSICS OF MORALS*

As has often been noted, however, the project of a "Metaphysics of Morals" goes back to the early days of Kant's philosophical development. He had clearly been concerned with it ever since observing in 1764 in the "Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality" that "even the fundamental concept of obligation" is inadequately understood and still in need of being properly analyzed, and that "the fundamental principles of morality in their present state are not capable of all the certainty necessary to produce conviction" (AA 2:298).<sup>7</sup> It became Kant's goal to change this unsatisfactory condition.

As early as December 31, 1765 he wrote to Lambert that he would publish "little essays" on the "Metaphysical First Principles of Natural Philosophy" and the "Metaphysical First Principles of Practical Philosophy" before going on to work "on the proper method of metaphysics and thereby also the proper method for philosophy as a whole."

<sup>6</sup> Allen Wood, "The Final Form of Kant's Practical Philosophy," in Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 1–21. See also the part of General Introduction in Kant, *Practical Philosophy* that bears the same title, i.e. xxx–xxxiii.

<sup>7</sup> All references in the text are to Immanuel Kant, *Gesammelte Schriften*, vols. I–XXII, Preussische Akademie der Wissenschaften (ed.); vol. XXIII, Deutsche Akademie der Wissenschaften zu Berlin (ed.); from vol. XXIV, Akademie der Wissenschaften zu Göttingen (ed.) (Berlin: Walter de Gruyter, 1907–). Since the CE includes the pagination of the AA, they can also be checked in the English translation.

Furthermore, “their materials [*Stoff*] lie ready” before him (Br. 10:56).<sup>8</sup> If they had seen the light of day, we would probably have heard much of sentiments and moral taste, and their relation to the theory of Hutcheson and other British writers. We know that these essays did not see the light of day.<sup>9</sup>

Still, it appears that Kant kept pursuing these subjects. Hamann wrote on February 16, 1767 that Kant was working on a “metaphysics of morals.” On May 9, 1768 Kant told his former student Herder that he had finally succeeded in his quest of knowing “the actual nature and the limits of human capacities and inclinations” at least as far as morality is concerned and that he was now working on a “Metaphysics of Morals,” in which he would present the “evident and fruitful principles of conduct and the method that must be employed” in it (HN 19:24). Within a year, if his health permitted, he would be done – or so he thought. Yet he was wrong again.

On September 2, 1770, after writing and defending the *Inaugural Dissertation*, he wrote to Lambert that in the coming winter he would be busy bringing order into his moral philosophy and completing his “investigations of pure moral philosophy, in which no empirical principles are to be found, as it were the Metaphysics of Morals” (Br. 10:97).

Three years later, in a letter to Marcus Herz written toward the end of 1773, he confided that he would be glad when he finished with his “transcendental philosophy, which is actually a critique of pure reason,” as he then could “turn to metaphysics,” which “has only two parts, the metaphysics of nature and the metaphysics of morals.” And he made the further claim that the metaphysics of morals would appear first. He also told Herz that he was very much looking forward to the metaphysics of morals (Br. 10:145). It was almost as if working on this project would be a relief from the critical business. Perhaps it would have been good, but it was not to be. Instead, he found it necessary to keep working on the *Critique of Pure Reason* for another seven years, and other works concerned with transcendental philosophy or the critique of pure reason.

In the *Critique of Pure Reason* itself, he distinguishes between the “speculative and the practical use of pure reason,” and based on this between metaphysics of nature and metaphysics of morals (KrV A841/B869). The metaphysics of morals is to contain all the principles that

<sup>8</sup> I have translated “*Anfangsgründe*” in the two projected works and the three works that appeared consistently as “First Principles” to make clear that the titles are indeed very similar.

<sup>9</sup> The “Remarks in the Observations on the Beautiful and Sublime” (AA 20:3–192) may very well contain the materials for these essays.

“determine action and omission a priori and make them necessary.” It will be “pure morality, which is not grounded on any anthropology (no empirical condition)” (KrV A841/B869). He contrasts metaphysics in general very sharply with critique, which is merely propaedeutic or preparatory, and not really “the system of pure reason.” What he has in mind for the metaphysics of nature at this point probably also holds for the metaphysics of morals: it “will be not half so extensive but will be incomparably richer than this critique, which had first to display the sources and conditions of its possibility” (KrV Axxi), even though he “set[s] it aside” in the *Critique of Pure Reason* because it does not really pertain to its end at this stage (KrV A842/B870). This claim should not be confused with another claim that he makes in the Introduction, where he argues that “the supreme principle of morality and the fundamental concepts of it ... do not belong to transcendental philosophy, since the concept of pleasure ... of desires and inclinations, of choice, etc. which are all of empirical origin, must there be presupposed” (KrV A14/B29). He claims here that the fundamental concepts and principles of morality cannot form part of the propaedeutic or preparatory project or the critical project. In other words, he claims that there will be no *Critique of Pure Practical Reason*.<sup>10</sup>

Kant’s *Groundwork* clearly was not meant as the preparation of a second *Critique*, but rather as an antecedent of the *Metaphysics of Morals*. It was to establish “a completely isolated metaphysics of morals, mixed with no anthropology, theology, physics, or hyperphysics, and still less with occult qualities (which would be called hypophysical)” (G 4:410). In such a metaphysics of morals – Kant insists on using the name, even though he thinks it is “decried” – “moral principles are not based on what is peculiar to human nature but must be fixed a priori by themselves, while from such principles to must be possible to derive practical rules for every rational nature, and accordingly for human nature as well” (G 4:410n), and when he enumerates “a few duties in accordance with the usual division of them into duties to ourselves and to other human beings and into perfect and imperfect duties,” he notes that he reserves “the division of duties entirely for a future *Metaphysics of Morals*,” and that the division he gives in the present context is merely one of convenience (G 4:421). And it is clear that at this point Kant still does not envisage that a *Critique of Pure Practical Reason* would need to be written and that all the preparatory work for the

<sup>10</sup> See also Lewis White Beck, *A Commentary on Kant’s Critique of Practical Reason* (University of Chicago Press, 1960), 9. Beck argues that Kant widened his conception of transcendental philosophy. One might, however, also say that he narrowed his conception of the basic concepts and principles of moral philosophy by eliminating any references to desires, pleasures, etc.

*Metaphysics of Morals* has been done. Indeed, some passages suggest that the step “into metaphysics of morals” is already taken in some parts of the *Groundwork* (G 4:427).

But be that as it may, it is clear that Kant wanted to proceed to the “complete elaboration” of the metaphysics of morals immediately after finishing his *Metaphysical Foundations of Natural Science* in 1785 (Br. 10:407).<sup>11</sup> Yet, it is just as clear that he did not succeed in doing so. Other, more pressing, matters interfered. In the very letter in which he said that he would immediately undertake the completion of the metaphysics of morals, he also agrees to write a review of the second part of Herder’s *Ideen*. There was also the second edition of the first *Critique* (1787). But more importantly, he came to think that he needed to write a *Critique of Practical Reason* (1788).

None of this means that Kant had given up on the *Metaphysics of Morals*. In the second *Critique* itself, he held out hope for “the system of science” that would go beyond the “system of critique” (KpV 5:8). Thus, in a letter to Jung-Stilling of April 7, 1789, he promised that “around the end of the summer I shall begin to work on my ‘Metaphysics of Morals,’” which also means that at this time he still had actually not begun work – or so I would think (HN 23:495). On May 26, 1789, he complained that his health was becoming “progressively worse” in his sixty-sixth year, and of the burden that “the last part of the Critique, namely, that of judgment” constitutes while he is also “working out a *system of metaphysics*, of nature as well of morals, in conformity with those critical demands” (Br. 11:49).

Friends and acquaintances expected the work to appear as early as Easter of 1791. On December 21, 1792 he reported to Erhard of working on the *Metaphysics of Morals*, referring to his discussion under the “heading of *Duties to Oneself*,” which he said he was treating “in a manner quite different from what is customary” (Br. 11:309). What is interesting is that, even though the letter goes on to discuss the social contract and natural law, Kant makes no reference to any discussion of such matters in the book he is working on. This is consistent with the fact that Fichte, who had visited Kant in 1792 and written his *Critique of All Revelation* in Königsberg, referred in letters to Kant to a “*Metaphysics of Morals*.” In a letter of May 12, 1793, Kant tells Fichte that, if he were not already seventy years old and held back by his age, he would probably already be dealing with the problem of revelation in the “planned *Metaphysics of*

<sup>11</sup> Zweig translates “völlige Ausarbeitung” as “full composition” (CE *Correspondence*, 229).

Morals” (Br. II:434). It looks very much as if Kant’s project at this time did not include the *Rechtslehre*, or, at the very least, that he had not begun working on it.

We do not know when Kant actually began the writing of the *Metaphysics of Morals* in earnest. It was probably sometime around 1795. In this context it is of some interest that Kant offered his usual lecture course on metaphysics in the winter semester of 1793–94 for the first and only time under the title of “Metaphysics of Morals or Universal Practical Philosophy in accordance with Baumgarten.”<sup>12</sup> While this new title does not necessarily mean that he was at this point already actively writing the two parts that now make up the *Metaphysics of Morals*, it does show that he more intensively dealt with the matters that make up this work.

It appears to me that we may divide Kant’s deferment of the proposed *Metaphysics of Morals* into three periods, with the first one dating from 1762 to about 1770, the second one from 1770 to 1785, and the third period from 1785 to 1797. In each period there are the pressures that come from the perceived need to finish the most important part of his transcendental philosophy, but also other kinds of interferences. The work on the proper method of metaphysics turned out to be much more time-consuming and demanding than Kant anticipated in 1765. It would lead to the *Inaugural Dissertation, De mundi sensibilis atque intelligibilis forma et principiis* of 1770, the *Critique of Pure Reason* of 1781, and the *Prolegomena* of 1783. And after he had written the *Metaphysical First Principles of Natural Science*, it became unexpectedly necessary to engage in further preparatory critical investigations in moral philosophy, namely the *Groundwork of the Metaphysics of Morals* of 1785 and the *Critique of Practical Reason* of 1788. Other work on theoretical and other matters resulting from the first *Critique* also interfered. And then there was the *Critique of the Power of Judgment* of 1790 and the work connected with it, not to speak of the *Religion within Reason Alone* that made his life difficult after 1794. Between 1785 and 1793 the “complete elaboration of the metaphysics of morals” was prevented by other works that were made more urgent by developments having to do with the reception of his thought, and by reasons of health caused by his advancing age.

<sup>12</sup> Actually: “Metaphysicam morum, sive Philosophiam practicam universalem, una cum Ethica ad compendia Baumgartiana.” The usual title would be something like “Metaphysicam, duce Baumgarteni” or “Metaphysicam praeunte Baumgarten explicat.” See Michael Oberhausen and Riccardo Pozzo (eds.), *Vorlesungsverzeichnisse der Universität Königsberg (1720–1804)*, 2 vols. (Stuttgart-Bad Cannstatt: frommann-holzboog, 1999).

In 1794 the so-called censorship edict made the less than ideal working conditions even more difficult. Kant had to promise that he would not publish on religious topics again, but any *Metaphysics of Morals* had to include a discussion of the relations between morality and religion. Thus he wrote on November 24, 1794 to a publisher that he could not promise a firm date for delivering any specific work, and explained that anything he might write would be problematic in any case, since his subject was “really metaphysics in the widest sense, and as such includes theology, morality (and thus also religion) as well as natural law (including public law [*Staatsrecht*] and international law [*Völkerrecht*]), though only to the extent that reason can address these subjects, but the hand of the censor lies heavily on all of these topics and one cannot be sure that all one’s work in any of those fields will not be rendered futile by a stroke of the censor’s pen” (Br. 11:531). He was hoping for peace and clear new rules about what was and was not admissible in publishing about religious and political matters.<sup>13</sup>

There is every reason to suppose that the final work on morals that appeared in 1797 is as different from the projected work of 1765 as the *Metaphysical First Principles of Natural Science* of 1786 is different from what he thought it would be in 1765.<sup>14</sup> The “system of metaphysics ... of morals, in conformity with ... critical demands” had to be very different from what Kant thought it would be in 1765 or 1770. On the other hand, there can also be little doubt that the *Metaphysics of Morals* was the fulfillment of the early hopes and the intermediate promises. We might, therefore, ask what, if anything, remained the same, and thus would justify the claims of some scholars that the practical philosophy is non- or precritical, and what precisely it was that that changed, and whether it goes to justify the claims of other scholars that it presents the final form of Kant’s critical moral philosophy. I think that the answer to both questions is a qualified yes. What remained the same is the actual content of Kant’s *Rechts-* and *Tugendlehre*. What changed was the perspective from which this content must be viewed according to Kant. This gave rise to tensions in Kant’s system, but they are not tensions that are fatal to Kant’s view, at least not when the content and form of his moral philosophy are properly understood.

<sup>13</sup> Kant is, of course, referring to the French Revolutionary Wars and the First Coalition, which was defeated in 1797. However, the Prussian army took part in the war only between 1792 and 1795. So, “peace” was indeed near.

<sup>14</sup> This has not kept scholars from arguing that there is a serious disanalogy.

3 THE RELATION OF THE *METAPHYSICS OF MORALS*  
TO KANT'S LECTURES

Though Kant found he had to postpone the completion of this project again and again, he clearly thought about it during all that time, if only because he was regularly teaching moral subjects. The lectures that were most relevant for this were those on natural law, on ethics, and on anthropology. While the lectures on metaphysics are important as well, they are less so. Kant lectured on moral philosophy twenty-six times between 1755 and 1793/94, twelve times during his so-called precritical period, and fourteen times during his critical period (1770–1804).<sup>15</sup> While the titles of these lectures varied widely, he seems to have always used the same textbooks, namely Alexander Gottlieb Baumgarten's *Initia philosophiae practicae primae acroamaticae* (Halle 1760) for the first part of the course, and his *Ethica philosophica* (Halle 1751) for the second part.<sup>16</sup>

Kant did not slavishly follow these textbooks in his lectures. Rather, he used them in a fashion similar to the way others had used commonplace books since the sixteenth century. His textbook served mainly as an outline of the discipline to be taught. It suggested the order in which the materials would be discussed in class. Perhaps more importantly, it also provided the “heads of inquiry” of the science in question. That Kant used them in this way is shown by the fact that he employed interleaved copies of textbooks and used both the margins and the separate pages for his reflections on the topics they covered. Accordingly, they were tools both in the preparation of his lectures and in his thinking about the subject matter at hand. Even the briefest look at the Table of Contents of Baumgarten's textbooks shows that they were eminently suited for such purposes. Furthermore, given Baumgarten's laconic style, he did not get as much in the way of independent reflections on these concepts as some other authors would have. This shows why Kant preferred his textbooks to those of more prolix authors.<sup>17</sup>

He was almost inevitably influenced by the textbooks in his thinking about moral matters. Thus Baumgarten's broad division of duties towards

<sup>15</sup> See Emil Arnoldt, *Gesammelte Schriften*, 6 vols., Otto Schöndörffer (ed.) (Berlin: Bruno Cassirer, 1907–09), vol. v, 338, who claims “with certainty” that Kant lectured 28 times on moral philosophy “under various titles,” but had intended to lecture at least 30 times.

<sup>16</sup> The *Initia* is reprinted in AA 19:7–21; the *Ethica philosophica* can be found at AA 27:737–1028.

<sup>17</sup> Karl Vorländer, in his Introduction to Immanuel Kant, *Metaphysik der Sitten*, 2nd edn. (Leipzig: Verlag der Dürr'schen Buchhandlung, 1907), ix–li, xxiii, 28 says that he used it as the “skeleton” for his lectures and notes. Arnoldt, to whom he refers, speaks of the “*Fachwerk der Einteilung*” or the “pigeon-holes of divisions” in this context.

God, towards oneself, and towards others in the *Ethica* had a lasting effect on Kant. Even though he relatively early rejected the idea that we have any duties towards God, the idea that we have duties to ourselves or that there are *officia erga te ipsum* and duties towards others or *officia erga alia* still informs the Doctrine of the Elements of Ethics in the *Metaphysics of Morals*. Even the subdivisions of these two broad divisions are clearly indebted to the *Ethica*. That we must differentiate between duties we have towards ourselves as animal beings as well as moral beings or rational selves is just as much indebted to Baumgarten as are many of the discussions of particular duties that are discussed in this context, like suicide (§252), misuse of one's sexuality or *crimina carnis* (§272–75), lying or *falsiloquium* (§343), avarice (§287), etc.<sup>18</sup>

It is possible to show which parts of the lecture notes taken by Kant's students are about which parts of the *Ethica*. Werner Stark has done this for one of the lectures from the 1770s, for instance, but it would be desirable for all the lectures.<sup>19</sup> And it is also possible to show that there is a relatively large degree of similarity between these notes from different times during the 1770s, 1780s, and 1790s. To be sure, different matters are emphasized at different times and new ideas are introduced, but there is no radical change in the parts of the lectures that concern Baumgarten's *Ethica*. Put differently, the contents of the sections that correspond to what later became the Doctrine of Elements remains relatively static.

The passages of the lectures that correspond to Kant's discussion of the *Initia* are very different. They exhibit large, and often radical differences over time. Thus, in the Herder lectures from the beginning of the 1760s, the discussion of the moral sense or moral feeling is absolutely central. Kant claims that "the moral feeling is unanalyzable, a basic feeling [*Grundgefühl*], the basis of conscience." While he already argues that in morality we only ask "for the formality of what is perfect in free actions," it is "only the consideration of free actions with moral feeling that is intrinsic to conscience" (H 27:5). "My reason can err, my moral feeling only if I mistake what is habitual for what is natural ... my ultimate criterion remains moral feeling" (H 27:6). In the Collins lecture notes from the middle of the 1770s, in which Kant's recent discovery of the

<sup>18</sup> There are also interesting differences. Avarice and lying belong according to Kant among the duties towards oneself, whereas Baumgarten discusses them in the context of duties towards others, but the similarities cannot be overlooked either. Duties towards others are characterized by Baumgarten by the notion of love, while Kant argues they consist in making others happy.

<sup>19</sup> Immanuel Kant, *Vorlesungen zur Moralphilosophie*, Werner Stark (ed.), with Introduction by Manfred Kuehn (Berlin: Walter de Gruyter, 2004), 419–28.



importance of anthropology for the application of moral principles is very apparent, Kant argues by contrast that “there must be a single principle emanating from the ground of our will” (C 27:253).<sup>20</sup> And the principle is already identified with “the moral imperative” that “expresses the goodness of the action in and for itself” and “shows that moral necessitation is categorical and not hypothetical. Moral necessity consists in the absolute goodness of free actions and that is *bonitas moralis*” (C 27:255–6). In the Mrongovius lecture notes from about 1785, we find a discussion of the categorical imperative as “the rule of a will [that is] intrinsically good” (M II 29:607) and in the lectures that came later, we do get essentially the same view as is found in the *Groundwork*.

These shifts in the meaning of the basic principle of morality also have *some* influence on the way the specific virtues are discussed in the lecture notes, but it does not go as far as one might expect. The particular duties and virtues Kant espouses do not really change as a result of the radical changes in his view of the basic principle of morality. The contents of morality, or the actual duties that human beings are said to have, remain more or less the same. What changes is Kant’s critical discussion of them, and this is what makes some of the lectures part of the critical corpus.

Similar things are probably true of the lectures on jurisprudence. The lectures on natural law, always listed as “*jus naturae*,” were always based on Gottfried Achenwall’s *Jus naturae in usum auditorum* (first published in 1763). The only set of lecture notes that has survived follows the outline of Achenwall’s compendium fairly closely. It is from the winter semester of 1785. While Kant seems to have followed the textbook closely, he does often criticize the author on specific points. One thing that makes these lecture notes interesting is that he rejects Achenwall’s definition of law as independent of morality and consistently argues against the consequentialism and eudaimonism present in his work. Again, it is the Introduction that is very interesting because it contains an extensive discussion of the human will, freedom, the distinction between hypothetical and the categorical imperative, which is also called the “unconditional imperative of wisdom” (F 27:1324), the distinction between actions in accordance with duty and from duty, and clearly shows that Kant discussed in this lecture some of the matters that are most characteristic of the *Groundwork*. But some of the basic distinctions of the *Metaphysics of Morals* are already

<sup>20</sup> For his view on morality see C 27:244. See also Manfred Kuehn, Introduction, in Immanuel Kant, *Anthropology from a Pragmatic Point of View*, Robert B. Louden (ed.) (Cambridge University Press, 2006), vii–xxix.

present. Thus we find that “Ethics is practical philosophy of actions in regard to dispositions [*Gesinnungen*]. *Jus* is practical philosophy of actions without regard to disposition. All that possesses obligation belongs to ethics, thus all duties. The law concerns duties and actions that are in accordance with the law and can be coerced. An action is called right when it agrees with the law, virtuous when it is based on respect for the law” (F 27:1327).

This set of lecture notes seems to confirm the conclusion we have reached with regard to the lectures on morals, while the particular legal subject matter is largely dependent on the textbook author (and may therefore be called precritical), the way this content is embedded in a discussion of the first principles of practical philosophy is at the very least on the way to critical philosophy.<sup>21</sup>

It is very tempting to see the discussions of Baumgarten’s *Initia* in Kant’s lectures as closely connected with Kant’s critical project. Put differently, the *Groundwork* and the second *Critique* come out of these discussions. They are in some ways just as closely related to his discussion of the first principles of practical philosophy as the substantive parts of the *Metaphysics of Morals*, the Doctrine of Law and the Doctrine of the Elements of Ethics are much more closely tied to Achenwall’s *Jus naturae* and Baumgarten’s *Ethica*. However, there is clearly one fundamental difference: while the doctrinal parts of the *Metaphysics of Morals* did not radically depart from the textbooks, at least insofar as the content is concerned, the *Groundwork* and the second *Critique* have rather little in common with what Baumgarten taught in the *Initia*.

If this is correct, then we may agree with those German critics who have claimed that Kant’s moral and legal philosophy remained ultimately precritical at least insofar as we admit that the contents of the Doctrine of Law and the Doctrine of the Elements of Ethics are not essentially new with Kant. These parts are not what is characteristic about Kant’s critical ethics. On the other hand, this would also mean that the Introduction to the *Metaphysics of Morals* (MS 6:211–21) as well as the Preface and Introduction to the *Doctrine of Virtue* (MS 6:375–413), which are attempts to make clear how the substantive parts of this work fit in with the critical project as a whole, more or less unequivocally belong to the critical project.<sup>22</sup>

<sup>21</sup> These results could be further supported by a careful discussion of Kant’s reflections on moral and legal subjects, but this is not possible in this context.

<sup>22</sup> I do not want to claim, of course, that the other parts of the *Metaphysics of Morals* are entirely uncritical. It is just that they are much more indebted to the textbooks Kant used.

We may ask what, if any, consequences this has for the question as to whether the *Metaphysics of Morals* is the final form of Kant's practical philosophy. It appears to me that the answer to this question needs to be equally measured. Historically speaking, it is just true that it is the final form Kant gave to his moral philosophy. It is also true that the development of a *Metaphysics of Morals* was Kant's ultimate goal throughout most of his philosophical life. But it is far from clear that what Kant ultimately produced is representative of his best intentions and fits unproblematically with his critical moral philosophy as developed in the *Groundwork* and second *Critique*. I think we need to be careful especially when we evaluate its substantial moral doctrines, such as his views on servants (MS 6:283) or "on defiling oneself by lust" (MS 6:424). But be that as it may, it is clearly more important in this context to take a closer look at how this work is related to the *Groundwork* and the *Critique*.

#### 4 WHAT MAKES THE METAPHYSICS OF MORALS METAPHYSICAL?

In the *Critique of Pure Reason*, Kant defines "metaphysics" as the "system of pure a priori knowledge" which "exhibits in systematic fashion the whole body of philosophical knowledge arising out of pure reason" (KrV A841/B869), and divides it into the "speculative" and the "practical employment of pure reason" or the metaphysics of nature and the metaphysics of morals. By "critique" he understood investigations propaedeutic to both divisions of metaphysics (KrV Axxi). The metaphysics of morals is to contain *all* the principles which "determine action and commission a priori and make them necessary." It will be "pure morality, which is not grounded on any anthropology (or any empirical conditions)" (KrV A841/B869).

In the *Critique of Practical Reason*, however, he declines to give a complete classification of all practical sciences just because "the special determination of duties as human duties, with a view of classifying them, is possible only after the subject of this determination (the human being) is cognized as he is really constituted" (KpV 5:8). The *Critique of Practical Reason*, however, gives an account of the possibility of the principles of practical philosophy "without special reference to human nature" (KpV 5:8). He makes essentially the same claims in the *Groundwork*, in which he radically differentiates moral philosophy from anthropology, claiming that the metaphysics of morals must precede practical anthropology, must be completely a priori, and must therefore be "purified"

or “cleansed” of anything empirical, a posteriori, or belonging to mere anthropology (G 4:388).

As a result, “Virtue” or “*Tugend*” does not play a significant role in Kant’s *Groundwork* and the *Critique of Practical Reason*. Virtue has a special reference to human nature. Indeed, it is the greatest achievement for a specific kind of reason, namely human reason. Kant makes this very clear in the early lectures on moral philosophy, where he argues that any kind of doctrine of virtue cannot capture moral philosophy as a whole because “virtue entails not just *morally good* actions, but at the same time the possibility of the opposite, and thus incorporates an inner struggle, this is therefore too narrow a concept, since we can also ascribe *ethics*, but not virtue (properly speaking) to the angels and to god, for in them there is assuredly holiness but not virtue” (H 27:13). As Kant also says: constant progress toward virtue is “the utmost that finite practical reason can effect” and “virtue itself ... as a naturally acquired ability can never be completed” (KpV 5:33).<sup>23</sup>

For this reason, it is not a topic for pure philosophy, but only one for pure moral philosophy *applied* to human beings. This is why it must play the most central role in the *Metaphysics of Morals*, where it is defined as the strength of resisting “what opposes the moral disposition *within us*” (MS 6:380). What opposes the moral disposition within us are our “sensible inclinations,” which cannot be effectively opposed by other sensible inclinations according to Kant, and therefore require “a moral end ... that must therefore be given a priori, independently of inclinations” (MS 6:381). And it is for this reason, the argument quickly proceeds, that “ethics can also be defined as the system of the ends of pure practical reason” (ZcF 8:381). Indeed, these are ends that are themselves duties.

It is probably unnecessary (and certainly impossible) to enter here into a complete discussion of Kant’s conception of ends that are also duties and the role it plays in the *Metaphysics of Morals*. But it is important to understand that ends that are also duties are absolutely central in Kant’s account of morality. For they are what supposedly gives rise to the two fundamental duties human beings are said to have, namely the duties towards ourselves and duties to others as the ends of one’s own perfection and the happiness of others (MS 6:385). All other particular duties, such as those that he found in Baumgarten’s *Ethica* and other textbooks on moral philosophy, are species of these two basic genera.

<sup>23</sup> In the *Grundlegung* the word appears only ten times. In the second *Critique* Kant addresses the question of the relation of happiness and virtue.

What is even more important to understand is how these ends that are also duties are related to the categorical imperative, and more particularly, which of these is more basic. Given the way that Kant is usually read, i.e. from the point of view of the *Groundwork*, we might suppose that only the categorical imperative can establish the duty that we need to perfect ourselves and that we need to make others happy. And are these two duties not used as the third and fourth examples concerning imperfect duties in the *Groundwork* itself?<sup>24</sup> And does not Kant say there that these are some “actual duties, whose derivation from the one principle cited above [i.e. the categorical imperative] is clear” (G 4:424)?

Well, actually Kant does not say that. What he does say is: “These are a few of the many actual duties, or at least of what we take to be such, whose *division* is clear from the one principle cited above” (G 4:424). The translation is based on an emendation by Hartenstein, who without any argument whatsoever substituted “*Ableitung*” for “*Abteilung*” in his 1838 edition of the *Groundwork*, and the majority of editors have followed him without any good reasons. “*Abteilung aus*” would indeed be an odd expression, but Kant is not using this expression. The sentence must be parsed differently. Kant is actually saying only that the *Abteilung* is clear “from” the principle given before or “*aus dem einigen angegebenen Prinzip.*” Kant introduces the examples in accordance with the “usual division,” says that the real division will follow in the *Metaphysics of Morals*, and then makes a point about what the categorical imperative shows about the “usual division,” namely that there are different kinds of contradictions when perfect and imperfect duties are considered. If this were not enough, he picks up the topic of the “division” of duties in the Introduction of the *Doctrine of Virtue*, claiming that “all the *divisions* of ethics will only have to do with duties of virtue” (MS 6:410).

The emendation is also at odds with the *Metaphysics of Morals*, where Kant finds that the categorical imperative, in which “I abstract from all ends,” shows only that maxims qualify for a possible universal law, but does not actually show that these maxims are universal laws. It is “only a negative principle (not to come into conflict with a law as such),” and there needs to be a law for the maxims of actions that goes “beyond this principle” (MS 6:389). And this is, according to Kant, “only the concept of an *end* that is also a duty, a concept that belongs exclusively

<sup>24</sup> It should perhaps also be noted that the other two examples are actually examples of legal duties. It was (and is) against the law to break contracts, and it was (and in some places still is) illegal to commit suicide.

to ethics, establishes a law for maxims of actions by subordinating the subjective end (that everyone has) to the objective end (that everyone ought to make his end)” (MS 6:389).

If we take this seriously, and it is difficult to take it seriously given the weight of the traditional reading, the categorical imperative alone is not sufficient for moral action. We need also ends that are duties.

However, I would argue that there is less here than meets the eye. For it might be argued that Kant’s discussion of the categorical imperative in the *Groundwork* in some sense already contains this notion of ends that are duties, for at least one of the three applications of the categorical imperative, which are based upon the categorical imperative as the formula of universalization, namely the formula of the “practical imperative,” does already contain “ends” in the required sense of “ends as duties” (G 4:430). This should not be too surprising, since it is, after all, also called the “principle of humanity” (G 4:431). He also speaks of this as the “matter [of all maxims], namely as an end,” and finds “that in this respect the formula says that a rational being, as an end by its nature and hence as an end in itself, must in every maxim serve as the limiting condition of all relative and arbitrary ends” (G 4:436). This matter is in the *Groundwork* determined only negatively as well.

The concept of ends that are also duties, as developed in the *Metaphysics of Morals*, is meant to extend this merely limiting condition and to transform it into a positive law. “Hence, if there is an end that is also a duty, the only condition that maxims of actions, as means to ends, must contain is that of qualifying for a possible giving of universal law” (MS 6:389).

Wood, in considering these very same matters, has argued that Kant “overwhelmingly prefers the Formula of Humanity as the formula in terms of which the moral law is to be applied.”<sup>25</sup> But it appears to me that he is seriously misconstruing the situation. It is not that Kant justifies “no fewer than nine of the sixteen ethical duties by means of the formula of humanity,” but rather that he argues these duties are duties that are at the same time ends. It is true that such arguments have a more or less “natural” connection with the formula of humanity, but they cannot be reduced to an application of it.

Indeed, the categorical imperative and its subformula play a rather subdued role in the entire book. One might get the impression that it is not really needed in the derivation of duties that are also essential ends of embodied rational beings. And one might ask why this is so. In the very

<sup>25</sup> Wood, “The Final Form,” 13.

first footnote of the *Doctrine of Virtue* (MS 6:375n) Kant makes a point about the distinction between someone who is merely versed in practical philosophy and a practical philosopher that may easily appear gratuitous. He claims that “someone *versed in practical philosophy* is not for that reason a *practical philosopher* already. A practical philosopher is one who makes *the final end of reason* the principle of his actions and joins with this such knowledge as is necessary for this.” To know *what* it is one’s duty to do is one thing. To transform the duty into “the inner principle of the will” in accordance with the knowledge necessary for it is quite another. And Kant suggests that the first kind of knowledge is “easily stated” because it has to do with “the ends all human beings have by their nature.” Even someone merely acquainted with moral philosophy can know what should be done. The second kind of achievement is more important. Only someone who has attained it deserves the title of “moral philosopher.” The “inner principle of the will” is the realization that “the consciousness of this duty is also the *incentive* to actions” (MS 6:375n).

This way of dividing up the work of morality suggests that the determination of a “mere duty of virtue” does not require mental acrobatics. We all know already what such duties are, at least insofar as we know who we are and what our essential ends consist in. There is no significant philosophical or moral problem here, and Kant contrasts this apparent ease of the determination of what our particular duties are with the deeper and more important problem, namely the one that has to do with understanding “the inner principle of the will.” And only someone who has joined these two things is truly a practical philosopher, or so Kant says.

If we call the first requirement “the principle of the individuation of particular duties” and the second one “the principle of the inner nature of dutiful action,” then Kant says just about as clearly as one might wish that we individuate duties by our essential ends. And I would submit that is what he actually does. He does not explicitly say what the more important second requirement or the “the principle of the inner nature of dutiful action” amounts to, but it is obvious to me that it is another way of stating the categorical imperative. It is the “supreme principle of morality” of the *Groundwork* or the “principle of the will” that makes for the moral worth of the action (G 4:400).

That this reading is plausible can also be shown by the distinction Kant makes between *Tugendverpflichtung* or *obligatio ethica*, on the one hand, and *Tugendpflichten* or *officium ethicum s. virtutis*, on the other. The first term is difficult to translate, and it is rendered in the Cambridge edition as “obligation of virtue,” while the other one is translated as “duties of

virtue.” Perhaps “ethical obligation” or “obligation to be virtuous” would be marginally better. In any case, Kant claims that we have many duties of virtue. Indeed, there are just as many duties of virtue as there are ends that are also duties. But there is only one *Tugendverpflichtung* or obligation to be virtuous, and this obligation precedes any particular duty that we may have. It also precedes any conception of ends and is

the virtuous disposition [*Gesinnung*], the subjective determining ground to fulfill one’s duty, which extends to duties of right as well although they cannot, because of this, be called duties of virtue. – Hence all the *divisions* of ethics have to do only with duties of virtue. Viewed in terms of its formal principle, ethics is the science of how one is under obligation without any regard for any possible external lawgiving. (MS 6:411)

To say that the categorical imperative as the supreme principle of morality has only to do with the “inner principle of the will” or with the principle of dutiful action amounts to saying that the categorical imperative has to do only with *Tugendverpflichtung* and does not go very far in telling us what our actual duties are. It explains what it means to act or will morally or dutifully, leaving unaddressed the question as to what our duties actually are, because that is not really perceived as a problem by Kant. And that is what one would expect from a critical discussion of morality that is “without special reference to human nature” (KpV 5:8), which is what the *Critique of Practical Reason* and the first two parts of the *Groundwork* are purported to be.

## 5 CONCLUSION

There is, of course, a problem that looms large in all of this, and this is the question of Kant’s large and largely essentialist claims about the a priori “system of the ends of pure practical reason.” What justifies us in assuming these? Kant introduces this idea far too quickly and does not sufficiently justify them. It is clear that they are supposed to be very different from anthropological claims that would be based on empirical observations. They are meant to be philosophical or a priori anthropology, which is another subject that Kant hoped to develop, but never did. It would have dealt with the a priori constituents of human nature as necessary for morality and filled the space between pure moral philosophy and empirical anthropology. Whether such a discipline would have been possible in accordance with the principles Kant established in the first *Critique* is highly doubtful.



In other words, I agree with Wood that “it is a mistake to think that rights and juridical duties for Kant rest on the moral imperative ... or that Kantian ethics has no place for ends or virtues.”<sup>26</sup> But this is a very mixed blessing. The unabashed essentialism about human nature present in the *Metaphysics of Morals* may not be enough to qualify “the final form of Kant’s practical philosophy” as precritical, but it shows that the historical Kant cannot have an answer to Nietzsche and others on their own terms. It would make no sense reflectively (or otherwise) to endorse essential ends of human nature or pure practical reason that are also duties, for instance. But then again, such essential ends of pure practical reason themselves may make no sense either.

<sup>26</sup> Wood, “The Final Form,” 20.

## CHAPTER 2

# *Reason, desire, and the will*

*Stephen Engstrom*

Much attention has been devoted to Kant's famous doctrine of autonomy, the proposition that morality finds its source in the will's self-legislation, depending neither for the content of its principle nor for its motivating power on any source, natural or transcendent, outside the will and its power of self-rule. But Kant also advances another striking proposition about the will, that it is nothing but practical reason. Though less extensively investigated, this idea is at least as important, both in its own right and for the light it throws on other parts of his ethics, including his doctrine of autonomy, which can seem unduly voluntaristic if not appreciated in its practical-cognitivist setting. According to tradition, the will is rational desire. Kant too understands the will in terms of reason and desire, but his way of combining these notions in his conception of a practical application of reason accounts for much of what is distinctive in his moral philosophy.

This chapter examines Kant's mature conception of the will, as presented in the Introduction to the *Metaphysics of Morals*. Kant approaches this conception from a definition of the faculty of desire. But before doing that or indeed anything else, he makes a few remarks about the system of philosophical rational knowledge within which the metaphysics of morals is situated.<sup>1</sup> Though somewhat fragmentary, these remarks recall the account he offered at the outset of the *Groundwork of the Metaphysics of Morals* – the work that, as its title announces, lays the ground for the metaphysics of morals. Before we turn to the faculty of desire, therefore, we should consider the conception of rational cognition that the metaphysics of morals presupposes.

<sup>1</sup> If recent textual scholarship is correct. According to Bernd Ludwig, the traditionally accepted order of the first two sections of the Introduction to the *Metaphysics of Morals* is the result of a typesetter's error and should be reversed. See his Introduction in Kant, *Metaphysische Anfangsgründe der Rechtslehre*, B. Ludwig (ed.) (Hamburg: Felix Meiner Verlag, 1986, 1998).

## I REASON

1.1 Kant begins the Preface to the *Groundwork* by marking two distinctions that partition philosophy into three sciences. Philosophy's rational knowledge divides into formal knowledge, comprising the a priori science of logic, and material, which concerns objects and the laws to which they are subject. Material philosophy divides into natural and moral, according as the laws it studies are laws of nature or of freedom. Laws of nature – laws “according to which everything happens” – determine how the objects they govern operate; laws of freedom – laws “according to which everything ought to happen” – determine for the things they govern how they are to act. Kant also draws a third distinction, within each of these material sciences, between an a priori, or metaphysical, part and an empirical part, thus marking off two a priori material sciences, a metaphysics of nature and a metaphysics of morals.

Though of Hellenistic provenance, Kant's taxonomy exhibits the principal differences separating his moral philosophy from the standard rationalist and empiricist approaches that have emerged from the Stoic and Epicurean traditions. Its identification of moral philosophy as a distinct form of material rational knowledge involves an implicit denial that ethics is subordinate to the theoretical knowledge of nature. And its identification of an a priori part of moral philosophy entails a rejection of empiricism. That there must be a pure moral philosophy, Kant argues, “is clear of itself from the common idea of duty and of moral laws,” by which these laws are recognized as having absolute necessity and strict universality, the hallmarks of a priori or rational knowledge (G 4:389).<sup>2</sup>

Of particular significance for present purposes, however, is that Kant's presentation of this taxonomy shows him to be taking for granted already at his point of entry into practical philosophy a conception of philosophy as a system of rational knowledge. His distinctions are divisions; they respect the unity of philosophy and of reason. Later in the Preface, with an eye to the second of the three divisions, Kant says it must be possible to show practical reason's unity with speculative, or theoretical, “for in the end there can be only one and the same reason, which is distinguished only in the application” (G 4:391).

If the difference between natural and moral philosophy, or, as Kant also expresses it, between theoretical and practical knowledge, is a difference

<sup>2</sup> My translations of passages from Kant's writings follow for the most part the Cambridge Edition of the Works of Immanuel Kant.

in the application of the same reason, then it must be possible to articulate the conception of this common reason and to describe how these applications differ. Let us take up these questions in order.

1.2 A starting point for reflection about the conception of reason is suggested by Kant's distinction between formal and material philosophy. Since logic deals "merely with the form of understanding and reason itself and the universal rules of thinking in general, without distinction of objects" (G 4:387), its formal account of reason is essential to an understanding of reason's use in both practical and theoretical applications.

Kant's broadly Aristotelian logic expounds the form of thinking in general by offering a formal account of the acts of the discursive intellect, treating first of concepts, then of judgments, and finally of conclusions. Kant frames his conception of reason by reference to the last of these three logical acts, in which a judgment is derived from another, or thought as necessary on account of its relation to the latter. In particular, reason is the faculty responsible for the mediately derived conclusion that constitutes a syllogism (*Vernunftschluß*: "conclusion of reason") (LJ 9:114, 120–21). In the principal case of the categorical syllogism, the conclusion is a judgment reached through the subordination (in the minor premise) of its subject concept under a universal judgment (the major premise), which serves as a rule or principle determining the attachment of the predicate (in the conclusion).

This logical conception of reason clearly underlies Kant's understanding of reason's theoretical and practical application. Reason is depicted as exhibiting the form of the syllogism both in deriving effects (what "happens") from laws of nature and in deriving actions (what "ought to happen") from laws of freedom (KrV A298–305/B355–61, A645–47/B673–75; G 4:412).

Reason so conceived is not merely a capacity to think consistently, or to calculate, or to infer one thing from another, nor an ability to figure out what to believe or to do, or to recognize the reasons one has to believe this or to do that. It is the capacity to know through a derivation of the form just noted. Knowledge gained through reason lies in a judgment conscious of its own necessity through its subordination to a universal cognition, or a principle. Rational knowledge is thus "knowledge from principles," or "knowledge of the particular in the universal" (KrV A300/B357). Since "a priori knowledge" is another name for knowledge from principles, "rational knowledge and a priori knowledge are the same" (KpV 5:12). Kant points out, however, that we speak of principles in a comparative as well as an absolute sense (KrV A300/B357). Many

conclusions of reason are derived from universal knowledge that is nevertheless contingent, having been acquired through experience. But knowledge that is a priori in the strict sense derives from principles having an absolute necessity, which marks them out as cognized through reason alone. Kant thus characterizes reason as “the faculty of principles” (KrV A299/B356; MS 6:214), indicating thereby that the principles of rational cognition have their origin solely in reason.

As the source of principles, reason must be conceived as spontaneity, “the capacity to produce representations itself” (KrV A51/B75), or as a self-active, self-determining power. An investigation that seeks to identify reason’s principles must therefore do so through self-consciousness, or reflection, abstracting from the conditions in which reason is exercised. For only self-consciousness provides an understanding not dependent on affections of receptivity and the contingent conditions they reflect; without self-consciousness there would be no thought of self, nor therefore of representations as *self*-produced, or spontaneous.

Such an investigation is just what Kant undertakes with regard to reason in both its theoretical and its practical application. He says the *Critique of Pure Reason*, which sets out the fundamental laws of nature, “rests on no facts whatsoever,” “taking nothing as given for its basis except reason itself” (Pr. 4:274). Likewise in morals, reason “need presuppose only *itself*” in its legislation (KpV 5:20–21), and philosophy must accordingly be the “sustainer of its own laws,” occupying a position “that is to be firm even though it is neither dependent on anything in heaven nor supported by anything on earth” (G 4:425). And in both cases the principles are grasped in self-consciousness. Kant argues in the *Critique of Pure Reason* that the fundamental laws of nature, being presupposed in experience rather than discovered through it, are recognizable only through reflection, in the act of theoretical cognition, on such cognition’s form. And he makes a parallel point in the *Critique of Practical Reason*, that the fundamental law of freedom is just the self-consciousness of pure practical reason (KpV 5:29, 42, 46).

1.3 How are we to understand the distinction in reason’s application? Kant’s way of drawing the contrast in the *Groundwork*’s Preface, by distinguishing between laws of nature and laws of freedom, might seem to suggest that he sees it as derived from a difference in the objects, or the material cognized. But this appearance dissolves when we consider other passages, where Kant says the objects determined in theoretical knowledge must be given from elsewhere, whereas practical knowledge need not wait for objects to be given in order to know them, but rather works

to make the object it determines actual (KrV Bix–x; KpV 5:46). In other words, in theoretical knowledge the actuality of the knowledge depends on the actuality of its object, whereas in practical knowledge the actuality of the object depends on the actuality of the knowledge. Thus, it is possible for practical and theoretical knowledge to share the very same object: so far as practical knowledge makes its object actual, the latter can be known also theoretically. What “ought to happen” and what “happens” are then one and the same. These points indicate that practical knowledge is distinctive in that it is efficacious. To make its object actual, it must have a certain causality.

Theoretical and practical knowledge can also be distinguished in another way, however, by saying that the former is of an object originally represented as distinct from the cognizing subject, whereas the latter is at bottom a form of self-knowledge, in which the object known is the knowing subject.<sup>3</sup> This way of marking the difference does not lie on the very surface of Kant’s text, but it is directly implicated in explanations he does offer, such as the ones just noted, which highlight the difference in the direction of existential dependence that theoretical and practical knowledge bear to their respective objects. He says, for instance, that theoretical knowledge concerns “objects that may be given to reason somehow from elsewhere,” while practical knowledge “can become the ground of the existence of the objects themselves” (KpV 5:46). To regard an object as given from elsewhere is clearly to regard it as distinct from oneself. Since the conception of self originates in self-consciousness alone, everything included in what is originally understood as self is contained in self-consciousness; and since in self-consciousness consciousness understands itself as identical with what it is conscious of, the latter cannot be conceived as given to consciousness from elsewhere. And as for the claim about practical knowledge, reflection on what it is for knowledge to become the ground of its object’s existence positions us to see that such knowledge must be self-knowledge. The crucial consideration here is that practical cognition’s causality is essentially self-conscious, hence not such as could be discovered only through experience, but originally represented in the knowledge itself. The essential self-representation of this causal relation is reflected in Kant’s characterization of an end – the object of such knowledge – as “the object of a concept, so far as the latter is regarded as the cause of the former (the real ground of its possibility)” (KU 5:220; cf. MS 6:384). Practical knowledge, then, in being conscious of its own efficacy,

<sup>3</sup> Although plural as well as singular subjects are possible, for present purposes it will suffice to consider the singular case, which is primitive.

represents itself as the cause of its object. To that extent, it is a form of self-knowledge, though not, of course, theoretical self-knowledge: it can be the cause of its object only through representing itself as the cause.

To prevent confusion, we should observe that in the foregoing discussion “object” has been used in two senses, reflecting the presence of two moments in the act of practical cognition. As a form of discursive cognition, practical cognition lies in an act of judgment, in which a practical predicate, the concept of a possible effect (e.g. to keep one’s promise), is attached to a concept of the subject. When Kant says practical knowledge can make its object actual, “object” refers to the effect the judging subject represents in the act of practical predication, or (what amounts to the same) to the content of the judgment. But when we say practical knowledge is self-knowledge, in which subject and object are the same, “object” signifies the judgment’s subject matter, what is thought through its subject concept. Thus, if the judgment is that I ought to keep my promise, “object” in the latter sense refers to myself, in the former to myself keeping my promise.

In sum, practical knowledge is efficacious rational self-knowledge. As *self-knowledge*, it lies in the practically cognizing subject’s attachment of a predicate to itself. As *rational*, it is knowledge in which the predicate’s attachment is derived from a principle, or representation of a law, to which the subject subordinates itself. It is thus an act of cognitive self-determination from a principle: one’s determination, derived from a principle, of what one ought to do. And as *efficacious*, it is an act of causal self-determination as well, a law-governed act of making the object cognized actual: through it, one determines oneself to do what one knows one ought to do, making happen what ought to happen. In this act – the conclusion of a practical syllogism, in which reason is employed to derive an action from a law – rational cognition and causality are united.

It deserves notice that it was the earlier characterization of reason as spontaneity that made it possible to appreciate the identity of cognition and causality in the act of practical knowledge. Had the possibility been ignored of investigating reason through self-consciousness – the one, original position from which reason’s spontaneity can be recognized – it would have seemed natural to regard reason as a form of receptivity, and hence natural to suppose, as Hume famously did, that “reason is perfectly inert.”<sup>4</sup> It would then have been difficult even to comprehend how cognition could itself be a form of causality. But no such impediment confronts the conception of reason as spontaneity. Though the concepts of spontaneity and causal power

<sup>4</sup> *A Treatise of Human Nature*, 2nd edn., L. A. Selby-Bigge and P. H. Nidditch (eds.), (Oxford: Clarendon Press, 1978), III.1.1, 458; cf. II.iii.3, 413–18.

are not the same (as the possibility of theoretical cognition shows), spontaneity, as self-active, can nevertheless, in a suitable condition, be also a causal power. We have next to consider this condition.

1.4 The distinction in human reason's application reflects an inherent limitation. If we consider the idea of an infinite intellect, which Kant occasionally deploys as a foil to highlight the finitude of our discursive reason, we will note that such a cognitive power, as omniscient, would be creative, causing the existence of the object it knows in the very act of knowing it (KrV B72). Such cognition would contain as a unity what is present only in a divided way in discursive rational cognition. For discursive reason, achieving completeness or perfection in its cognitive activity involves establishing a harmony across its two applications, an agreement between its knowledge of what happens and its knowledge of what ought to happen. In infinite cognition, this harmony is immediate, pertaining to different aspects of a simple act; such cognition contains no discrete acts of creation and inspection, being at once knowledge of what is and knowledge of what is to be.

The limitation of discursive reason implies that its knowledge is in a sense subject to certain external conditions. But the limitation cannot spring from these conditions. Reason's finitude cannot arise from an externally imposed limitation on an originally infinite cognitive power, for the idea of such a power excludes the possibility of such limitation. Nor can reason's limitation stem from inner conflict, from one component's restricting or infringing another. For such conflict is incompatible with the unity that the self-consciousness of cognitive activity establishes as essential to all cognition and hence to the cognitive capacity, even if finite. Nothing originally understood as absolute unity – as the cognitive power is – could conceivably limit itself by opposing itself. The inner limitation must rather lie in a certain lack of completeness in the capacity, entailing a reliance upon external conditions for its exercise, grounding a distinction between power and act. The external conditions must include things whose existence lies outside discursive consciousness, among them subjects – bearers of discursive reason – coexisting with the rest in such a way that their reason can be exercised. And since reason is spontaneity, the ground of this way of coexisting, so far as it lies in the subject, under the name of sensibility, must contribute to reason's exercise, not by determining it, but by enabling reason to determine itself. As enabling, this subjective ground must be a cooperating representational power; as cooperating without determining, it must be determinable material to which reason can apply itself, namely receptivity, the capacity to acquire representations through being affected.



Though it reflects reason's inner limitation, the distinction in its application cannot come into view so long as we attend, as we did earlier (section 1.2), merely to what is the same across the different applications. Kant's claim that "in the end there can be only one and the same reason" (G 4:391) can have no basis but the recognition that there is at bottom but one thing it ever does. The distinction must therefore depend on some division in the subjective condition of reason's exercise, in sensibility. The division is easily found. As the subjective ground of the reason-enabling mode of the subject's coexistence with other things, sensibility can reflect the two sides of coexistence, passive and active. It can accordingly include two powers, each able to cooperate with reason, enabling it while at the same time being determined by it. The first is *sense*, the capacity to acquire representations of objects so far as those objects affect the mind; the second is *feeling*, the capacity to acquire desires for or aversions to objects through being affected – pleased or displeased – by sense representations of them (MS 6:211–12n). The first is recognized through consciousness of certain representations as effects; the second through consciousness of certain representations as having a causality of their own. On the consciousness of sense is founded the consciousness of a power of perception (or a capacity to be conscious of objects' actuality); on the consciousness of feeling is founded the consciousness of a faculty of desire, which as we shall see Kant defines as a representation-involving causal power, through which objects can be produced, or made actual. Being dependent on a sensibility that can enable its exercise through these two powers, one and the same reason can distinguish itself in application. In relation to the capacity for perception, this spontaneity constitutes the capacity for theoretical knowledge, under the title of the *understanding*; in relation to the faculty of desire, it constitutes the capacity for practical knowledge, under the heading of the *will*.

We next examine Kant's account of the faculty of desire, then consider how the will is constituted through the relation reason bears to this faculty.

## 2 DESIRE

2.1 The section of the Introduction to the *Metaphysics of Morals* containing the discussion of desire and will bears the caption "Of the relation of the faculties of the human mind to moral laws."<sup>5</sup> Although Kant holds

<sup>5</sup> I thank Barbara Herman and Andrews Reath for very helpful conversations regarding this section of the Introduction.

there to be at bottom three such faculties – the faculty of knowledge, the faculty of desire, and the feeling of pleasure and displeasure (KU 5:177) – here his eye is chiefly trained on the faculty of desire. The faculty of knowledge is not directly discussed, though as we shall see it is implicated in his account of the faculty of desire. Feeling receives considerable attention, not however as an independent faculty, but only insofar as its operation is combined with desire, as cause or effect.

Some readers, suspicious of talk of faculties and capacities, may wonder what the point could be of introducing the notion of a faculty of desire, preferring instead simply to speak of desires. This notion is needed, however, to represent a grounding for all desire in the single nature of a living, or animal, being. As Kant notes elsewhere, life-power expresses itself in the faculty of desire (KpV 5:23). An animal's desires belong to it, not in a sheerly accidental way, as mere elements of a "motivational set," but as modes of its living, as determinate actualizations of its life-capacity. By representing desires as grounded in a power in the living being, the notion of a faculty holds in view their relation to that being's nature. On account of that relation, there are "laws of the faculty of desire"; at one point Kant even characterizes life as the capacity to act in accordance with such laws (KpV 5:9n). Some of these laws must be discovered empirically, through pleasure and displeasure, pursuit and avoidance, but there are certain things relating even to such laws that can be inferred a priori from the concept of life itself, as a form of organized natural being, notably that all of an animal's desires, to the extent that they are healthy expressions of its nature, will be in systematic harmony with one another (KU 5:430).

Kant defines the faculty of desire (*Begehrungsvermögen*) – or the capacity to desire, as it might also be called – as "the capacity [*Vermögen*] to be, through one's representations, the cause of the objects of those representations" (MS 6:211). In speaking of the cause of the objects of one's representations, Kant means the cause of those objects in respect of their *actuality*, or *existence*. Thus, in the *Critique of Practical Reason*, this faculty is defined as the capacity of a living being "to be through its representations the cause of the actuality of the objects of those representations" (KpV 5:9n). If for instance the representation of one's health is included among the representations through which one is such a cause, then one's faculty of desire will include the capacity to be the cause of one's health, the cause through whose action one's health is effected, or made actual.

It follows that the representing through which, in desire, the subject is the cause of the existence of the represented object differs from the

representing that figures in theoretical knowledge. For the very idea of such knowledge implies that the actuality of the representation depends on that of the object represented, not the reverse. For a similar reason, desiderative representation cannot lie in perception. Perception represents its object through sensation, which, as a modification of consciousness dependent on the workings of the outer senses (sight, hearing, etc.), is the effect an already existing object present to the senses has on the subject's capacity to represent. Perception also includes, however, the exercise of the imagination (KrV A120n), the capacity to represent objects in intuition even without their presence (KrV B151; ApH 7:153, 167). Insofar as this capacity can be used to represent an object not already present, it can furnish representations suited to figure in the exercise of the faculty of desire. The possibility that desiderative representation might be a concept, or even an idea or principle, will be considered in due course.

2.2 Two points about Kant's definition are particularly significant for our purposes here. The first is that it situates desire under the broad headings of causality and action rather than affection and passion. Despite the long-standing tendency to conceive of desire in terms of want and passion, which entail need, dependence, and passivity, Kant rejects as tendentious definitions that build such notions in from the beginning (KpV 5:9n). To claim that all desire arises from affection – from the feeling of pleasure or displeasure, as a hedonist would hold – would be to advance a substantive and disputable thesis. It is true that desire implies limitation and dependence in one sense, owing to its essential relation to life, a form of organized and therefore merely contingent natural being.<sup>6</sup> But this limitation lies, not in the way the faculty of desire is determined, but in its efficacy, in the productive power residing in the animal's capacity to be the cause of the object it represents. Such a limitation entails a dependence on external conditions, a dependence on account of which the desired object (one's health, say) may not be made actual if such conditions are unfavorable; but it does not imply that all desire arises from affection. It is also true that there is a form of desire – namely sensible desire – that is essentially passive, depending on affection in the form of a feeling of

<sup>6</sup> Organized natural being (the subject matter of biology, embracing the vegetable as well as the animal kingdom) is naturally self-productive (see KU 5:369–72), but as a product of nature, indeed a self-organizing product, its existence (unlike that of bare matter) lacks natural necessity. What marks life (animality) out as a distinct form of organized natural being is that in its case the (contingent) self-production generically characteristic of the latter is self-production *through representation*. The faculty of desire is precisely the natural capacity for such representation-guided self-production.

pleasure or displeasure, and Kant notes that the term “desire” (*Begierde*) is sometimes used in a correspondingly narrow sense (MS 6:212). But this dependence is not already implicated in the original concept of desire. That concept does, on the other hand, include the thought of causality. As Kant observes, such causality figures even in bare wishing. “As *striving* (*nisus*) to be by means of one’s representations a *cause*, a desire is ... always causality” (MS 6:356).

This point has an obvious importance for our investigation, comparable to the significance we found in Kant’s characterization of reason as spontaneity (section 1.3). In order to conceive of the will as at once cognition and causality – reason and desire – it is necessary to avoid not only the assumption that reason is a passive power of apprehension and hence inert, but also the assumption that the faculty of desire is determinable only from without, through the effects objects produce upon the mind in perception and experience. Having introduced the faculty of desire through a definition free of any such assumption, Kant keeps open a path by which he can approach a conception of a form of faculty of desire that is spontaneously determinable.<sup>7</sup>

The second point is that, by identifying desiring as a type of causality distinguished from other forms by the involvement of representation, the definition marks an internal dependence of the faculty of desire on the faculty of representation. Whether the latter capacity can be exercised without the involvement of the former is a nice question, which might be disputed by some,<sup>8</sup> but it is not even conceivable that a being might desire without representing.

This second point is significant in part because it reflects Kant’s recognition that living beings generally have a nature that comprises the dynamically interrelating representational capacities of perception and desire. Earlier, we noted the relation these capacities bear to *rational cognition* on account of their being integral to the subject’s reason-enabling mode of *coexistence* with other things (section 1.4). But appreciating how representation is involved in desire brings into view the essential relation these capacities have

<sup>7</sup> Though Kant keeps his conceptions of reason and desire free of the assumptions just mentioned, he does not aim in the *Metaphysics of Morals* to show that the cognitive power itself has causality, that pure reason is practical. That task was addressed earlier, in the *Critique of Practical Reason*.

<sup>8</sup> Aristotle holds at least that where there is the *power* of sense there is also the *power* of desire. See *De Anima* 11.3 414b1–2 (*The Complete Works of Aristotle: The Revised Oxford Translation*, J. Barnes (ed.) (Princeton University Press, 1984), vol. 1, 659). Kant too supposes that any animal having the first of these powers also has the second. But he does so on the grounds that life entails desire, which in turn entails sense. And he seems to hold that only a theoretical application is already implied by the bare idea of the faculty of discursive cognition (G 4:395; KU 5:178).

to *one another* in the integral roles they play in the very *existence* of an animal being. A causal power can be representational only in cooperation with a power of *perceptual* representation. The striving in which desire consists comes to rest only in experience or perception – that is, in a representation of the existence of the object the desire works to produce. And conversely where perceptual and desiderative representations thus coincide, a feeling of pleasure, or consciousness of the desire's reinforcement, is necessarily involved. Thus, the perception sustains the desire inwardly through feeling, just as the desire sustains the perception outwardly through action. Since it is in the faculty of desire that life-power expresses itself, this relation of mutual furtherance between perception and desire constitutes the specific form of self-production distinctive of life as a form of self-organizing natural being, and pleasure resulting from the act of this faculty is accordingly life's consciousness of its own self-production.

The dependence of the faculty of desire on the faculty of representation is significant for another reason as well. In the paragraphs following his definition of the faculty of desire, Kant introduces increasingly specific forms of desiderative capacity by characterizing them in terms of increasingly specific forms of representational capacity. In particular, he proceeds from (i) the *faculty of desire* through (ii) the *faculty of desire in accordance with concepts* to (iii) the *will* by moving from (a) the *faculty of representation* through (b) the *faculty of concepts* (the understanding, or the faculty of knowledge) to (c) the *faculty of principles* (reason, the faculty of a priori knowledge). But though various forms of desire are distinguished, it will emerge as we follow his exposition that they are related as moments in a single sensibly affected but rationally determinable desiderative capacity in a human being under moral laws.

In what follows, we shall ascend this ladder, advancing from receptivity to spontaneity, from sensible desire to the will.

2.3 *Desire in the narrow sense (sensible desire)*. Kant holds that desire always has pleasure (or displeasure) combined with it. He calls this pleasure *practical* to indicate that it is necessarily combined with desire, in that it accompanies a representation of the *existence* of an object (i.e. a *sensation-involving* representation of the object, as in experience or perception), making the self-sustaining causality integral to it as pleasure (KU 5:220) identical with that of a desire for the object. But as we noted, Kant's definition of the faculty of desire nevertheless leaves unspecified whether desire always arises *through* such a feeling. He goes on to consider briefly the form of desire that does arise in such a way and to contrast with it

the form that precedes the pleasure. In discussing the former – sensible desire, or desire in the narrow sense – he introduces two further notions, those of inclination, or habitual desire, and interest.

No account is offered of inclination, but the points recorded above concerning the faculty of desire, along with remarks Kant offers elsewhere, suggest the following view. In a living being, the faculty of desire includes within its natural constitution certain propensities and instincts, owing to which the experience or perception of certain objects is pleasing (cf. RGV 6:28–29n). When such an object is experienced, the pleasure determines the faculty to an act of desire, a representation of the object of the pleasing experience (the object, not its existence), and the animal is thereby and to that extent moved to pursue or to produce the object. That is to say, the animal is moved to reproduce the pleasing experience of the object by keeping the thing present to the senses thereby maintaining that experience of it, or by bringing about another such experience of such an object. The returning pleasure is a further causal stimulus, reinforcing the original desire. Thus, desires arising from pleasing experiences naturally tend to become habitual, establishing themselves as inclinations of the faculty of desire. (To the extent that they are also grounded in the animal's instincts, which have as their objects its "true natural needs," they qualify as *natural* inclinations.)

The habitual character of sensible desire shares such desire's nature as receptive. The habit arises in the faculty of desire without needing to be an object represented in any act of that faculty. It is possible, of course, to desire to instill habitual desires in oneself or in another and to inculcate them successfully. Indeed, such habituation is integral to the cultivation of the faculty of desire in moral upbringing. But this possibility depends on the natural constitution of the faculty of desire, on account of which habitual desire would arise willy-nilly in any case, provided only that the object be experienced and re-experienceable in conditions not unfavorable for the habit's acquisition. The habitual character of the pleasure's connection with the faculty of desire no more belongs to sensible desire's object than does sensible desire itself; it arises without needing even to be noticed at all by the subject.

Next in the ascending sequence is the concept of interest: "the combination of the pleasure with the faculty of desire is called *interest*, so far as this connection is judged through the understanding to be valid according to a universal rule (if only for the subject)" (MS 6:212). The combination, or connection, here referred to seems clearly to be the one already implied in the notion of practical pleasure, so in speaking of such

a connection so far as it is judged valid by the understanding according to a universal rule, Kant evidently has in mind the sort of general combination implicated in habitual desire, but as represented as such by the understanding. Such a relation is not just the generic connection that any pleasure in an object's existence has with the faculty of desire, but a specific connection of the pleasure figuring in the enjoyment of a certain object (an apple, say, or the company of another) with that faculty, a connection residing in an habitual determination of the latter – an inclination – to have that same thing as its object. And while as we noted this connection might hold without the subject's being aware of its generality, it is only through such awareness that there is any interest. (So interest, unlike inclination, “is never attributed to a being unless it has reason” [KpV 5:79].) Kant immediately goes on, however, to identify an interest of the sort in question – an “interest of inclination” – with the pleasure itself, not, as he initially suggested, with the combination of the pleasure with the faculty of desire; and other passages too suggest that he regards interest as a pleasure.<sup>9</sup> Presumably his thought is that once an inclination is in place, as one term in a stable, homeostatic connection between the pleasing experience of some object and the faculty of desire, an animal may, if it has understanding, notice this general connection and represent it conceptually, through a rule, in which, on account of the rule's generality, the pleasure must itself be represented through a *concept* of the object the representation of whose existence it accompanies: in such a case, the pleasure will count as an interest. Thus, unlike the object of inclination, which may be no more than what is represented through an animal's reproductive imagination, an object of interest is always represented through a concept figuring in a rule by which the subject represents the object's relation to its own faculty of desire, or to itself as a living being. Kant does not identify this conceptual representation as an act of the faculty of desire. And rightly, for it cannot be the inclination, though based on it, nor is it a wish, a choice, or an exercise of the will. But he does see it as the basis for a certain form of desire, for he takes maxims, which are exercises of the power of choice (MS 6:226), to be founded on interest (KpV 5:79). We shall return to this relation, but we can anticipate that a maxim goes beyond interest in that it first introduces a conception of action through which the object of interest is to be made actual.

<sup>9</sup> “The delight [*Wohlgefallen*] that we combine with the representation of the existence of an object is called interest” (KU 5:204; see also KU 5:296).

Kant concludes his discussion of practical pleasure with a comment on the other way such pleasure may be combined with desire, as the latter's effect rather than its cause. Here the pleasure "must be called an intellectual pleasure and the interest in the object an interest of reason." Kant is clearly supposing that in this case the faculty of desire is determined spontaneously, by the intellect, or reason. He is also clearly thinking that insofar as reason determines the faculty of desire, it does so in a *practical* use, under the name of the will. As he remarks in the *Critique of Judgment*, "to will something and to delight in the existence of the same, that is, to take an interest in it, are identical" (KU 5:209).

Kant's implicit claim that the faculty of desire is determinable in just the two ways indicated has been questioned from different angles. Some have doubted his contention that, aside from reason, only pleasure can determine this faculty;<sup>10</sup> others have seen reason as having no genuine practical use, but at most a theoretical power to apprehend or to intuit goodness or value.<sup>11</sup> While detailed consideration of such challenges lies outside our present concern, we can throw some light on the issues while pursuing our immediate purpose if we pause briefly to appreciate how the exhaustive and exclusive character of Kant's distinction between the two ways the faculty of desire can be determined flows from the exhaustive and exclusive character of his distinctions between theoretical and practical and between sensible and intellectual representation. As we noted, a living being's capacity to represent may be operative either as ground or as consequence of its object's actuality. In the former case, it constitutes the faculty of desire; in the latter, the capacity to represent existence ("what is") in perception or experience. Where, as in the human case, the capacity to represent is a cognitive capacity, this distinction is between its practical and theoretical applications. Now in both theoretical and practical cognition, the distinction between the sensible and the intellectual is drawn in terms of the contrast between the two representational powers that cooperate in discursive cognition – externally determined receptivity and self-determining spontaneity. Both are requisite, self-determination for cognition's formal unity, external determination for its material content.

<sup>10</sup> See esp. KpV 5:21–25. For discussion, see Andrews Reath, "Hedonism, Heteronomy, and Kant's Principle of Happiness," in his *Agency and Autonomy in Kant's Moral Theory* (Oxford University Press, 2006), 33–66; and Barbara Herman, "Rethinking Kant's Hedonism," in her *Moral Literacy* (Cambridge, Mass.: Harvard University Press, 2007), 176–202.

<sup>11</sup> Kant would presumably regard such a view as subject to his general criticism of heteronomous theories of morality. See John Rawls, "Themes in Kant's Moral Philosophy," in *Kant's Transcendental Deductions: The Three "Critiques" and the "Opus postumum,"* Eckart Förster (ed.), (Stanford University Press, 1989), 95–98.



But the external determination takes different forms in accordance with the difference in sensibility that, as we saw (section 1.4), grounds cognition's division into theoretical and practical. In theoretical cognition, the external determination is of *sense* (in sensation) by *objects*, whose existence is thereby represented; in practical, it is of *feeling* (of pleasure or displeasure) by the resulting *representations* of those objects' existence, whereby those objects become objects of desire or aversion. Thus, in the former, there is always sensation connected with the capacity to represent existence, but in a human being, where this capacity is not a bare animal power (of sensible representation, or perception) but a capacity for theoretical cognition, the sensation need not always precede the latter's representations; it does in the case of sensible representation, but it follows where the representation is a self-determining act of the intellect (reason in the guise of the understanding). In parallel fashion, there is always feeling connected with the faculty of desire, but in a human being, where this faculty is not a bare animal power (of sensible desire) but a capacity for practical cognition, the feeling need not always precede the latter's representations; it does in the case of sensible desire, but it follows where the representation is a self-determining act of the intellect (reason in the guise of the will).

### 3 THE WILL

3.1 *The faculty of desire in accordance with concepts.* Having outlined separately Kant's conceptions of reason and desire as powers of cognition and causality, we next consider them together, in the idea of the will. Kant approaches the will by situating it under the general heading of "the faculty of desire in accordance with concepts" (MS 6:213).<sup>12</sup> This latter

<sup>12</sup> Between his treatments of practical pleasure and the faculty of desire in accordance with concepts, Kant pauses to distinguish concupiscence from desire, describing the former as a "stimulus to the determination" of the latter, and as "a sensible determination of the mind but one that has not yet grown to an act of the faculty of desire" (MS 6:213). Kant's meaning is unclear, but possibly he has in view that moment in the feeling of pleasure that traces to the condition of our sensible nature (a condition knowable only empirically) on account of which it is possible for choice to be contrary to or only reluctantly in conformity with reason's moral law: the condition, that is, owing to which this law operates in us as an imperative or constraint – a principle that necessitates – making moral perfection for us a matter of virtue rather than holiness. In the *Doctrine of Virtue* he says this constraint "applies not to rational beings in general (there could perhaps also be *holy* ones) but rather to *human beings*, rational *natural beings*, who are unholy enough that pleasure can induce them to overstep the moral law, though they recognize its authority, and, even when they follow it, to do so *reluctantly* (with opposition from their inclinations)" (MS 6:379). But since this topic receives no further attention in Kant's advance from desire to will in the passage we are examining, we need not pursue it here. For discussion, see Herman, *Moral Literacy*, 231–34.

faculty is not defined.<sup>13</sup> But the reference to concepts shows that Kant means to designate a form of desire that depends not just on the capacity to represent but specifically on the understanding, the capacity for discursive cognition, a form of representation that lies in the combination of concepts in thought and judgment. The preposition “in accordance with” (*nach*) might seem to leave open the possibility that this faculty’s operation is distinct from the use of the concepts with which it accords, so that the desire could stand to the understanding’s act in something like the way an inclination instilled in a young, uncomprehending child stands to the thought that guided its parent’s training of it. But this is plainly not what Kant has in view, since the acts he ascribes to this capacity are wishing, choosing, and willing, none of which are inclinations. His idea seems rather to be that this faculty’s acts are thoughts or judgments that are themselves desires.

To understand how thinking can be identical with desiring, two points noted earlier must be borne in mind: that the cognitive power’s exercise is self-conscious and spontaneous, and that the concept of desire does not already imply that all desire arises from affection, or external determination (from some pleasing experience). Appreciating these points enables us to comprehend the identity, and spelling out the implications of that identity will bring into focus the form of desire common to wishing, choosing, and willing. Since the act of the faculty of desire in accordance with concepts is a judgment (or thought) that is also a desire, the self-consciousness essential to it as a judgment must also include consciousness of itself as a desire, as a causality, a striving. And for essentially the same reason, the act’s self-consciousness must include awareness that the judging subject and the desiring subject are likewise the same. The judgment thus contains awareness of itself as efficacious, and the judging subject is therein necessarily aware of itself as the agent of the action the judgment strives to effect. Consider choice, for instance. In my choice to pursue some object, to make it actual, I am aware, at least implicitly, that I, the choosing subject, am the very agent who through this choice is to pursue that object. The exercise of this faculty thus always has two moments – one cognitive, the other causal – united in a single act. The former lies in the subject’s judgment, its self-conscious determination of

<sup>13</sup> Some of the few other passages where similar expressions occur may suggest that, for Kant, speaking of the faculty of desire in accordance with concepts is just another way of speaking of the will (e.g. KU 5:220). From the way the expression is used here, however, it is clear that it has a broader sense in the present context.

what it should do, the latter in the causality of which the subject is conscious *in* and *through* that judgment, and by which that subject, as agent, does (barring unfavorable conditions) what, in that judgment, it sets itself to do. This two-sided self-consciousness will prove significant. For the moment it will suffice to note that it entails a twofold self-conception, in which the subject regards itself first and originally as a knower, a bearer of the capacity for cognition, and second as also an agent, a subject with the power, through its judgments, to make the objects therein represented actual. The first moment is basic, since self-consciousness belongs originally to cognition and only through it to anything else.

From the implications just traced, we can outline the general form of the act of this faculty. As a judgment, this act must lie in a self-conscious combination of concepts. Being also an exercise of causality, it must lie in a *reflexive* use of the concept of cause, in which causality and the thought of it are the same. In addition, the subject making the judgment is necessarily conscious of itself as identical with the subject the judgment concerns, the agent. Finally, since determinations of the faculty of desire lie in representations of the objects they themselves work to make actual, so must exercises of the faculty of desire in accordance with concepts. From these points it follows that the judgment in which the exercise of the latter faculty consists must include a concept of the subject – indeed the twofold conception just noted – and a predicate, representing some effect to be produced (the object); and the judgment itself must consist in the *use* of these concepts, the attachment of the predicate to the concept of the subject, and specifically the *practical* use of them, that is, a *reflexive* act of attachment, conscious of itself as both a judgment and a desire, as at once the cognition of causality and the causality of cognition. We can thus distinguish in the act of practical judgment the materials – the concept of the subject and the concept of the effect – and the act of combination, which constitutes the practical cognition of the very thing it also (to the extent that conditions permit) effects, namely the action, or the relation of causal dependence of the object on the subject, the effect on the cause. Also to be noted, however, is an asymmetry in the way the subject and predicate concepts figure in the combination. The judgment begins with the concept of the subject, not the predicate, and the attachment of the predicate yields an enlarged, more determinate conception of the subject. It is therefore a condition of the judgment's validity that the predicate it attaches be in agreement with the subject concept. Hence, so far as there are formal, or necessary, elements of the self-conception that serves as the subject concept for the act of practical judgment, there will

be formal, universal conditions of self-agreement to which the judgment's act of attaching its predicate will be subject.

Now this self-conception does, as we saw, necessarily include two moments, consciousness of oneself as a cognizer, a bearer of reason, and consciousness of oneself as a cause, a subject with a faculty of desire. The act of the faculty of desire in accordance with concepts – the act of practical judgment – can accordingly be considered in relation to each of these moments. As will be suggested below, the power of choice and the will lie in the capacity these two moments of self-consciousness have to contribute to the determination of practical judgment: so far as such judgment is determined by the subject's cognizance of its causal power, it lies in the exercise of the power of choice; so far as it is determined by the subject's consciousness of its cognitive capacity, it is determined by the will.

3.2 *The power of choice.* Kant says of the faculty of desire in accordance with concepts that “so far as it is combined with the consciousness of the capacity [*Vermögen*] of its action to produce the object, it is called *power of choice* [*Willkür*]; but if it is not combined therewith, its act is called a *wish*” (MS 6:213). At first glance, it might seem that wish is mentioned here in an aside, to provide a contrast for the power of choice. But closer examination reveals this act to have a fundamental importance for the metaphysics of morals and to be the proper starting point for a consideration of choice. To an inattentive eye, Kant might appear to be suggesting that wish is combined with consciousness of a lack of capacity to produce the object. Certain other passages, dealing with wishes for things that are impossible or beyond one's power (e.g. KU 5:177–78n), can contribute to the impression. But while Kant does in places seem to have such a conception in view, in others he employs a broader notion. The present passage is a case in point. Wishing is said to be marked, not by consciousness of lack of capacity, but by lack of consciousness of capacity. Rather than regarding the object as unattainable, wish represents the action indeterminately and problematically, leaving open the possibility that choice might be reachable by finding a more determinate representation of the action, one that renders it recognizably within one's capacity while still sufficient for the production of the wished-for object.<sup>14</sup> Wish is accordingly the beginning, a problematic major premise, in the exercise

<sup>14</sup> This broad conception of wish is also expressed in other passages, e.g. MS 6:482 and AA 28:254. Since Kant accepts many of the Scholastic doctrines that survived in the tradition of German school-philosophy to which he belonged, it is not surprising to find that his conception of wish is in broad agreement with the account presented in Aristotle's *Nicomachean Ethics*, 111.2–4 (in Barnes [ed.], *Complete Works*, vol. 11, 1755–58).

of the power of choice, and choice is the conclusion, reached through deliberation. So understood, wish can also be captured in other terms Kant uses. As an act that will be immediately expressed in choice where, in the absence of constraints and limitations, one has “the capacity to do or to refrain as one likes” (MS 6:213), it counts as an immediate “liking” (*das Belieben*). It can also be characterized as the “inner employment” of the power of choice, its first use, in maxims of ends, whereas the choice issuing from deliberation would constitute this power’s “outer employment,” in maxims of action.<sup>15</sup>

To appreciate this deliberation in its proper setting, we must of course bear in mind not only the human subject’s empirically gained theoretical cognizance of the extent of its power as an agent, but also an additional, empirically determined component of its self-conception, touched on in our discussion of sensible desire and interest (section 2.3). As a human subject, one is conscious through the feeling of pleasure that one’s faculty of desire is determinable by one’s experience of certain objects,<sup>16</sup> and one can thereby gain acquaintance with the empirically modified constitution of that capacity. So far as one becomes cognizant of one’s inclinations, one can conceptually represent their objects, now objects of interest, and regard them as elements of one’s happiness. The concepts of these objects provide the materials for practical predicates one can attach to oneself in wishes, or maxims of ends. Deliberation ensues, as in each case one’s power of choice strives inwardly to reach a practical judgment in which the attachment of the practical predicate containing the concept of the object of interest is “combined with the consciousness of the capacity of its action to produce the object.” Securing such consciousness will require a deployment of reason in the service of this interest of inclination, a specification of the wish’s practical predicate in accordance with one’s awareness of one’s power of agency, including one’s empirically determined cognizance of its extent, what one can and cannot do. The office of the power of choice, then, is to close the apparent gap between the habitually pleasing objects of one’s interest and one’s limited causal capacity. In choice, one subjects one’s practical judging to the condition of agreement with one’s empirically determined cognizance of oneself *as agent*,

<sup>15</sup> Kant’s distinction between the outer and the inner employment of the power of choice (MS 6:214) is closely related to his distinction between outer and inner freedom and his division of the *Metaphysics of Morals* into its two parts, the *Doctrine of Right* and the *Doctrine of Virtue* (cf. MS 6:406).

<sup>16</sup> More precisely, such experience determines the faculty of desire’s *lower*, receptive power, yielding sensible desire, which *affects*, but does not *determine*, the *higher*, spontaneous power, the faculty of desire in accordance with concepts (MS 6:213; KpV 5:32).

ensuring that ends set in acts of choice are possible, within one's capacity, and that the actions chosen are sufficient for reaching them.

3.3 *The will.* From the power of choice Kant advances to the will, ascending from the practically judging subject's conception of itself *as agent* to its conception of itself *as knower*, as bearer of the "one and the same reason" exercised in all judging, practical as well as theoretical: "The faculty of desire whose inner determining ground, consequently even whose liking, is met with in the subject's reason is called the *will* [*Wille*]" (MS 6:213). Now a conception of oneself as knower depends on an understanding (in self-consciousness) of the cognition of which one conceives oneself as capable, an idea of its form. So we can take Kant to be identifying the will with the capacity one's consciousness of the form of rational cognition in general has to determine one's capacity for practical judgment (in particular, as we shall see, one's power of choice), making its exercise agree with its form. We noted earlier that rational cognition is "knowledge from principles," or knowledge derived from universal knowledge originating in reason alone (section 1.2), and that such cognition, in the specific form it takes in reason's practical application, is efficacious self-knowledge (section 1.3). Cognition having such a form is precisely cognition belonging to a subject capable of acting from a recognition of universal law. The will thus lies in the capacity the practically judging subject's consciousness of this cognitive form – the form of universal law – has to determine its capacity for practical judgment, or its faculty of desire in accordance with concepts. So far as this consciousness has this capacity, the subject will (unless influenced by sensible desire) judge in conformity with that form. Through this capacity, then, reason subjects maxims to "the condition of suitability to be universal law" (MS 6:214).

From his initial characterization of the will, Kant draws the following conclusion, comparing and relating the will to the power of choice and identifying it with practical reason:

The will is therefore the faculty of desire, considered not so much (as is the power of choice) in relation to the action as rather in relation to the ground determining the power of choice to the action, and has itself properly no determining ground, but is, so far as it can determine the power of choice, practical reason itself. (MS 6:213)

On the proposed interpretation, the two relations in which the faculty of desire is here considered – "to the action" and "to the ground determining the power of choice to the action" – correspond respectively to the two ways the practically judging subject conceives of itself: as agent, through

whose action the object represented can be made actual, and as knower, or bearer of reason. The first correspondence is reasonably clear. The second may seem less so, owing to Kant's somewhat obscure description of the will as the faculty of desire considered "in relation to the ground determining the power of choice to the action." But Kant's meaning becomes clearer once we recall the earlier characterization of wish as the beginning of the exercise of the power of choice, the first framing of a maxim of an end. Just as it is through the exercise of the *power of choice* that we reach, in *choice*, the *representation of an action* that determines our *power of agency*, so it is through the exercise of the *will* that we frame, in *wish*, the *representation of an end* that determines our *power of choice*. Not that wish is an act of bare will. As the "inner employment" of the power of choice, wish also depends, for its materials, on the interests one acquires through cognition of one's inclinations. But to the extent that the will is efficacious, one attaches to oneself the practical predicates representing the objects of those interests only so far as the resulting maxims of ends are suitable to be universal laws. It is here in the first act of the faculty of desire in accordance with concepts – the determination of the ground that in turn determines the power of choice – that the will first brings its cognitive form to bear.

It is noteworthy that the will is here represented as having a certain precedence over the power of choice, even in the latter's inner employment. Kant underscores this supremacy in the next sentence by situating both wish and the power of choice *under* the will, and commentators have marked it too, identifying will and power of choice as, respectively, the "legislative" and the "executive" functions of practical reason.<sup>17</sup> But our investigations of reason and of desire position us to appreciate a metaphysical significance in the subordination. Will and power of choice, we saw, are distinguished through the difference between cognitive and causal self-consciousness. But we also noted that since self-consciousness belongs originally to cognition and only through it to anything else, the cognitive moment in our self-conception is prior to the causal moment (section 3.1), just as representation is prior to desire (section 2.2). The primacy of will over power of choice thus registers an essential priority of reason and cognition over desire and causality: knowledge is the rule for the use of the desiderative power in the conduct of life – not an instrument that serves it, but a determining form and pattern that governs it.

<sup>17</sup> See for instance Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason* (University of Chicago Press, 1960), 198–202; Henry E. Allison, *Kant's Theory of Freedom* (Cambridge University Press, 1990), 129–36.

This self-consciously recognized priority of cognitive form is of a piece with the will's autonomy, its character of being a law to itself. Since the will is practical reason, its autonomy is just the spontaneity of rational cognition in its practical application, and the will itself is just the capacity this self-determining faculty of knowledge has to determine the faculty of desire. Self-rule is the rule of knowledge.



## CHAPTER 3

### *Justice without virtue*

*Katrin Flikschuh*

#### I INTRODUCTION

An Aristotelian friend recently proposed to me an interpretation of the *Doctrine of Right* according to which Kant is best read as distinguishing between “acting justly” and “being just.” On that distinction, it is possible for a person to *act* justly without, in so acting, being just. Alternatively, in acting justly a person may simultaneously *be* just. The first person acts justly without making it her maxim so to act; the second person makes acting justly her maxim. This Aristotelian reading finds support in Kant’s characterization of duties of justice as “indirect” ethical duties (MS 6:219): all juridical duties are at the same time indirect ethical duties in that we should make it our maxim to act *from* juridical duties rather than merely *in accordance with* them. The reading does not deny that acting in mere outward conformity with justice is in some sense sufficient for the fulfillment of our juridical duties; however, the claim is that the person who also makes it her maxim to act from those duties is more fully just. It is an intuitively attractive reading: we do not want persons to act in mere outward conformity with justice – we want them to “internalize” the demands of justice, to act justly from inner conviction. Many believe that this must be Kant’s position, given his apparent claim in the *Groundwork* that our capacity for self-legislation constitutes the ground of morality in general.

I find the proposed reading troubling. I accept that, of the two, the person who acts in inner conformity with the demands of justice is the morally better person. However, I do not believe that this person is more

This chapter was written while I was in receipt of a Leverhulme Research Fellowship awarded to work on Kant’s political philosophy. I gratefully acknowledge the support of the Trust. I would also like to thank Christoph Horn, Reza Mosayebi, Amanda Perreau-Saussure, Heather Roff, Jens Timmermann, and Lea Ypi for helpful comments and stimulating conversation during my work on this chapter. Finally, my thanks to Lara Denis for her careful reading and helpful suggestions on improving the penultimate draft.

*just* than the other – I would say simply that he is more virtuous. On my reading, there is no distinction in *law* between acting justly and being just. One who acts in mere outward conformity with the demands of justice is as fully just as one who also makes it her maxim so to act. Indeed, I am inclined to believe that one who makes outer conformity of action with principles of justice conditional upon the latter's possible inner adoption as her maxim may be more likely to end up acting unjustly. I take this view because I believe that Kant's morality of Right is a public, not a private, morality. Universalizability of subjective maxims cannot serve as a test of the justness of proposed laws of external freedom that *are*, not *could be*, binding upon everyone.<sup>1</sup>

Some will find the claims in this last paragraph more contentious than others. Among those who focus on the *Doctrine of Right* many will concur that the morality of Right is a public morality which cannot be interpreted or evaluated in terms of criteria – good will, moral worthiness, autonomy of will – that belong to Kant's ethics. Yet those for whom the *Groundwork* offers Kant's definitive position on morality in general will object that the exclusion especially of the principle of self-legislation from the *Doctrine of Right* amounts to the latter's expulsion from Kant's moral philosophy. Not only do many Kant scholars take this view, but also contemporary Kant-inspired political philosophers who, following John Rawls, have taken the categorical imperative of the *Groundwork* to offer a decision-making procedure for the generation of valid principles of justice, and who in so doing have transformed Kant's ethics of self-legislation into a political morality of co-legislation.<sup>2</sup> There are, therefore, two sorts of objections to my proposed view that mere outward conformity of a person's action with general principles of justice is *all* that can be required by *justice* even if not by ethics. The first is the systematic objection which says that, if this is all Kantian justice amounts to, it can form no part of Kantian morality. The second is the substantive objection which says that no proposed principles of justice are legitimate that have not passed the test of possible self-legislation. In what follows I shall engage with both these objections, though I shall here pay more attention to the first and shall flag the second objection and my response to it more as an indication of the wider relevance of the issues at stake. I do not want to lose sight of my Aristotelian friend's initial proposal, nor of my opposition to it. I am struck by the strength of my opposition, but I suspect that it has

<sup>1</sup> I explain this distinction in section 4 below.

<sup>2</sup> John Rawls, *Lectures on the History of Moral Philosophy* (Cambridge, Mass.: Harvard University Press, 2000), 143–328.

less to do, in the end, with worries about a Kantian Aristotelianism than with worries, to the contrary, about liberal Kantianism.

In the next two sections I set out the systematic objection to the possible inclusion of Kant's philosophy of Right in his moral philosophy; I suggest that despite powerful textual support, it is mistaken. I then go on to argue that if one starts from central concerns of the *Doctrine of Right*, the text's distinctly public morality comes into focus. I draw on Kant's late *Wille/Willkür* distinction in order to tease out the distinctive nature of public juridical willing: surprisingly, in the political context, autonomy as self-legislation is simply irrelevant. Finally, I ask how individual political judgment relates to public willing; here we shall discover crucial differences between Kant's political morality and Kant-inspired current liberalism.

## 2 DOUBTS ABOUT THE MORALITY OF RIGHT

Doubts about the moral status of Right arise from Kant's characterization in the general Introduction to the *Metaphysics of Morals* of Right and virtue as distinct though complementary domains of moral lawgiving.<sup>3</sup> It is evident that Kant assumes the moral status of Right – he speaks explicitly and repeatedly of the *moral* concept of Right, of our juridical *duties*, of our (a priori) *moral obligation* to enter the civil condition. The question is whether he has systematic warrant for this assumption. Those who answer this question negatively tend to take it as settled that the capacity for autonomy – the capacity, that is, of moral self-legislation – lies unshakably at the core of Kant's account of moral obligation in general. The *Groundwork* offers seemingly decisive support for this view. When introducing the principle of autonomy between the second and third variants of the basic formulation of the categorical imperative, Kant asserts that:

In accordance with this principle all maxims are repudiated that are inconsistent with the will's own giving of universal law. Hence the will is not merely subject to the law but subject to it in such a way that it must be viewed as also giving the

<sup>3</sup> In the Anglo-American context, these doubts were first brought to general attention by Allen Wood's "The Final Form of Kant's Practical Philosophy," in Mark Timmons (ed.), *The Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 1–22. Discussion of the systematic status of the *Doctrine of Right* has a longer pedigree in German scholarship. A good overview can be found in Ralf Dreier, "Zur Einheit der praktischen Philosophie Kants," *Perspektiven der Philosophie* 5 (1979), 5–37. An early attempt at resolving the tension between virtue and Right is Otfried Höffe, "The Principle of Justice as Categorical Imperative of Law," in Y. Yovel (ed.), *Kant's Practical Philosophy Reconsidered* (Dordrecht: Kluwer Academic Publishers, 1989), 149–67.

law to itself *and just because of this as first subject to the law* (of which it can regard itself as the author). (G 4:431, emphasis added)

The message seems clear: if the autonomous will is subject to the law *because* it gives the law to itself, autonomy as self-legislation grounds moral obligation and possibly even the moral law itself. So: no moral obligation – possibly no moral law – without capacity for self-legislation. We also know from the *Groundwork* that the autonomous will is in a very particular state; it is a will that is receptive to objective determination through the mere idea of a law – a will cleansed of all subjectively valid determining grounds, such as desires and inclinations. So: no autonomous will without absence of inclinations as the will’s determining ground on the one hand and consciousness of the idea of law as its sole determining ground on the other hand. Contrast these claims from the *Groundwork* with the notorious passages in the Introduction to the *Metaphysics of Morals*:

In contrast to laws of nature, the laws of freedom are called *moral* laws. As directed merely to external actions and their conformity to law they are called *juridical* laws; but if they also require that they (the laws) themselves be the determining grounds of actions, they are ethical laws, and then one says that conformity with juridical laws is the *legality* of an action and conformity with ethical laws is its *morality*. (MS 6:214)

Juridical laws are here classed as moral laws and are distinguished, together with ethical laws, from laws of nature. However, the idea of autonomy – consciousness of the law as determining ground of actions – is implicitly restricted to ethical lawgiving. Kant both affirms that juridical laws are moral laws and implies that they are not autonomously legislated laws. If they are moral laws we are presumably obligated by them, but if autonomy of the will is not required in relation to them, autonomy is presumably not the ground of our obligation in relation to these laws. This implies, contrary to the *Groundwork*, that capacity for autonomy is not the ground of juridical obligation. Some commentators believe that juridical obligation is in fact non-moral. This view finds support in an extended further passage in the Introduction to the *Metaphysics of Morals*, which begins as follows:

In all lawgiving there are two elements: first, a law, which represents an action that is to be done as *objectively* necessary, that is, which makes the action a duty; and second, an incentive, which connects a ground for determining choice to this action *subjectively* with the representation of the law. (MS 6:218)

Here Kant suggests that we can represent the “objective necessity” of the law independently of making that law the “incentive” of our action. We

recognize a law as being objectively necessary – as being a duty for us – and we “connect” this recognition with an incentive as subjective determining ground of action in accordance with that law. The incentive can be either consciousness of and respect for the law itself, or “an incentive other than duty.” This seems to conflict with the autonomy passage in the *Groundwork* cited above, whose claim is (or appears to be) that to make the law one’s maxim of action just *is* to acknowledge its objective necessity, hence moral bindingness for one. The *Groundwork* passage appears to elide the distinction between objective necessity and subjective determining ground emphasized in the *Metaphysics of Morals*. In effect, the claim in the latter work is that we need not make the law our maxim to acknowledge its objective necessity for us:

All lawgiving can therefore be distinguished with respect to the incentive. That lawgiving which makes an action a duty and also makes this duty the incentive is *ethical*. But that lawgiving which does not include the incentive of duty in the law and so admits of an incentive other than duty itself is *juridical*. It is clear that in the latter case this incentive which is something other than the idea of duty must be drawn from pathological determining grounds of choice, inclinations and aversions, and among these, from aversions; for it is a lawgiving which constrains, not an allurements which invites. (MS 6:218–19)

Having started with the suggestion that the domain of morality is wider than the *Groundwork* led one to suppose, having then gone on to suggest that autonomy may not be the principle of moral obligation in general (even if it is still the principle of ethical, self-obligation), Kant appears finally to endorse inclinations – in particular, aversions – as typical subjective determining ground of action in accordance with juridical laws. In one sense, none of this need be very startling: Kant may simply be saying that, in relation to juridical duties, persons typically act from self-interest. This need not mean that self-interest is the ground of juridical obligation. Kant clearly is not saying that: his distinction between law and incentive affirms that the law’s necessitating character holds independently of the incentive to act in accordance with it. But nor does Kant say – and herein lies the perceived problem – that although persons typically act from inclinations in relation to their juridical duties, the ground of these duties’ validity lies nonetheless in each person’s capacity for autonomy. So although Kant does not say that inclinations are the ground of juridical obligation, neither does he say that autonomous willing is. It is the apparent rejection of the capacity for autonomy as the ground of juridical obligation that has led interpreters to conclude that if we take the *Groundwork* argument from autonomy as decisive for Kant’s account

of moral obligation in general, the *Doctrine of Right* cannot belong to his moral philosophy. This is the conclusion of Allen Wood and Marcus Willaschek respectively. For Wood, the *Doctrine of Right* can form part of the *Metaphysics of Morals* only insofar as juridical duties are treated as indirect ethical duties:

Of course, right (*Recht*) along with ethics (*Ethik*), in the context of *The Metaphysics of Morals*, both belong to practical philosophy or “morals” (*Sitten*). Both parts involve categorical imperatives, because Kant holds that juridical duties *as such* are also ethical duties. Insofar as juridical duties are regarded as ethical duties, they can be brought under the principles of ethics. To this extent, it may be correctly said that Kant’s theory of right falls under or can be derived from the principle of morality. That is, this may be said *in so far as juridical duties are regarded not merely as juridical but also as ethical duties*. Considered simply as juridical duties, however, they belong to a branch of the metaphysics of morals that is entirely independent of ethics and also of its supreme principle.<sup>4</sup>

Wood takes this line because of the apparent irrelevance of the idea of autonomy to the *Doctrine of Right*: since “Kantian morality is entirely about enlightened individuals autonomously directing their own lives”<sup>5</sup> the latter cannot form part of Kant’s moral system – not, at least, independently of ethics. Willaschek reaches substantively identical conclusions. For Willaschek, the universal principle of Right as stated in the *Doctrine of Right* cannot be a categorical imperative, so cannot have prescriptive authority for us: “We must distinguish between obeying an imperative and acting in accordance with it. Someone acts *in accordance with* an imperative if she acts as the imperative prescribes. But this may be quite accidental ... *Obeying* an imperative, by contrast, means that one acts as one does because this is what the imperative demands.”<sup>6</sup> Insofar as the Introduction to the *Metaphysics of Morals* explicitly reserves acting from duty for ethical legislation, “juridical laws cannot find expression in categorical imperatives because juridical laws do *not* require obedience for their own sake.”<sup>7</sup> And because juridical laws cannot find expression in categorical imperatives, they cannot be moral laws.

<sup>4</sup> Wood, “Final Form,” 9.   <sup>5</sup> Ibid.

<sup>6</sup> Marcus Willaschek, “Which Imperatives for Right? On the Non-Prescriptive Character of Juridical Law in Kant’s *Metaphysics of Morals*,” in Timmons (ed.), *Kant’s Metaphysics of Morals*, 65–88, 70. See also Marcus Willaschek, “Why the *Doctrine of Right* Does Not Belong in the *Metaphysics of Morals*,” *Jahrbuch für Recht und Ethik* 5 (1997), 205–27.

<sup>7</sup> Willaschek, “Why the *Doctrine of Right* Does Not Belong,” 205–07.

### 3 THE IRRELEVANCE OF AUTONOMY TO THE MORALITY OF RIGHT

A striking feature of the debate over the moral status of Right is its usually exclusive focus on the above-cited passages from the Introduction to the *Metaphysics of Morals* together with some further remarks in the Introduction to the *Doctrine of Right* concerning the relation between Right and coercion. Those whose systematic orientation lies within the *Groundwork* rarely venture beyond these two introductions in reaching their verdicts regarding the non-moral status of Kantian Right. There is something odd about such an approach: one would not normally expect a question of such systematic importance to be settled by a look at the introduction rather than the main body of the text. Why assume, moreover, that if it does not fit the *Groundwork*, the argument in the later work cannot be a moral one? Why not assume, as seems more natural, that if the later *Doctrine of Right* departs in certain respects from the earlier *Groundwork*, this suggests certain developments in Kant's thinking about morality?

There is good reason to be circumspect about the idea of autonomy. Some Kant scholars have argued that the principle of autonomy establishes us as legislators of the moral law in a strong sense – we ourselves will the moral law into existence (albeit necessarily so).<sup>8</sup> Others have been keen to restrict the reach of autonomy even within Kant's ethics: autonomy as self-legislation should be read as free submission under the acknowledged authority of the moral law, where the latter is not itself a possible object of human legislation.<sup>9</sup> Self-legislation means not that we legislate the moral law in the sense of “making law,” but that we self-enforce the independently valid moral law against ourselves. This restricted interpretation of autonomy as freely willed obedience to independently valid moral law is often thought to take the excitement out of Kant's ethics. The magnificence of Kant's ethics is thought to lie precisely in its Enlightenment promise of the human being as maker and arbiter of his own fate,

<sup>8</sup> I take the principal representative of this position to be Christine Korsgaard. See especially, *The Sources of Normativity*, Onora O'Neill (ed.) (Cambridge University Press, 1996).

<sup>9</sup> See especially Onora O'Neill, “Four Models of Practical Reasoning,” and “Agency and Autonomy,” in her *Bounds of Justice* (Cambridge University Press, 2000), 11–28 and 29–49. O'Neill's interpretation is not moral realist: her position is not that there is an independently knowable moral law with the prescriptions of which we seek to bring our maxims into conformity. O'Neill emphasizes judgment where a more voluntaristic reading such as Korsgaard's emphasizes decision-making. On O'Neill's account we judge what we take the moral law to prescribe in any given situation; on Korsgaard's more voluntaristic reading we decide what it is going to be.

including his moral fate. If we merely submit, through “*Selbstzwang*” (TL 6:394), to the authority of a law whose validity holds independently of our autonomous willing, much of the allure of Enlightenment emancipation seems to disappear. It is worth pointing out, however, that the restricted reading makes one’s task easier in relation to the *Doctrine of Right*: if self-legislation is self-enforcement of independently valid moral law – if, as “good-willing,” self-legislation is merely the ethical mode of subjective comportment towards the moral law – the problem of the moral status of Kant’s philosophy of Right is much diminished.<sup>10</sup> We then have, as suggested in the Introduction to the *Metaphysics of Morals*, (our idea of) an independently valid moral law and two possible modes of subjective comportment towards it – the ethical mode of autonomous comportment and the juridical mode of non-autonomous comportment. Again, however, this raises fresh problems. If autonomy is not the ground of the moral law, then what is? To this question, I shall for now simply answer that it is the idea of freedom. In the *Metaphysics of Morals*, freedom is not synonymous with autonomy. Kant there distinguishes between internal freedom and external freedom, confining the idea of autonomy to internal freedom – more of that distinction below. He also distinguishes between *Wille* and *Willkür* – “will” and “power of choice.”<sup>11</sup> *Wille*, as possible determining ground of the agent’s power of choice, is “practical reason itself.” *Willkür* is the capacity “to do or refrain from doing as one pleases” – it is the capacity for self-determining *action* as distinct from self-determining *willing*. *Willkür* determined to action through *Wille* as its determining ground is “*freie Willkür*” – free power of choice. The *Wille/Willkür* distinction chimes with the above distinction between the law’s objective necessity and the subjective determining ground of action in relation to the law. Will as practical reason represents the law as objectively necessary; the agent’s power of choice accepts that law as its determining ground of action either for the sake of duty or on the basis of some other incentive – or fails to accept it as its determining ground entirely. Whether or not, and

<sup>10</sup> The problem does not disappear altogether: irrespective of how one answers the question as to the origin of the moral law, the ethical necessity of self-enforcement implies the impermissibility of other-enforcement. As we shall see below, the *Doctrine of Right* precisely champions the legitimacy of other-enforcement. See section 4 below.

<sup>11</sup> There has been some debate about whether this late distinction creates more problems than it is meant to solve. Compare L.W. Beck, *A Commentary on Kant’s Critique of Practical Reason* (University of Chicago Press, 1960), 176–208; Henry E. Allison, *Kant’s Theory of Freedom* (Cambridge University Press, 1990), 129–35; and Nelson Potter, “Does Kant Have Two Concepts of Freedom?” in G. Funke and J. Kopper (eds.), *Akten des Vierten Internationalen Kant Kongresses* (Berlin: Walter de Gruyter, 1974), 590–96.



on whatever grounds, the agent does accept the authority of the law does not impugn its *objective* necessity relative to him.

An interesting implication of the *Wille/Willkür* distinction as just sketched is the possibility that the two may not always be united in the same agent. *Wille* as practical reason could be located elsewhere than within the agent whose *Willkür* it directs. Where the *Willkür* of an agent is determined by an externally legislating *Wille*, that agent has not willed the law autonomously: he has not himself willed the law as a law for himself. Instead, external *Wille* has willed the law as a law for him; the agent merely directs his power of choice accordingly. In the *Groundwork*, *Wille* and *Willkür* are unified in the same person – they may even still represent a conceptual unity – such that the law’s necessitating character is seen as a function of its subjective endorsement as a law. With the *Wille/Willkür* distinction of the *Metaphysics of Morals*, a *Willkür* that acts in conformity with *Wille* conceived as an external source of lawgiving is non-autonomous. Yet since *Wille* is practical reason, and since the laws that issue from *Wille* are laws of freedom, the agent whose power of choice is so determined acts non-autonomously but nonetheless freely. The agent is non-autonomously – i.e. externally – free.

In the main text of the *Doctrine of Right* Kant invokes the idea of the general united will. This a priori idea of a collective will contrasts with the unilateral wills of individual agents in the *Groundwork*. The general united will is not an *aggregate* of individual wills but represents a multitude of individual wills *unified* into one will under a sovereign head.<sup>12</sup> The juridical authority to make law in accordance with a priori principles of Right is vested in this idea of the unified public will. The sovereign, who makes public law in accordance with the idea of the general united will, pronounces judgments of Right as judgments valid for every constituent member subsumed under that will. As spokesperson for the idea of the general united will the sovereign thus stands to the individual citizen as objective *Wille* stands to subjective *Willkür*: the general united will makes laws in accordance with a priori principles of Right and these laws are non-autonomously legislated laws of freedom which each citizen is subject to. This conceptual separation between the legislative authority of *Wille* and the executive function of *Willkür* together with the stated relationship between law-giving sovereign and law-receptive citizen is crucial

<sup>12</sup> I develop this interpretation in more detail in “Elusive Unity: The General Will in Hobbes and Kant,” *Hobbes Studies* (in press).

to a proper understanding of the distinctly public morality of Right. I shall return to this thought in section 5.

#### 4 THE PUBLIC MORALITY OF RIGHT

In the last section I provisionally employed the *Wille/Willkür* distinction of the *Metaphysics of Morals* to move us from the *Groundwork's* focus on autonomous willing to the later work's concern with possible public willing. The already mentioned distinction between internal and external freedom is of relevance in this context, and I shall now elaborate on it in order to indicate why laws of external freedom *require* legislation through a public as opposed to a private will. The intrinsically coercive character of Right becomes pertinent here.

Internal freedom is autonomy (strong or restricted) in the *Groundwork* sense: that person is internally free who wills acting from duty as his maxim, thereby ensuring the non-heteronomous determination to action of his power of choice. Internal freedom, good will, and moral worth are closely related terms. By contrast, external freedom is the use of one's power of choice through actual choices and actions consonant with relevant laws of external freedom.<sup>13</sup> An externally free agent typically acts from inclination though in accordance with laws of external freedom: though her maxims may well fail to conform with the laws of external freedom, her choices and actions must so conform. In contrast to laws of internal freedom, which constrain actions by way of constraining maxims, laws of external freedom constrain actions irrespective of maxims. From the perspective of the *Groundwork* this sounds puzzling. In the *Groundwork*, the idea of an unconditionally good will is all but synonymous with consciousness of and respect for the moral law *within* us: consciousness of the moral law accounts for capacity for autonomy. By contrast, in the *Doctrine of Right* laws of external freedom enter the picture not through inner moral consciousness but via our unavoidable coexistence with others. Under conditions of unavoidable coexistence, the valid claim to the use of his power of choice by each is limited by the equally valid claims to external freedom of everyone else. The universal principle of Right accordingly makes the possible coexistence of each

<sup>13</sup> Laws of freedom are not extraneous to the idea of freedom for Kant: one is not first free to act, in the sense of being unconstrained, and then accepts certain laws regulating the free actions of all as necessary constraints upon one's freedom. Rather, action unconstrained by laws of freedom is "wild" or "lawless" action and in that sense not so much free as arbitrary.

agent's exercise of their freedom with that of everyone else the criterion of rightness of action:

Any action is *right* if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law. (RL 6:230)

As Willaschek notes,<sup>14</sup> the universal principle of Right is non-imperative – it does not have the form of a categorical imperative. The principle offers a criterion by means of which *anyone* can judge *anyone's* actions juridically right or wrong. Conceivably, given the reference to maxims in the second subclause, the universal principle of Right could serve as a (non-imperative) principle of self-legislation: a person could test the juridical probity of her maxim by employing the criterion of freedom-compossibility. Yet if what matters is not that our maxims come out right but that our actions do, juridical self-testing of maxims is strictly irrelevant. Significantly, the universal *law* of Right, stated a few lines later, is imperative yet omits all reference to maxims:<sup>15</sup>

so act externally that the free use of your choice can coexist with the freedom of everyone in accordance with a universal law. (RL 6:231)

The universal law of Right,

is indeed a law that lays an obligation on me, but it does not at all expect, far less demand, that *I myself* should limit my freedom to those conditions just for the sake of this obligation; instead, reason says only that freedom *is* limited to those conditions in conformity with the idea of it and that it may also be actively limited by others. (RL 6:231)

Here we have both the explicit confirmation of the universal law of Right as a moral law and its explicit repudiation as a self-legislated law. The universal law of Right is valid because “reason says” it is, meaning that the law's restriction is analytic to the idea of freedom itself. This is consistent with what I have called the restricted interpretation of autonomy in the *Groundwork*: the objective necessity of the moral law holds independently of our autonomous subjection under its authority such that I *am* obliged whether or not I make acting in accordance with that obligation my maxim. On both the restricted and the strong interpretation of autonomy, where I fail to make the moral law the maxim of my action no one

<sup>14</sup> Willaschek, “Which Imperatives for Right?”

<sup>15</sup> Oddly, the imperative form of the universal *law* of Right is missed by Willaschek; yet it is the law, not the principle, which Kant takes to be relevant for “strict,” i.e. coercively enforceable, Right.

else can compel me to do so. Good will is non-coercible. This injunction against ethical compulsion is upheld by the universal law of Right: I cannot be compelled to make the universal law of Right my *maxim* of action. Others can, however, “actively limit” my *actions* to the conditions of law-governed external freedom in general. Indeed, it is not necessary for *me* so to limit my actions: the authority to do so rests with *others*. Why should this be so?

Right, Kant holds, is intrinsically and reciprocally coercive.<sup>16</sup> Under conditions of unavoidable coexistence my exercise of my freedom of choice and action necessarily limits yours. Likewise, your exercise of your external freedom necessarily limits mine. The power of choice of one necessarily restricts the power of choice of the other. The relation is coercive in the formal sense of one *Willkür* acting as a constraint upon another irrespective of either’s consent. It is this reciprocally coercive relation between agents’ respective “power[s] of choice” which renders the idea of a public will necessary. Again, the contrast with the principle of autonomy is instructive. In testing the maxim of her action, an agent does not prescribe the moral law to anyone else. The maxim she adopts as permissible determining ground of her action *could* be adopted by everyone else but *need not* be. Others may act on different, equally permissible maxims in relation to the same general duty.<sup>17</sup> Since the categorical imperative is a test of maxims it cannot serve as a *binding* law for anyone other than the maxim-testing agent. The universal law of Right is not a test for maxims: it prescribes a morally necessary action. The ground of that moral necessity is located not in the agent’s inner consciousness of the moral law and related good will but in others’ legitimate claim to external freedom. But in that case, why cannot individual agents who raise reciprocally coercive freedom claims against one another enforce the universal law of Right against one another?

A unilateral will cannot serve as a coercive law for everyone ... since that would infringe upon freedom in accordance with universal laws. Only a will putting everyone under obligation, hence only a collective general (common) and powerful will [can serve as a coercive law for everyone]. (RL 6:256)

<sup>16</sup> I say more about this in “Reason, Right, and Revolution: Kant and Locke,” *Philosophy and Public Affairs* 36 (2008), 375–404. See also Arthur Ripstein, “Authority and Coercion,” *Philosophy and Public Affairs* 31 (2004), 2–35.

<sup>17</sup> Cf. O’Neill, “Consistency in Action” in her *Constructions of Reason: Explorations of Kant’s Practical Philosophy* (Cambridge University Press, 1989), 81–104; and Barbara Herman, “Moral Deliberation and the Derivation of Duties” in her *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), 132–58.

The universal law of Right is a coercive moral law: it authorizes enforcement of the freedom condition against everyone, even against each of their subjective wills. The ground of coercion lies in the validity of everyone's freedom claim. A unilateral will cannot serve as a coercive law for everyone in three related senses of "cannot." First, a unilateral will cannot coercively legislate the universal law to itself – my will cannot bind my will, or rather: my will is at liberty to unbind itself from itself. This is inconsistent with the universal law of Right which is a coercive law for everyone, including myself. Second, my will cannot coercively bind yours. If my will could bind yours, then given innate equality between us (RL 6:231), your will could bind mine. But this shows that, given innate equality of wills, neither can coercively bind the other.<sup>18</sup> Third, to say that any action is right, the freedom of which is compatible with the freedom of *everyone else*, is to affirm the systemic character of rights relations: the actions of any one person affect the possible actions of *everyone else*, not just of this or that identifiable other person. The systemic effects of any agent's particular actions are beyond her control. Relatedly, if rights pronouncements regulate the freedom of *everyone*, every rights pronouncement issued in relation to any particular rights dispute is at the same time a pronouncement that holds for everyone. But a unilateral will lacks necessary systemic dominion. Only a will that is not itself party to any particular rights dispute, so is not itself subject to the coercive law of Right, yet has dominion over systemic freedom relations as a whole can issue coercive pronouncements of Right whose validity for one constituent member of that system entails validity for all members. This is the general united – i.e. public – will. The formal distinction between *Wille* and *Willkür* in the *Metaphysics of Morals* and the substantive argument for the idea of a general united will in the *Doctrine of Right* are thus internally related.

##### 5 JUSTICE WITHOUT VIRTUE

Summarizing the argument across the preceding sections, we can say that to assume that for its argument to have moral status the *Doctrine of Right* must fit the pattern of argument of the *Groundwork* is to assume that the morality of Right is a species of the morality of virtue. Whether or not our rights claims against others are met then depends on others'

<sup>18</sup> Contractual binding is of course possible for Kant, but requires the presence of a coercive public authority as its condition.

good will. This is deeply counterintuitive, and not Kant's view. Agents have valid, coercible external freedom claims against one another. Since each agent is both claimant and addressee of others' valid claims, these freedom-based rights claims are, moreover, systemic in nature. No private will can make coercive rights pronouncements that are valid for everyone: both because a private will necessarily excludes itself from its coercive pronouncements and because it lacks systemic jurisdiction. The constitutive unsuitedness of private willing in relation to the morality of Right entails the necessary idea of a general united will which, while not itself a member of any rights-relation, is so connected to each individual will as to be able to claim dominion over the whole. In the *Doctrine of Right*, therefore, *Wille* and *Willkür* are not "located" in the same agent. *Wille* is located in the sovereign power that makes public law in accordance with the idea of the general united will; *Willkür* is the power of choice of each individual agent to whom *Wille* legislates the laws as determining ground of *Willkür's* actions. In relation to Right, agents are non-autonomous in that they do not themselves will juridical laws as valid laws for them; and yet agents are externally free in that they may do or refrain from doing as they please within the constraints of public laws of external freedom.

I want now to return to the distinction between acting justly and being just. There seems to me to be a connection between the systematic position which holds that unless the categorical imperative serves as its ground juridical obligation in Kant cannot be moral, and the substantive view which holds that for a person to be fully just he must adopt principles of justice as his maxim rather than merely acting in outward conformity with them. As I said, I do not doubt that the person who makes all juridical duties indirect ethical duties for her is virtuous; I only deny that such a person is more just than one who acts in mere outward conformity with laws of Right. Under an Aristotelian scheme, the distinction within law between acting justly and being just makes sense. Where justice is a virtue no person who acts justly without being just can be fully just. In that sense Kant's systematic distinction between Right and virtue as distinct domains of morality merely betrays his non-Aristotelianism. Duties of Right are perfect, externally enforceable duties; duties of virtue are imperfect, self-legislated duties.<sup>19</sup> Right is not a virtue among others – nor is it a cardinal virtue. Right is a moral concept that simply does not fall

<sup>19</sup> There is, however, good reason to believe that the perfect/imperfect distinction does not fit onto the Right/virtue distinction as neatly as Kant seems to want to suggest. Excellent here is Willaschek, "Why the *Doctrine of Right* Does Not Belong."

under the category of “virtue” at all. The reason for this is the constitutive non-self-enforceability of intrinsically coercive rights relations.

However, the current temptation among Kant-inspired liberals to demand inner endorsement of externally legislated principles of Right is not inspired by Aristotelian concerns about the importance of cultivating the virtue of justice for reasons of character. That temptation stems instead from familiar liberal worries about the coercive character of Right. The worry is that unless individual citizens take an active part in the legislative process, controlling it by means of each of their subjective wills, they render themselves vulnerable to the arbitrary exercise of sovereign political power. The requirement among Kant-inspired current liberals that a proposed public law must pass the test of possible self-legislation thus responds to the classic liberal concern to limit arbitrary political power.

I accept the force of the classic liberal worry; however, the remedy proposed by current Kant-inspired liberals strikes me as misconceived. The proposed remedy employs Kant’s ethics of self-legislation as a criterion by which to judge the justness of public laws. Formally, it has us start with a proposed public law. We are each to ask ourselves whether we could adopt a proposed public law as a possible maxim of action for us. If we answer in the affirmative, we judge the law just; if in the negative, unjust. The proposed test treats a public law as though it had the same general structure and scope as a subjective maxim; it assumes, moreover, that the mode of introspective personal evaluation is equally suitable in relation to public law-making as it is in relation to ethical self-legislation. We have seen, however, that coercive public laws are systematically distinct from ethical laws in several respects. A coercive juridical law *cannot* be self-legislated, nor is such a law appropriately evaluated from a private standpoint.

A Kant-inspired liberal may concede that the *Groundwork*’s categorical test for maxims cannot without modification serve as a test for public laws. He may nonetheless insist that the essential insight of Kantian self-legislation is of cardinal importance in relation to the concern to limit arbitrary public law-making. The Kant-inspired liberal will introduce certain modifications to Kant’s categorical imperative test. When judging the justness of proposed public laws individuals must advance public, not private, reasons in support of their judgment. Moreover, individuals’ judgments need not be seen as putative acts of positive legislation that claim validity for everyone. Instead these judgments are best interpreted as negative injunctions against the adoption of particular proposed laws. Nonetheless, moral veto power must lie with each individual citizen subject to coercive public law: if they cannot themselves endorse a proposed

public law as a possible coercive law for them, citizens should not, given the hazards of political power, be obliged to consider themselves bound by it.

The liberal requirement for a connection between inner disposition and outer conformity is thus motivated by concerns different from the Aristotelian ones with which I began. Under the Aristotelian view, justice is a virtue. It follows that the truly just person seeks to make a habit of acting justly, thereby acquiring a just character. By contrast, the liberal rendering of the “being just” requirement is motivated by the thought that unless a person can subjectively endorse a proposed public law as legitimately applying to him that law lacks proper juridical authority for him. The liberal concern is not with virtuous character but with political legitimacy. The Aristotelian view seems to me broadly consonant with Kant’s claim that duties of justice are indirect ethical duties: *if* X is a duty of justice you ought as a matter of virtue to make it your maxim to act from X. But for Kant this is, as I said, a concern of virtue, not of justice, and belongs within the domain of ethics. The liberal view, by contrast, threatens to undermine Kant’s categorical distinction between private and public willing and lawgiving within the domain of Right.

We do, as individual citizens, judge the justness of operative public laws, policies, and public institutions. We judge this law unjust, that policy misguided, this institution incompetent. At times of official failure we think how we would have done things differently and more competently. Kant does not deny the indispensability of individual public engagement to the vitality of a given system of public laws. The question is whether moral judgments relating to public processes should be theorized in the ethical mode. The temptation to fall back into the ethical mode is very strong – especially when palpable social injustices are at stake. In Britain, one such issue is failed child protection. The public shows a strong reaction to cases in which relevant public authorities fail to protect vulnerable children from their (retrospectively predictable) deaths at the hands of parents. The tendency is to adopt the ethical perspective: to imagine how one would have comported *oneself* to ensure prevention of these children’s usually shockingly cruel deaths. Officials’ defense of their public actions or inactions – procedures that had to be followed, evidence to be collated, families to be kept united – are brushed aside in public discourse as irrelevant excuses: everyone knows quite well what his or her (ethical) duty would have been with regard to the cases at hand.

Liberals would agree that, though understandable, the ethical mode of evaluation is inappropriate in such instances. Concerned citizens should



appeal to public reasons in their evaluation of officials' actions, asking themselves whether, in their view, officials followed relevant procedures with sufficient care and discrimination. Since the issue is not about officials' maxims, any inference from what one retrospectively believes one's own maxim would have been to what one believes the officials' maxim should have been is irrelevant as well as being misguided. The critical point here is that to the extent to which the liberal agrees that the object of citizens' public judgments should be how well appropriate procedures were followed, not whether officials adopted the right maxims, a decisive step has been taken away from the applicability in such contexts of any possible categorical imperative test. As a test of moral goodness of maxims, the categorical imperative is simply inapplicable in the domain of Right. Yet the liberal may respond that the above example misconstrues their position. The categorical imperative test may be inappropriate in relation to cases of public judgment of the above kind. But it is nonetheless appropriate in theorizing reasoned citizen agreement on basic public laws – it is relevant in theorizing agreement at the constitutional level. To be able to bind herself to a proposed coercive scheme in general, the individual citizen must be able to ask herself whether she could will its basic legal elements as public laws for herself. Admittedly, the agent is not testing her *maxims* – she is evaluating the justness of proposed public laws. Still, the crucial point from the liberal perspective is the idea of self-binding to proposed public laws by means of personal judgment and assent: a citizen who judges that she could adopt a proposed public law as one that is valid for her is able to consent to its coercive imposition upon her. She will then not act in mere outward conformity with the law, fearful of its coercive sanction in cases of her non-performance, but will think of herself as having had an active part in the legislative process and as having bound herself, through her participation, to that law.

It seems to me a confusion to suggest that a person can legislate a public law to herself alone. Given their systemic reach, public laws are laws that hold for everyone. A person who deems a public law acceptable for herself necessarily deems it acceptable for everyone else. In legislating a public law to herself a person cannot say “this proposed law is endorsable as a public law for me, but I don't know about everyone else.” If she believes she can say this, she treats the proposed law as though it were a private maxim: others then *may* but need not consider themselves bound by it. If, on the other hand, she (correctly) believes that in endorsing the proposed law as a public law for herself she is also necessarily endorsing it for everyone, she is prescribing law to others. This raises the question of

her authority to do so. Kant denies that a unilateral will can make coercive law for everyone, itself included. Only a publicly constituted will can make such law. This does not mean that those subject to the public will are merely passive recipients of its arbitrary directives. Citizens can and should have input in the legislative process. However, they lack authority to *will* proposed laws *as* public law. I remarked above that the Kantian general united will is not an aggregate of individual wills – a multitude of individual wills that stand in a symmetrical relation to one another, each with equal legislating powers. Perhaps the Rousseauan general will has that shape: individuals judge, each by him- or herself, though on the basis of public reasons, which of a set of proposed laws he or she can endorse as valid public laws. Taking away what Rousseau calls the “pluses and minuses” from aggregated individual judgments one arrives at an overlapping consensus: this is the public will. By contrast, Kant’s citizens are neither self-legislators nor co-legislators within the public will. Kant’s idea of the general united will includes a relation between superior and subordinate:

Between the *commander* (*imperans*) and the *subject* (*subditus*) [of the *civil union*] there is no partnership. They are not fellow-members: one is *subordinated to*, not *coordinated with* the other; and those who are coordinate with one another must for this very reason consider themselves equals since they are subject to common laws. The civil union *is* not so much a society but rather *makes* one. (RL 6:306–07)

The civil union results from submission of subjects under the commander; the consequence is equality between citizens but not between citizens and sovereign. If the union consists of commander and asymmetrically related subjects, citizens cannot be co-legislators: if they were co-legislators they would, collectively, *be* the commander. No asymmetrical relation could then obtain between commander and subject. Insofar as the Kantian union requires a commander, Kantian citizens cannot be co-legislators so much as equally subject to the laws issued by the commander. When Kant says that “the legislative authority can belong only to the united will of the people” (RL 6:313), he has in mind the *commander* as legislator of the general united will. This is consistent with the above analysis of the constitutive unsuitability of a unilateral will for public law-making. An aggregate of unilateral wills is not less unsuited to that task than a single unilateral will. Only the unison of unilateral wills *under* one omnilateral will brings about the qualitative transformation required for public, i.e. coercive, law-making. Submission of subjects under the commander as necessary spokesperson for the idea of the general united will does not make them passive recipients of

the latter's arbitrary will. Subjects are constituent if subordinate members of the general united will. They lack coercive law-making authority but have the capacity to participate in the process of public will formation. They can and should ask themselves whether the sovereign could legislate a proposed law as a law coercively valid for all.<sup>20</sup> Where a citizen judges that the sovereign cannot will a proposed law as a public law, she has a public responsibility to alert the sovereign to this effect and to make a case for legal reform. The individual cannot, however, judge a given public law as non-binding for her on the grounds that it is not a law which, in her judgment, the sovereign could not (and therefore should not) have willed as a public law. She cannot do so because she has no public law-making authority. As subordinate member of the public will she remains bound by a law which, in her judgment, the sovereign should not have passed.

#### 6 ACTING JUSTLY AS BEING JUST

The crucial if routinely overlooked difference between Kant's account of citizenship participation and liberal approaches lies in the fact that, for Kant, the citizen asks herself whether, in her judgment, the *sovereign* could have passed a given law as public law *for everyone*. On the liberal account, by contrast, each citizen asks himself whether *he* could have passed the law as public law *for himself*. While Kant's citizen acknowledges the sovereign as appropriate locus of public law-making, the liberal citizen regards himself as (co-)possessing the relevant authority. For the liberal, therefore, that person is just who acts in outward conformity with given public law insofar as he has inwardly endorsed their validity for him. By implication, the person who acts in outward conformity with public laws despite inner dissent from them acts unjustly. For Kant, by contrast, that person is just who acts in outer conformity with a given public law even when she personally judges that law to be morally deficient as public law. Given her personal dissent from the sovereign's judgment she campaigns for reform; however, she continues to act in outward conformity with the unreformed law despite her inner dissent from it. She thus acknowledges the sovereign as appropriate spokesperson for the general united will, hence as appropriate law-making authority. It turns out, then, that there is a sense in which Kant can be said to distinguish, within Law, between acting justly and being just – but it is not the sense

<sup>20</sup> See, most famously, "An answer to the Question: What is Enlightenment?" ([1784] (AA 8), translated by Mary J. Gregor in CE, Practical Philosophy); also "Theory and Practice," section II.

which either the Aristotelian or the Kant-inspired liberal typically has in mind. The Aristotelian requires a fit between external and inner conformity with justice: since justice is a virtue, we ought to make a habit out of acting justly so as to acquire a just character. Where justice is a virtue the private/public distinction is inapplicable. Kant shares the private/public distinction with the liberal tradition more generally. However, the Kant-inspired liberal makes acting justly conditional upon being just in the sense of advancing criteria of inner subjective evaluation as appropriate criteria for judging public laws. On the Kant-inspired liberal account, any public law that cannot serve as a possible maxim of action for me cannot therefore be a just public law for me. My continuing to act in outer conformity despite my inner dissent then amounts to my being and acting unjust(ly). By contrast, Kant holds that anyone who despite their express personal dissent from the sovereign's judgment regarding the moral probity of a given public law continues to act in outward conformity with it (whilst lawfully campaigning for reform) is just. So for Kant that person is just who does not make her outer conformity with a given public law conditional upon that law's suitability as a maxim for her. One who acts in accordance with a public law even whilst dissenting in her private judgment from the sovereign's judgment regarding that law is not passively following the directives of an arbitrarily legislating public will: she has grasped the crucial difference between juridical willing and ethical willing, and the ground of that difference as lying in the public character of the former as opposed to the private nature of the latter.

## CHAPTER 4

# *Kant's innate right as a rational criterion for human rights*

Otfried Höffe

### I INTRODUCTION

Which rights, and even which sets of rights, are to be counted as human rights is a politically as well as a philosophically contentious issue. But the plural has now been settled on. Whether one recognizes only negative freedom rights, as does classical liberalism, or whether, referring back to Georg Jellinek's *System der subjektiven öffentlichen Rechte*,<sup>1</sup> one distinguishes three types of claims a legal<sup>2</sup> subject can make, granting them all the status of human rights – namely personal freedom rights (*status negativus*), rights to democratic participation (*status activus*), and social and cultural rights (*status positivus*) – in every case one speaks of several human rights. Under the heading “There is only one innate right” (MS 6:237:27f.) Kant defends the distinct opposing view. The plural is replaced by a decisive singular.

His *Doctrine of Right* argument is, however, so short that, if only because of its brevity, we cannot expect it to offer the kind of superior clarity that comes only when a thought is explained step by step, various ramifications are considered, and potential misunderstandings are taken into account. In order to understand Kant's excessively concise arguments and the occasional cryptic allusion, we must fall back on other texts, especially the Introduction to the *Metaphysics of Morals* and some of the differential-analytical definitions in the *Doctrine of Virtue*. I have arranged the requisite explanations in six sections. The first discusses Kant's distinction between two basic questions of right (section 2). Thereupon follow reflections concerning the moral concept of right

This chapter was written in German and translated into English by Michael Ludwig, a student of the author's, in consultation with the author and the editor.

<sup>1</sup> Tübingen: Mohr, 1892, 1905.

<sup>2</sup> The adjective “legal” is always used to translate the German “rechtlich” or equivalent compound nouns whenever using “right” would be cumbersome.

and the moral law of right (section 3), the authorization to employ force (section 4), and the turn of phrase “by virtue of his humanity” (section 5). I then analyze how Kant clarifies his thesis of an innate right in four statements, which only appear to refute his singular use of “human right” (section 6). Finally, the conclusion further sharpens the focus already brought to Kant’s thesis by mentioning two quasi-human rights.

## 2 MORAL AS OPPOSED TO POSITIVE RIGHT

The idea of a single innate right rests, as does the entire first part of the *Metaphysics of Morals*, upon a double distinction. Kant differentiates between two fundamentally distinct basic questions – what is “laid down as right” (*Rechtens*) and “what is right” (*Recht*) – and two fundamentally distinct doctrines of right, an empirical and a natural one. These are not, of course, to be understood as two kinds of jurisprudence, a rational one and a positive one, but rather as two conceptions of right. The German term “*Lehre*” does not only mean “doctrine” (in Latin, *doctrina*), that is a teaching, but also “the entire content of a course, that which is taught about a branch of knowledge.”<sup>3</sup> Kant makes use of the rather ponderous terms “doctrine of right,” “natural doctrine of right,” and “doctrine of virtue” to identify the systematic element of each field. The term “doctrine of right” thus designates the systematic aspect of the law, in which we recognize the idea of a “doctrine,” that which can be taught about a field of enquiry. Kant’s addition of “*Ius*” (MS 6:229:6) shows that by “doctrine of right” he means specific laws themselves (line 5). Thus “natural doctrine of right” does not stand for the science of natural right, but for natural right itself: Kant specifically says “*Ius naturae*” (line 13f.). Accordingly, “empirical doctrine of right” does not refer to the empirical science of right but to positive, established law.

Kant calls the primary object of the current “science of right,” positive law, a “merely empirical doctrine of right” (MS 6:230:5). It is concerned with the question of what is “laid down as right” (*quid sit juris*), that is, what the laws in a certain place and at a certain time say or have said” (MS 6:229:22–24).

<sup>3</sup> J. and W. Grimm (eds.), *Deutsches Wörterbuch*, 33 vols. [1854–1971] (photomechanical reimpression, Munich: Deutscher Taschenbuch Verlag, 1984), vol. XII, 554.

Kant is very critical of a positive science of right, which runs the risk of overestimating itself and claiming its own object as absolute. This critique not only points us to an object that is no longer empirical (MS 6:229:13), it also explains that the other kind of right, natural right, is indispensable. That which is most important falls by the wayside when only current or past established laws are dealt with. A doctrine of right, to follow Kant's provocative analogy, is then like a wooden head "that may be beautiful but unfortunately has no brain" (MS 6:230:6).

A brainless head is like a mindless mind, an entity shorn of its essential ability, here: thought. It is this thinking, and its organ, the brain, for which Kant declares his object of study to be responsible. This natural right, that is, the metaphysical concept of right, is to "supply the immutable principles for any giving of positive law" (MS 6:229:14f.). The two terms "any" and "immutable" are important.

A specific legislation such as the codification of the German *Bürgerliches Gesetzbuch* or the French *Code civil* can certainly be based on principles that an empirical science of right could determine. The judgment for which Kant is searching is not to be understood as comparatively but rather as absolutely valid. He means the "strict universality" that characterizes a priori judgments in the first *Critique* (KrV B3f.), thus something that is fundamentally inaccessible to an empirical doctrine of right.

Since antiquity practical philosophy has accepted a strict alternative with respect to the matter of validity: something can be valid either *nomo* (by decree) or *physei* (by nature). One also speaks of empirical or positive validity as well as normative or supra-positive, that is, moral validity. Kant draws on this distinction. For he defines moral right – that is, non-positive, what he calls "non-statutory" (RL §36, 6:296:16), pre- and supra-positive right – as "natural." He also speaks of "natural right" and defines it as the right "that can be cognized a priori by everyone's reason" (line 16f.).<sup>4</sup> Because we are here concerned with something a priori, we require a genuinely philosophical science, a science of reason, which is also called "a metaphysics." In the case of right, it is not a theoretical metaphysics dealing with knowledge, but a practical one dealing with action. A theoretical metaphysics studies aspects of knowledge that are pre-empirically valid; a practical metaphysics by contrast studies impulses that are pre-empirical and that thus also exclude well-being as a determining ground.

As its object has an a priori character as a result of its strict universality, the corresponding principles belong to pure practical reason. They

<sup>4</sup> For the more general concept of moral laws, see MS 6:215:16f.

are above all experience and change and thus, as principles of “any” positive legislation, “immutable” (MS 6:229:14f.). These qualifications apply also to the innate right. It is a strictly pre-empirical and immutably valid right, and thus a right of pure reason. It is, further, a principle of pure legal-practical reason and also a legal-moral principle.

The task that is accomplished for knowledge by the pure forms of intuition, the pure concepts of understanding, and the regulative ideas, is taken over in the field of action by the “moral” “laws of freedom.” Of the two kinds of laws of freedom, juridical and ethical, the *Doctrine of Right* deals only with the juridical, the laws of freedom that are “directed merely to external actions and their conformity to law” (MS 6:214:13–15). They include the innate right.

That which is right according to external laws is called just (*iustum*), and its opposite unjust (*iniustum*) in the Introduction to the *Metaphysics of Morals* (MS 6:224:7f.), whereas in the Introduction to the *Doctrine of Right* (§B) the contrast is between “right” and “wrong.”

The very concept of human rights implies a pre- and supra-positive significance. For they extend not simply to many persons, but strictly and without exception to all persons. To be more precise, they apply to all accountable subjects. Since Kant calls such subjects “persons” (MS 6:223:24f.), we should in fact speak of “person rights” instead of “human rights.” Nevertheless, the usual term does not evidence any species preference, any so-called “speciesism.” In fact, it represents merely an entrenched *pars pro toto*, which simply refers to one kind of accountable subject, the only one we know of: human beings.

Whether they are called “human rights” or “person rights,” because of their pre- and supra-positive meaning they belong in a natural, not an empirical, doctrine of right. Yet the distinction between positive and supra-positive validity remains relevant. For human rights are usually laid down in specific declarations, which, from a methodological point of view, have a merely positive relevance. They are issued by an empirically recognizable authority, and thereby meet Kant’s definition of “laid down as right”: they hold “in a certain place and at a certain time.” Philosophy opposes to this the moral question of a pure legal-practical reason and asks about right and wrong. Kant answers this question with the innate right. The answer claims to offer a rational criterion for a positive declaration of human rights, in short, for human rights. And since it is to be the highest criterion, it necessarily appears as a singular, from which its singular character results. Consequently, Kant already offers such a unique criterion in the section dealing with the moral concept of right (§B “What is Right?”). It consists in the yet-to-be-defined law of right.



Today, Kant's claim that positive right receives a proper brain only through natural or metaphysical right appears excessive. For what jurisprudence would recognize the rational concept of right as that indispensable precondition that "must supply the immutable principles for any giving of positive law"? Kant provides a strong argument, however, which undoubtedly remains convincing. We can think of it in terms of a thought experiment that comes very close to the current legal reality.

Imagine an "external lawgiving" that "would contain only positive laws." Even such a purely positive lawgiving cannot be understood as purely positivistic, for "a natural law would still have to precede it, which would establish the authority of the lawgiver (i.e. his authorization to bind others by his mere *choice*)" (MS 6:224:33–37). In other words, no purely positive legal order can uphold itself; it requires a foundation beyond positive law, a supra-positive foundation.

We can find support for this in Thomas Hobbes, a philosopher who is readily, though too hastily, touted as the father of contemporary legal positivism. In his *magnum opus*, *Leviathan*, in chapter 26, "Of Civil Laws" (§2), the philosopher presents the commanding character of the law by means of a qualification: only that command is a law which is addressed by one person to another "formerly obliged to obey him."<sup>5</sup> It is precisely this previous obligation that has the pre- and supra-positive character demanded by Kant.

A further confirmation is supplied by Hans Kelsen, the great legal positivist of the twentieth century, and a thinker close to the neo-Kantianism of that time. Even if we subscribe to his famous statement that any content can be made into law,<sup>6</sup> the corresponding lawgiver must be authorized to give laws. Whether, as with Kelsen, we trace the final authorization to a fundamental norm or to the people, as in democracies, only a normative – in the end, a moral – justification empowers the relevant authority to give the law.

### 3 ON THE MORAL CONCEPT OF RIGHT AND THE MORAL LAW OF RIGHT

Unlike mathematics, philosophy cannot begin with definitions; its concepts must be extracted from the matter itself. In the case of the moral

<sup>5</sup> T. Hobbes, *Leviathan*, [1651] E. Curley (ed.) (Indianapolis, Ind.: Hackett, 1994), 173.

<sup>6</sup> H. Kelsen, *Reine Rechtslehre*, 2nd edn. (Vienna: Auflage, 1960), 201. For an English translation of this work, see *Introduction to the Problems of Legal Theory: A Translation of the First Edition of the Reine Rechtslehre or Pure Theory of Law*, Bonnie Litschewski Paulson, Stanley L. Paulson (trans.) (Oxford University Press, 1997).

concept of right, and of the moral law of right that ensues, this “matter” has two moments: the moral point of view and the object to be considered from it. The latter is concerned with moral obligation, with legitimacy as opposed to positive validity, to (juridical) legality; the former with the conditions to which moral obligation must submit, the question of that for which law is responsible.<sup>7</sup>

Because of this moral obligation, the concept of right we seek is dependent on a metaphysics, in this case a metaphysics of morals. This metaphysics concerns the objects of experience, and in the case of right, human beings, which is why it also requires an anthropology. Thus the legitimatory strategy can be titled “metaphysics plus anthropology.” Both moments fulfill two fundamentally different but complementary tasks. The metaphysics is responsible for the moral perspective, the anthropology for the conditions of its application. For “a metaphysics of morals cannot be based upon anthropology but can still be applied to it” (MS 6:217:6–8).

Kant’s *Doctrine of Right* is basically concerned with that coexistence of responsible beings (persons) that conforms to the point of view of morality. If we turn to Kant’s criterion of universal law to describe this point of view, we then come to the famous formula in §B of the *Doctrine of Right*: “Right [morally understood] is therefore the sum of the conditions under which the choice of one can be united with the choice of another in accordance with a universal law of freedom” (MS 6:230:24–26).

Because its terseness can lead to ambiguity, the phrase warrants a few comments. The first concerns the matter of its scope. A moral concept of right usually evokes the idea of a standard that distinguishes the morally legitimate from the morally illegitimate. But the moral concept of right could also maintain that to structure human coexistence according to right is morally legitimate, perhaps even morally required. In the first, more modest, case the concept provides a norm for right; in the second it also or instead legitimizes right. Though Kant does not make this distinction himself, the interpretation to follow will show that his moral concept of right fulfills both tasks.

<sup>7</sup> Prior to this chapter, see O. Höffe, *Kategorische Rechtsprinzipien. Ein Kontrapunkt der Moderne* (Frankfurt am Main: Suhrkamp Verlag, 1995), ch. 5; in English: *Categorical Principles of Law: A Counterpoint to Modernity*, M. Migotti (trans.) (Pennsylvania State University Press, 1990). See also “Der kategorische Rechtsimperativ. Einleitung in die Rechtslehre,” in O. Höffe (ed.), *Immanuel Kant, Metaphysische Anfangsgründe der Rechtslehre*, Klassiker Auslegen XIX (Berlin: C.H. Beck-Verlag, 1999), 41–62.

Further, when Kant speaks of a “law of freedom” in the moral concept of right, he seems to take back the distinction between right and disposition that is constitutive of his theory of right. But by “freedom” he does not here mean freedom of the will, that is, morality or the disposition towards right. In the Introduction to the *Metaphysics of Morals* he had already distinguished two types of (freedom-) lawgiving “with respect to the incentive”: “That lawgiving which makes an action a duty and also makes this duty the incentive is *ethical*. But that lawgiving which does not include the incentive of duty in the law ... is *juridical*” (MS 6:219:1–6). Accordingly, he explicitly explains in §C, which follows upon the definition of the concept of right, that one need not make the principle of right one’s maxim (6:231:3ff.). It is only ethics which makes the extensive demand for a disposition towards right which recognizes the principle of right “out of duty,” in which Kant refers us to the second part of the *Metaphysics of Morals*, the *Doctrine of Virtue*. He qualifies the term “law” with “freedom” inasmuch as a moral law, being distinct from natural laws, becomes real only if it is freely accepted.

Third, the qualification of the law as “universal” must not be thought of as explicative. That a law containing no proper names has a certain level of universality is too obvious for it to have to be stated twice (MS 6:230:22 and 25f.). With this universal characteristic Kant is “only” reminding us of his general moral criterion: the moral concept of right fits into his program of a universal morality. The moment of the universal moreover silently overrides personal and collective well-being as determining grounds. The object, a moral concept of right, is at the same time assigned to a genuine metaphysics of morals, that is, not a theoretical metaphysics but one that is practical and concerned with motives.

Fourth, it should be clear that the anthropology that extends the moral point of view is not of an empirical nature, such as that which Kant excludes in his Preface to the *Groundwork*. The anthropology of which he is speaking there pertains to motives and, as an empirical concern, stands in opposition to the practical metaphysics he is looking for. The anthropology included in the concept of right, an anthropology of right, consists in nothing else than the conditions of application of right: how multiple persons can coexist.

The moral concept of right requires that the anthropological aspect of right be organized according to the principle of universality. The moral concept thus corresponds to the legitimacy strategy of “practical metaphysics plus anthropology.” Within that strategy, metaphysics restricts itself to the genuinely moral aspect, the moment of strict universality.

As Kant includes no further metaphysical elements, the metaphysics contained in the moral concept of right is not problematic.

The mutual compatibility of freedom of action, in which consists the concept of right, has – fifthly – a negative side, which Kant does not mention because it was all too obvious to him: that each person’s freedom of action must be restricted. This restriction applies factually before a constitutional right is founded, even before private right, before an external mine and yours. This restriction of freedom that is pre-constitutional and precedes even property rights has a scope that is sometimes overlooked in Kant interpretation, and even more often within the human rights debate: the human right that follows from the moral concept of right concerns a pre-constitutional restriction of freedom. It does not concern a lawgiver or the writers of constitutions, but rather “natural subjects of right,” that is, those persons who do not have any public legal order but who act with and against one another within a shared environment. Public legal relationships are, with respect to this “state of nature,” certainly indispensable yet of only secondary importance (see RL §§41–42).

We tend to begin the story of human rights with medieval concepts of freedom, with those documents whereby kings granted rights to secular and religious estates, beginning with the Magna Carta Libertatum (1215) and the Golden Bull of Andrew II of Hungary (1222). But from a philosophical point of view – and this is where the greater emphasis lies – the history of human rights begins much earlier. At the latest it begins when specific actions are declared to be crimes, and actually even earlier, namely as soon as any right existed at all, making human beings into persons with inalienable rights and corresponding inescapable duties.

This state of affairs is also to be found in the relation of human rights to the state. The preceding account of rights undermines the widely held conception that human rights consist in defensive rights against the state. In truth, the fundamental human right might be endangered neither only nor even primarily by the state. Two points are nevertheless correct. First, the idea of negative rights to freedom is at the source of human rights. Second, they are definitely defensive rights. From a systematic point of view, they do not only, or even primarily, defend against the state, but rather against fellow subjects of law. Moreover, rights and duties always appear together at this level: the innate right consists in the legal authority to impose a duty upon all others (MS 6:237:18f.).

As already mentioned, however, Kant does not explain why freedom should be restricted at all. But the following argument suggests itself. Because people live in the same environment, two things can happen

at any point in time: that someone wants to settle where someone else already lives, and that someone wants the same thing, a good or a service, as another. Wherever the one or the other happens, there is *eo ipso* a conflict. And above all, freedom of action is restricted simply because this conflict situation persists, regardless of how any particular conflict is resolved – be it violently or peacefully, whether someone gives in or a compromise is found.

This argument, which expands on Kant's own considerations, corresponds to the thought experiment of an original state of nature. Accordingly, this systematically first restriction of freedom is not a moral phenomenon, but rather a contingent "natural given" that necessarily results from the coexistence of persons.<sup>8</sup> As soon as several persons share the same environment, no one can make claims on a living space without thereby restricting everyone else's possible living spaces. Pointedly: where there is one world but several persons who influence one another's actions to some extent, they will unavoidably restrict one another's freedom.

The sixth explanation is that Kant himself raises only the second aspect of the restriction of freedom, the positive aspect that is of a morally legitimate kind: that one person's choice "can be united" with the choice of another (MS 6:230:26). If freedom is strictly and equally restricted, it is also thereby protected. Restricting and protecting freedom are two sides of the same process. Only when freedom is limited by a universal law is it equally protected for everyone.

The principle of right (§C) follows upon its moral concept. It addresses the same issue but from a different perspective. The moral concept of right deals with an objective right, whereas the moral principle of right defines the corresponding subjective right. Here Kant's human right can already be discerned. Lastly, because Kant understands subjective right as the entirety of all actions which objective right permits (MS 6:230:29–31), he establishes a singular and not a plural human right. The principle of right constitutes the moral standard by which to measure all subjective claims as permitted by moral right.

Because these claims arise prior to and independently of positive legal acts, pre- and supra-positive rights, that is, innate rights or human rights, do indeed exist. Kant's moral principle of right is therefore equivalent to a, actually to *the*, criterion of human rights. Kant himself says as much in the

<sup>8</sup> Cf. O. Höffe, *Politische Gerechtigkeit. Grundlegung einer kritischen Philosophie von Recht und Staat* (Frankfurt am Main: Suhrkamp Verlag, 1987); in English: *Political Justice: Foundation for a Liberal Philosophy of Law and the State* (Cambridge: Polity Press, 1995).

section “There Is Only One Innate Right”: “Freedom (independence from being constrained by another’s choice), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right belonging to every man by virtue of his humanity” (MS 6:237:27–31).

The moral concept of right corresponds to the idea of justice, an idea that has been differentiated into various types since Aristotle.<sup>9</sup> Because of the reciprocity of the restrictions on freedom, the justice that Kant deals with is neither distributive justice nor corrective justice, but commutative justice (justice in exchange). Coexistence is not, from a systematic point of view, primarily a matter of distributing goods and services but rather consists in an exchange between and among persons who live in the same place. This basic exchange does not pertain only to economic goods. One of the most important human rights, the protection of life and limb, consists in an exchange of restrictions on freedom: each gives up the right to kill another in case of conflict, whereby each then *eo ipso*, and not only as a consequence thereof, receives a subjective right to life and limb.

Kant finally adds a third variation on the moral right, the universal – and again, moral – law of right. That which the concept of right and the principle of right already dealt with is here formulated as a moral, that is, a categorical, imperative. The law of right has the rank of a, actually of *the*, categorical imperative: “so act externally that the free use of your choice can coexist with the freedom of everyone in accordance with a universal law” (MS 6:231:10–12). Inasmuch as the core statement concerning the universal compatibility of freedom of action is already contained in the concept and in the principle of right, they too, in fact, are both equivalent to categorical imperatives of right.

#### 4 AUTHORIZED COERCION

Whoever provides someone with a legal authorization thereby generally states that everyone else is legally prohibited from hindering the authorized action (or omission). The same applies to the special case of human rights. For example, one has a right to life and limb only when everyone else is legally barred from using physical force. Conceptually, legal authorization includes a second-level authorization which gives one the right to enforce that which is required or prohibited. A moral concept of right is therefore incomplete as long as it does not deal with legitimating authorized coercion.

<sup>9</sup> See *Nicomachean Ethics*, S. Broadie (ed.), C. Rowe (trans.) (Oxford University Press, 2002), book V, 158–76.

We may reproach the most influential theories of justice of the last few decades for a deficiency in this respect. For example, in *A Theory of Justice*, John Rawls derives the principles of justice from primary social goods.<sup>10</sup> But he does not provide a satisfactory answer to the question as to why these goods can generate claims that may then be enforced if necessary.

Rawls' opponent Robert Nozick makes note of this deficiency, yet he also fails to provide a convincing answer to the question of how to justify the authorization to coerce compliance with basic principles of justice. Kant, however, sketches an argument that promises to remedy this justificatory deficit. He does not claim, as does Nozick,<sup>11</sup> that an authorization to coerce can be justified only in two steps. Instead, he sees in authorized coercion an element in the definition of any moral-legal obligation. He carries out a single-level justification by arguing that authorized coercion is "connected with (subjective) right by the principle of contradiction" (§D; cf. §E and MS 6:233:34f.).

One could justify authorized coercion from an antinomy, namely the opposition between a positivistic thesis, according to which a legal order is given a blank check which makes the use of coercion everywhere legitimate, and an anarchistic antithesis that rejects any use of coercion. The antinomy would be resolved by a justification of coercion that simultaneously restricted its scope. Legitimation would thus be linked to a limitation, the standard for which would consist in Kant's moral criterion of strict universality.

This kind of argument would better reveal the characteristic features of the legitimacy task. It would revisit the considerations pertaining to the conditions of application of right, the thought experiment of a primitive state of nature. Both sides of the antinomy of right contradict themselves. Contrary to philosophical anarchism, the thought experiment points to a first level of coercion: as soon as several persons share the same environment, a coercion comes into play that is primary from a systematic point of view, which operates behind everyone's backs and is answerable to no one.

The thought of a reality principle opposed to a pleasure principle originates with Sigmund Freud.<sup>12</sup> Since this reality principle has its systematic origin in a primary social coercion, Freud is wrong when he attributes

<sup>10</sup> J. Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971).

<sup>11</sup> R. Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 131ff.

<sup>12</sup> See S. Freud, "Formulierungen über die zwei Prinzipien des psychischen Geschehens," in *Studienausgabe*, 8th edn. (Frankfurt am Main: S. Fischer, 1997), vol. III, 13–24; in English: "Formulations on the Two Principles of Mental Functioning," in *The Case of Schreber, Papers on Technique and Other Works*, Standard Edition of the Complete Psychological Works of Sigmund Freud, vol. XI (1911–13), J. Strachey (ed.) (London: Hogarth Press, 1958), 218–26.

responsibility for it to culture. This inescapable social coercion follows not from culture but simply from humans' coexisting in the same environment. That is why it has been emphasized ever since antiquity that human beings' social nature has a "dark side." As a consequence, one should not emphasize simply the positive aspect of social nature. Without just as one-sidedly emphasizing the negative aspects, however, one should view humans' social nature as a judge would, rising above an exclusively optimistic or a purely pessimistic perspective. Human sociability is a burden as well as a pleasure.

The suggested thought experiment must now be carried out with respect to unavoidable social coercion. The concept of a moral claim of right disintegrates when subjective right consists in an *arbitrary* authorization of coercion as required by strict legal positivism. Whatever one claims for oneself – whether rights to property or life and limb – can just as well be promised to another and again claimed by a third. And if there is no coercion at all, as per the anarchistic antithesis, then one is dependent upon others' goodwill to have one's subjective rights recognized. As a consequence, one can speak at best of quasi-rights; no hard claims of right can be made either for life and limb or property.

The suggested thought experiment now conforms to both sides of the moral concept of right, and further expands upon what it claims: mutually ensuring freedom can only be accomplished by also mutually restricting it. By not carrying the thought experiment through, Kant fails to explore all the possibilities of a critical theory of right.

Kant sets up his actual argumentation in section C and carries it out in sections D and E. The very formal argumentation aims at an authorization of coercion that is *immanent* in the concept of right. In non-Kantian words: to the moral right of one person corresponds a moral duty of everyone else because one person's right to a universally compatible freedom corresponds to others' duty to be content with that universally compatible freedom. And because both sides – right and duty – are juridical and not only ethical, they include an authorization of coercion.

The term "coercion" is often taken to refer to physical violence. There are however plenty of other types of coercion, and not only direct coercion, but also indirect coercion, such as that which results from social dependence or economic poverty. Since Kant does not describe the broad spectrum of possible kinds of coercion, one often assumes he is employing a restricted concept that is reduced to physical violence. In fact it remains open in Kant's text whether the coercion is physical or economic, whether it can be easily perceived or instead is hidden, and whether it



applies directly or indirectly. And Kant is well advised to avoid these questions, as they do not affect the basic argument. The decisive question is rather whether, and if so, how far, coercion is morally permitted. Kant's answer to this question is as simple as it is convincing: every use of coercion is legitimate inasmuch as it opposes an illegitimate use of coercion, the wrong.

The answer is aimed at the antithesis of the antinomy of right, the idea of anarchy. It quietly rejects a strict anarchism. "Coercion is permitted," says Kant. It is, however, justified only under two restrictive conditions, with which he contradicts the thesis advanced by a strict legal positivism. First, coercion is permitted only where it has already been used, namely where another is intruding into my legitimate space. Legitimate coercion does not attack but only protects itself; it is not an aggressive but a defensive coercion, a counter-coercion. Second, within the purview of defensive coercion, only that coercion is legitimate which opposes a wrong. Otherwise a thief would have a moral right to prevent his victims from recovering their goods. The thief is opposing the use of coercion, but though his action is defensive, it is also illegitimate; it does not prevent a wrong, but rather reinforces it.

There are two guises, a preventive and a restitutive one, under which morally legitimate coercion, resisting a wrong, appears. If a robbery is underway, one is allowed to prevent it; if it has already taken place, one is permitted to retrieve the stolen goods. This argument tacitly links the argumentation for authorized coercion with its limitation. Be it preventive or restitutive, a counter-coercion is justified only when opposing a wrong. Whoever prevents a robbery but also intentionally wounds the thief, and whoever retrieves more than what was stolen, is himself committing a wrong.

Kant employs the terms "hindering an effect" and "resistance that counteracts" a hindrance (MS 6:231:24f.) to justify coercion in the sense of resisting a wrong. Both concepts fit under the higher concept of a "practical negation." "Hindrance" points to a simple negation of action, "resisting a hindrance" to a double negation of action.

When an action is morally legitimate, hindrance is a simple negation, a moral wrong (cf. MS 6:230:32–34). Conversely, a wrong is nothing other than the hindrance of a legitimate free action (6:231:26f.). Whoever negates the negation reinstates the original position, as Kant rightly states. To resist a wrong is to reverse it and thereby acknowledge the right: "if a certain use of freedom is itself a hindrance to freedom in accordance with universal laws (i.e. wrong), coercion that is opposed to this (as a hindering

of a hindrance to freedom) is consistent with freedom in accordance with universal laws, that is, it is right" (6:231:28–32).

Kant is here simply arguing with the moral concepts of right and wrong and the logical operation whereby a double negation brings you back to the original position. That is exactly why authorized coercion is contained in the concept of a principle of right, and why Kant can claim his legitimation is analytical: "there is connected with right by the principle of contradiction an authorization to coerce someone who infringes upon it" (lines 32–34). Because authorized coercion belongs in the definition of an authorized right, that which is permitted by right includes a second-level permission, namely to enforce that which is permitted at the first level: "Right and authorization to use coercion therefore mean one and the same thing" (MS 6:232:29).

Since the categorical imperative of right results from applying a moral perspective to an empirical fact (the coexistence of responsible subjects), and since authorized coercion is analytical with respect to the categorical imperative of law, authorized coercion also fulfills the programmatic aim of a Kantian theory of right. The authorization is justified through reason alone as concerns the empirical datum. Like the categorical imperative of right, the authorized coercion of right is an a priori right, albeit a relational one, as it is based upon an anthropology of right. The argumentative framework "practical metaphysics plus anthropology" stands.

##### 5 "BY VIRTUE OF HIS HUMANITY"

The hardest part to understand in Kant's thesis of an innate right is the turn of phrase "by virtue of his humanity." The German "*kraft*" means both "by virtue of" and "as a result of." The significance of the argument is therefore clear. Humanity is the ground of the innate right. But what is to be understood by "humanity"?

The term is familiar from the *Groundwork*. There it does not designate the human species, but the nature of human beings, their "rational nature" (G 4:429:2). It is claimed that it exists as "an end in itself" (line 3), which leads to the famous second, "material" formulation of the categorical imperative: "So act that you always use humanity ... always at the same time as an end" (lines 10–12).

The term "humanity" does not appear in the Introduction to the *Metaphysics of Morals*. Nor is it to be found in the Introduction to the *Doctrine of Virtue*, though Kant clarifies there many concepts that are important for the entire *Metaphysics of Morals*. Yet it appears in the

Introduction to the *Doctrine of Right*, in both of its sections, the “General Division of Duties of Right” (section A) and the “General Division of Rights” (section B); it then receives its most extensive definition in the “Division of the Metaphysics of Morals as a Whole.” In this last passage (MS 6:239:23–28) the meaning is clear.

Under “humanity,” Kant understands the human being inasmuch as he can be represented “in terms of his capacity for freedom.” This capacity is “wholly supersensible,” so that humanity is represented as a “personality independent of physical attributes,” the “*homo noumenon*.” The “*homo phaenomenon*” is the complement, a human being with physical attributes. Of Kant’s two points of view, the noumenal and the phenomenal, the term “humanity” belongs to the moral one since “the humanity in his own person” is nothing else as the human being “regarded merely as a moral being” (MS 6:429:4f.). This point of view is naturally not missing from Kant’s Introduction to the *Metaphysics of Morals*. It is to be found in at least two places: in moral personality, which is “nothing other” than “the freedom of a rational being under moral laws” (MS 6:223:25f.), and again when the human being is regarded “as a noumenon ... merely as an intelligence” and the opposition between sensible and intelligible beings is discussed (6:226:19ff.).

The *Groundwork*’s humanity-as-end-formula requires that you 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” (G 4:429). It therefore applies to duties to oneself (“in your person”) just as to duties towards others (“in the person of any other”). We expect that right deals with the relationships between persons. Surprisingly, in the *Doctrine of Right* Kant explicitly speaks first of “the right of humanity in our own person” and bases upon it a duty to oneself: “Do not make yourself a mere means for others but be at the same time an end for them” (MS 6:236:27f.). It is that duty to rightful honor with which Kant explains the first Ulpian formula: “*Be an honorable human being (honeste vive)*.”<sup>13</sup>

This – clearly ethical – duty is astonishing inasmuch as it requires “asserting one’s worth as a human being in relation to others” (line 25f.). This is, in my opinion, a novel type of duty that Kant here adduces in his philosophy of right and which he himself only fully understood later on.

<sup>13</sup> For a commentary on the three Ulpian formulas, see O. Höffe, “Kategorische Rechtsimperative nach Ulpian” in *Königliche Völker*, Zu Kants kosmopolitischer Rechts- und Friedenstheorie (Frankfurt am Main, 2001), ch. 7; in English: “Categorical Imperatives of Right According to Ulpian,” in *Kant’s Cosmopolitan Theory of Law and Peace*, A. Newton (trans.) (Cambridge University Press, 2006), 119–33.

In Vigilantius' notes on the "Lecture on the Metaphysics of Morals" in the winter semester 1793–94, Kant still assigns the "*honeste vive*" maxim to the counterpart to the section on right, that is, to the "ethical duties" that are distinguished from "duties of right" (V 27:527). This new kind of duty, which is also new within Kant's own philosophical development, is one of the elements that shows Kant's theory of right to be both revolutionary and provocative.

The new duty consists in a basic, not physical but rather legal-moral self-assertion. Kant places it at the core of a rightful honor in contrast to which the usual understanding of rightful honor – that one has committed no legal wrong – sounds almost banal. Kant is concerned with an honor that is of fundamental importance for a theory of right. It is of constitutive importance because it is a necessary condition of a relationship to right and consists in self-recognition, not in being recognized by others.

There are two ways for a primary recognition to establish a legal personality. You become a legal entity by being recognized either by others or by yourself. Kant rejects the first position, which is a fundamental legal paternalism, whereas he agrees with the second, just as fundamental a legal liberalism. Right begins with a negative personal contribution. Only those who refuse to be simply instrumentalized and fundamentally degraded in matters of right become legal entities.

Kant does not claim that self-recognition precludes the need for an initial recognition by others. He considers self-recognition primary, however. Only those who constitute themselves as legal entities are suitable subjects of legal relationships. Why? Only they are able to claim a space within which to be free, so only with respect to them can any restriction or any symmetrical granting of freedom be spoken of. According to its definition as "the capacity for putting others under obligation" (MS 6:239:19f.), right presupposes the duty to constitute oneself as a moral – actually only as a legal-moral – entity.

In Kant's table of duties the "right of humanity in our own person" is understood, as a duty, under three headings (MS 6:240). The first two are self-explanatory: under "duty of right," the duty in question is a duty of right, not of virtue; under "perfect duty," it is valid without exception, as a perfect rather than imperfect duty. The third designation is, however, surprising, even irritating. This is where the revolutionary novelty and provocation are to be found. Though right is concerned with one's relationship to others, it originates not with a social relationship, but with a relationship to self – instead of with a "duty to others," with a "duty to

oneself." In the case of right, a human being is first of all "accountable to the humanity in his own person" (MS 6:270:19f.). This is the place to cite the following from the *Doctrine of Virtue*: "Humanity is itself a dignity; for a human being cannot be used merely as a means by any human being (either by other or by himself)"; and with this dignity he raises himself "above all other beings in the world that are not human beings" (TL §38, 6:462:21ff.).

Two further contrasts reinforce and clarify the unusual character of a duty of right towards oneself. For one, Kant opposes a "right of humanity" to a "right of human beings," which amounts to a duty towards others. For another, it is a duty of right not of virtue.

By moving from rightful honor (section A) to an innate right (section B), the "right of humanity in our own person" takes on the meaning that a human being as a practically rational creature, and as a moral creature, has the right to be considered as a legal entity in relation to other human beings. He must, however, acknowledge this right; and to do so he must refuse the legal degradation mentioned above. Only he who meets this fundamental requirement of rightful honor and never allows himself to be degraded to a mere means, he alone establishes himself as a legal entity and acknowledges that which "belongs to everyone by nature, independently of any act that would establish a right" (MS 6:237:22f.); and only he has claim to a freedom that can be shared by every other in accordance with a universal law. The word "independently" here indicates that we are dealing with an inner and not an outer mine and yours, that it is therefore innate and not acquired, hence original and not somehow subordinate.<sup>14</sup>

## 6 FOUR IMPLICIT HUMAN RIGHTS

Kant is not satisfied with the freedom formulation of the innate right. He explicitly includes an explanation as to what is already implied by the freedom formulation. Kant adduces four rights, which are just as valid as, because included in, the innate right. Each corresponds to a human right, which creates a plural out of the singular after all. There is a multiplicity within the one human right, but not in addition to it: the one innate right can be spelled out in four innate rights.

<sup>14</sup> One discrepancy should be mentioned. According to the "General Division of Rights" something can only be internally mine or yours, since something external that is mine or yours must be acquired (MS 6:237:24–26). Nevertheless Kant will, in the course of the section on private right, use the name for "an external belonging, though an ideal one" (RL §35, 6:295:7f.).

These implied rights clarify, as does the innate right to freedom, the meaning of the two concepts of freedom and equality that usually appear at the beginning of a declaration of human rights. They are, if only for that reason, more basic than many elements of the standard catalogue of human rights. Kant refers neither to property, happiness, and safety as the “Virginia Bill of Rights” does, nor to property, safety, and resisting oppression as in the French “Déclaration des droits de l’homme et du citoyen,” nor to life, liberty, and security of person as the “Universal Declaration of Human Rights” has it. What is more, he does not include human rights such as freedom of belief, conscience, and religion, or freedom of expression and of the press, or artistic and scientific freedom.

The four implicit human rights follow in a systematic order. Kant thereby tries as far as possible to include everything that the one human right implies. To do so, he begins with the most obvious implication and works his way toward more distant ones.

The first implication is clear: inasmuch as every human being is entitled to a well-defined freedom by virtue of his humanity, all must be considered as equals. This innate equality is strictly equivalent to freedom. The first implicit human right declares: no one has more obligations than anyone else; everyone is legally permitted to do whatever falls within the purview of the reciprocal restrictions on freedom.

The “hence” introducing the next statement points to a logical consequence. The next implication does not follow directly from innate freedom, but from what it implies, namely innate equality, which is why Kant deduces it as a second step: a human being’s quality of being “his own master” – that is, according to the concept of thinghood (*realitas*) in Kant’s first *Critique*.

It is best to read this cryptic formulation in light of the concept of a person. In the Introduction to the *Metaphysics of Morals*, Kant defines the person as a responsible subject (MS 6:223:24f.). This person, he continues, is “subject to no other laws than those he gives to himself” (lines 29–31). He adds two options in an aside which can be associated respectively with the *Doctrine of Right* and the *Doctrine of Virtue*: one gives the laws “either alone,” which is to be associated with the duties of virtue, “or at least along with others,” which pertains to duties of right.

In contrast to those “human beings without a [legal] personality,” slaves and those in bondage, whoever is his own master has a legal-moral authorization to order his own life. He claims all the rights that have their origin in the innate human right. This is why the second implicit human

right stipulates that every human being has the right to be a legal entity, a person with rights, and not a slave or a man in bondage.

The third implicit human right, which is likewise introduced as a consequence of the first, consists in the right to be beyond reproach. This right only superficially appears to resemble the criminal law principle of presumed innocence; in fact, it is much more basic. As long as a human being has not carried out a legal act, he has committed no wrong against anyone. As Kant explains: as long as one does not do anything legally relevant, one cannot be at fault with respect to anyone. For example, whoever violates his duty to help the needy or more generally to be charitable has indeed acted wrongly, but not in a legal sense. The “spirit of brotherhood,” which the first article of the UN “Universal Declaration of Human Rights” requires that human beings show toward to one another, therefore does not have the rank of an innate human right for Kant.

According to the fourth implicit human right, one is permitted to act in any manner towards others as long as one does not curtail their rights, their inner or outer mine and yours. Kant includes an example that is surprising at first sight. It permits that which the philosopher usually forbids, to “tell an untruth,” that is, to lie.

We can refer back to the “Division of a Metaphysics of Morals” to understand what is meant: “Accordingly the giving of the law that promises agreed to must be kept lies not in ethics but in *Ius*” (MS 6:220:1f.). Kant’s explanation does not restrict this duty, but instead simply indicates that were “external constraint” absent, as is the case in ethics, “the idea of duty by itself would be sufficient as an incentive” (line 4f.). Unlike in the fourth implicit human right, Kant does not make any restrictions here with respect to the command to keep one’s promises. He says simply: “It is no duty of virtue to keep one’s promises but a duty of right, to the performance of which one can be coerced” (lines 11–13).

Kant is uncompromisingly strict in the *Doctrine of Virtue* (§9): “A lie,” defined as any “intentional untruth in the expression of one’s thoughts” (MS 6:429:7f.), “annihilates” one’s “dignity as a human being” (line 23f.). To this rigorous verdict he however appends a more specific and restrictive supplement, namely “in the ethical sense of the word” (6:430:1).<sup>15</sup> Kant thereby plays on the distinction between ethics and *Ius*, and more pointedly explains that an intentional untruth “in the Doctrine of Right”

<sup>15</sup> Cf. MS 6:429:9, “in ethics.”

deserves the “harsh name” of a lie only “if it violates another’s right” (6:429:9).<sup>16</sup>

Kant therefore strictly distinguishes between an untruth that is relevant to right and one that is not. He explains in the *Doctrine of Right*, in a similar way to his exposition in the *Doctrine of Virtue*, that one may speak insincerely and untruthfully, and even make promises, so long as one does not deprive another of what is his (MS 6:238:32). He, and only he, is guilty of lying in a legal sense who “violates another’s right” (MS 6:429:9).

## 7 CONCLUDING REMARKS: TWO QUASI-HUMAN RIGHTS

Even though Kant considers there to be only one innate human right, two further pronouncements are relevant to human rights. They correspond to both sections of the *Doctrine of Right*, the “Private Right Concerning What is Externally Mine or Yours” and the “Right of a State.” Kant’s philosophy of human rights thus covers his entire doctrine of right. Though his explicit philosophy of human rights consists “only” in the theory of “what is internally mine and yours” outlined above, it implies that (external) private right and the right of a state also pertain to human rights.

We cannot go into detail here. Yet these two pronouncements should at least be mentioned and their relation to human rights indicated. Kant bases his theory of property, which comes before the right of states and is thus termed “private right,” neither on empirical or historical, nor anthropological or pragmatic reasons. Nor does he claim that property rights have existed for a long time, or that they exist in almost all cultures. He does not justify property either by reference to biological specificities of the human species or by to personal or collective benefits. Property right in the sense of what is externally mine and yours, in its three modes of objects, services (“contract right”), and personal status (“matrimonial, parental, and domestic right”) is justified according to Kant by no other reason than an argument of pure practical reason that pertains to the conditions of application of right, namely the external freedom of multiple persons living in the same world.

According to this basis in reason, Kant introduces the concept of intelligible ownership (RL §1) and then asserts as a legal postulate of [pure]

<sup>16</sup> Cf. MS 6:430:2–4 where a harmful lie is a “violation of the rights of others.”



practical reason that any law according to which “an object of choice would *in itself* (objectively) have to *belong to no one* (*res nullius*)” is “contrary to rights” (MS 6:246:7f.)<sup>17</sup> as it contradicts the moral concept of right. A few lines later we read: “It is therefore an a priori presupposition of practical reason to regard and treat any object of my choice as something which could objectively be mine or yours” (lines 32–35).

Kant's argumentation, which can here be only indicated, begins with the counter position and asks whether it can be maintained along with the legal-moral concept of tolerance, with “the freedom of everyone in accordance with a universal law.” Now he explains that there can be “for practical reason none other than formal laws.” That is why one cannot distinguish legitimate from illegitimate objects in the case of merely potential property. As a consequence, all objects are either to be permitted or prohibited. But an absolute prohibition suspends one task of the moral concept of right, namely to allow external freedom, in this case, to pursue your own chosen goals. So the only alternative remains to allow all objects without restriction to be claimed as property: nothing may become “*altogether* (objectively) *ownerless* (*res nullius*).”

Given the a priori character of this thesis, it is a strictly pre- and supra-positively valid statement, with the inalienable status that corresponds to the level of a human right. Kant is not here establishing a human right to property. He is neither declaring that every human being has an innate right to a certain title of ownership, nor claiming that every human being has a right to at least own something. Hence, I am not speaking of a human right, but of something that is as significant as a human right.

This thesis is addressed to a legal order, certainly a private and not a public legal order. From a legal-moral point of view, this legal order may not prohibit any objects of human choice from becoming something that is mine or yours. Positively put, every human being has a right to live in a legal order that permits every object of human choice to become mine or yours.

There is also a quasi-human right in the second half of Kant's *Doctrine of Right*. It consists in the right of a human being to live in a public legal order. Kant had already referred to a public right three times in the section on private right. These anticipatory remarks, which appear in all three

<sup>17</sup> In the CE *Metaphysics of Morals*, the passages cited here can be found at MS 6:250. The author is following AA ordering and Gregor departs from it. See her “Translator's Note on the Text of the *Metaphysics of Morals*,” in Immanuel Kant, *Practical Philosophy*, M.J. Gregor (ed. and trans.) (Cambridge University Press, 1996), 355–59, esp. 355–57.

main sections (§§8–9, §15 and §§41–42) are not evidence of poor composition. Rather, they show that simple private right, in each main section, at first only reaches the level of a state of nature in which something can be mine or yours only provisionally; it is only in a public legal order that something can be mine or yours preemptorily.

This state, however, exists only wherever others enter into it. That is why there is a duty, as the “General Division of Duties of Right” already had it, to “*enter* a condition in which what belongs to each can be secured to him against everyone else” (MS 6:237:7f.). This state has three dimensions<sup>17</sup> as we learn in “Section II, Public Right”: a human being has a right to a state, specifically to a republic (RL §§43–49: “The Right of a State”), the right to a relationship of right between states (§§53–61: “The Right of Nations”) and to a cosmopolitan right to visit, but not to settle (§62: “Cosmopolitan Right”). All three possible kinds of public right require that a preemptory state of right ensures that the innate right is not merely provisionally but rather preemptorily acknowledged.

<sup>17</sup> Cf. ZeF, second paragraph.

*Intelligible possession of objects of choice*

B. Sharon Byrd

## I INTRODUCTION

Kant's main problem when discussing private law in the *Doctrine of Right* is justifying why individuals have a right to have external objects of their choice as their own. The original innate right to freedom from coercion by others does not give us the right to have external objects of choice as our own. The original innate right to freedom is the right we have to move around without restriction with our own bodies, our bodies being inextricably attached to ourselves. Why is it that we have a right to things external to ourselves? Indeed not only a right to pick up and use them, but a right to put them down and expect no one to take or use them when we leave them?

Kant's inquiry is more far-reaching than perhaps first meets the eye. Kant asks his questions not simply for physical things, but for any object of choice. An object of choice is something external to me that I have the ability to use, namely: (1) a physical thing (including animals), (2) someone else's freedom of choice to perform an act, and (3) someone's status in relation to me (RL 6:247:18–21).<sup>1</sup> Kant's discussion thus includes not only rights to property but also contractual rights and family rights.<sup>2</sup> Kant divides his discussion of rights to external objects of choice into one general part and three specific parts. The general part is devoted to explaining what it means to have something as mine. The three specific parts are devoted to explaining the nature of a property, a

<sup>1</sup> The translations of the Kant texts are my own. I am proceeding according to the AA ordering of the RL sections and not according to *Metaphysische Anfangsgründe der Rechtslehre*, 2nd edn., Bernd Ludwig (ed.) (Hamburg: Meiner, 1998), which Gregor uses in the Cambridge Edition.

<sup>2</sup> The group of rights to persons akin to rights to things also includes a variety of rights in relationships where the right-holder acts through another person, such as the mandate or any fiduciary relationship. See B. Sharon Byrd and Joachim Hruschka, *Kant's Doctrine of Right: A Commentary* (Cambridge University Press, 2010), ch. 12.

contractual, and a family (or similar) right and how one acquires these rights.

In this chapter, I map out Kant's arguments on how to have something as one's own. His arguments turn on his concepts of possession. These concepts move from an empirical concept of possession as having something in one's hand to an intelligible concept of possession as having something as one's own based on a duty all others have not to interfere with what one intelligibly possesses (section 2). His arguments depend on the postulate of practical reason. This postulate has been interpreted to provide a justification. A justification, however, suggests that what would otherwise be wrongful or prohibited conduct is rightful conduct because of the situation. My position is that there is nothing wrongful about taking something external to oneself and calling it one's own. The taker thus does not need any justification for doing so. In this chapter I rely on an alternative interpretation of the postulate as a power-conferring norm. The postulate empowers us to have external objects of our choice as our own (section 3). Although we may unilaterally impose an obligation on all others to respect what we have declared to be ours, this power flows from our will being compatible with the universal united will. Nothing in Kant's arguments for individual rights to have objects of choice as one's own depends on the existence of a state. Indeed Kant notes that without a right to property and other objects of our choice there would be no duty to move to the civil social order. Property rights therefore are rights we have in the state of nature. They do not depend on social approval any more than our right to freedom of choice in general depends on social approval and recognition. The sole purpose of the state for Kant is securing rights we already have before leaving the state of nature and moving to the civil state. The state secures our right to freedom and our rights to external objects of our choice (section 4).

## 2 KANT'S CONCEPTS OF POSSESSION

Kant has three concepts of possession: (1) empirical physical possession, (2) possession as a pure concept of the understanding, and (3) intelligible possession. Empirical physical possession is possession with detention. I physically possess an apple I have in my hand (RL 6:247:28–29). If someone grabs the apple out of my hand, they violate the *internal* mine, namely my freedom (RL 6:248:1–4). By grabbing an apple out of my hand a person hinders my freedom, and any resistance I then make to the grabbing

would be compatible with the freedom of all, according to the principle of contradiction.<sup>3</sup> Nonetheless, the empirical concept of physical possession does not go beyond the idea of freedom from coercion or the freedom from violation of the *internal* mine.

Possession as a pure concept of the understanding (RL 6:253:6–7) abstracts from conditions of space and time and focuses on my having something under my physical control (RL 6:253:7–11; 6:268:15). I possess something in this sense when I lock and leave it, such as the house I live in. Possession as a pure concept of the understanding is an abstraction from the empirical concept of possession. Still, the physical aspect of possession remains because the possessor has the thing under her physical control (RL 6:253:11).

Intelligible possession is based on duty, namely a duty others have to leave what I intelligibly possess alone. Intelligible possession is purely legal possession. It does not depend on physical contact with the thing, as does empirical physical possession, or on having it under one's control, as does possession as a pure concept of the understanding. Kant's main question in the *Doctrine of Right* is: how is it possible to have something in one's intelligible possession (RL 6:249:30–32)? How is it possible to have an object of choice as one's own? To solve this problem, Kant begins with the postulate of practical reason, to which I turn in the next section.

### 3 THE POSTULATE OF PRACTICAL REASON

Kant introduces the postulate early in his discussion of how one can have something as one's own:

It is possible to have any external object of my choice as mine, i.e. a maxim according to which if it were law an object of choice must become in itself (objectively) masterless (*res nullius*) is wrongful. (RL 6:246:5–8)

The argumentation in support of the postulate is a *reductio ad absurdum*. Kant considers external objects of choice as objects we have the physical capacity to use. If we did not have the *legal* capacity to use those objects “then freedom would rob itself of the use of its choice in regard to objects of choice by placing *usable* objects beyond any possibility of *use*” (RL 6:246:13–16). Kant calls the postulate a “permissive law (*lex permissiva*)” because it gives us an:

<sup>3</sup> Hindering a hindrance to freedom is compatible with freedom in accordance with universal laws, i.e. right (RL 6:231:30–34).

authorization which we cannot derive from pure concepts of right, namely to impose an obligation on everyone else they otherwise would not have to refrain from using certain objects of our choice because we were the first to take those objects into our possession. (RL 6:247:1–6)

Two aspects of Kant's argument must be emphasized. The first is that having the physical capacity to use an object of choice means that its use is within what Kant calls one's "power" (*Macht*). Having something in one's power does not mean possessing that thing, but instead being able to possess it through an act of choice (RL 6:246:25–32). I have an apple in my power, although not under my control, if the apple is lying on the ground – ground that belongs to no one – and I am closer to the apple than anyone else. Through an act of choice – walking to the apple and picking it up – I can possess the apple in either of the two meanings of physical possession Kant defines. If I can possess the apple, I can use it because "the subjective condition of the possibility of use in general is possession" (RL 6:245:11–12). Still, as Kant notes, it is sufficient to conceive of the object as an object of my choice if I have the object in my power to possess and use.

The second aspect of the argument deserving emphasis is that Kant's postulate and argumentation proceed from the proposition that a maxim prohibiting the use of a usable object of choice is *wrongful* (*rechtswidrig*). This aspect of the postulate and argument deserves special emphasis because the vast majority of Kantians continue in the presumption that the permissive law provides a justification for otherwise wrongful conduct.<sup>4</sup> Kant, however, defines a situation in which the conduct is simply not wrongful. That situation too has been misrepresented in the literature. It is the situation in which my taking the apple would violate no

<sup>4</sup> See e.g. Reinhard Brandt, "Das Erlaubnisgesetz, oder: Vernunft und Geschichte in Kants Rechtslehre," in R. Brandt (ed.), *Rechtsphilosophie der Aufklärung: Symposium Wolfenbüttel 1981* (Berlin: Walter de Gruyter, 1982), 233–85, esp. 244, 255; Reinhard Brandt, "Das Problem der Erlaubnisgesetze im Spätwerk Kants," in O. Höffe (ed.), *Immanuel Kant: Zum ewigen Frieden* (Berlin: Akademie Verlag, 1995), 69–86. Although not using the word "justification" Brandt states "Something that is prohibited 'per se' is provisionally permitted and thus required to let the legal claim to prevention not become effective." ("Es wird etwas 'an sich' Verbotenes provisorisch erlaubt und damit geboten, den Rechtsanspruch der Verhinderung nicht wirksam werden zu lassen."); Katrin Flikschuh, "Ist das rechtliche Postulat ein Postulat der reinen praktischen Vernunft? Zum Endzweck der Kantischen Rechtslehre," *Jahrbuch für Recht und Ethik* 12 (2004), 299–330, esp. 316; Katrin Flikschuh, "Freedom and Constraint in Kant's *Metaphysical Elements of Justice*," in S. Byrd and J. Hruschka (eds.), *Kant and Law* (Hants: Ashgate, 2006), 87–108, esp. 102–04, arguing in line with Brandt that the permissive law justifies what otherwise would be a violation of the universal principle of Right; Wolfgang Kersting, *Wohlgeordnete Freiheit* (Berlin: Walter de Gruyter, 1984), 133–34. Brandt's thesis is discussed exhaustively and refuted in Joachim Hruschka, "The Permissive Law of Practical Reason in Kant's *Metaphysics of Morals*," *Law and Philosophy* 23 (2004), 45–72, esp. 62 n. 48 and 64 n. 54.

one's freedom of choice. Thus the taking could not be a violation of the universal law of Right and needs no justification.<sup>5</sup> Indeed, preventing me from taking the apple or any other object of my choice would itself be a violation of the universal law of Right.

The argument that the permissive law provides a justification for taking an external object of my choice is based on the assumption that Kant's permissive law in "Toward Perpetual Peace" is the same as the permissive law expressed in the postulate of practical reason in the *Doctrine of Right*. A simple review of the context in which Kant discusses permissive laws in "Toward Perpetual Peace" indicates that it is far from the context in which Kant introduces the postulate in the *Doctrine of Right*. In "Toward Perpetual Peace" the permissive law is a law that tolerates a wrongful state of affairs until that wrongful state can be gradually reformed and brought into line with the requirements of law. The first mention of the permissive law in "Toward Perpetual Peace" is in connection with the preliminary articles. Kant indicates that preliminary articles 2, 3, and 4 contain permissions to postpone full execution of their commands until a more suitable date (ZcF 8:347:20–24). Kant also gives us an example of what he means in his discussion of the second preliminary article, which states: "No independent state (small or large being here irrelevant) shall be acquirable by another state through inheritance, exchange, purchase, or gift" (ZcF 8:344:14–16). As Kant points out, this prohibition applies with full force to any acquisition, making the acquisition of one state by another simply wrongful. Nonetheless, the permissive law tolerates an already existing state of possession (*Besitzstand*) as long as this possession is gradually discontinued. The purpose of the permissive law here is to allow the acquiring state to take its time in restoring the acquired state's freedom and as such the permissive law justifies the wrongful state of possession for this limited period of time.

Similarly, when Kant discusses the permissive law a second time in "Toward Perpetual Peace," he does so in connection with a wrongful state of affairs that can be corrected gradually rather than immediately. One example Kant gives is of a state with a despotic constitution that is strong with respect to enemy states. Kant says that the despotic state cannot be expected to immediately lay its constitution aside and run the risk of being swallowed up by these enemy states. Instead, a permissive law allows the despotic state to delay reforming its constitution until a better

<sup>5</sup> The universal law of Right is: "Act externally so that the free use of your choice could be compatible with everyone's freedom according to a universal law" (RL 6:231:10–12).

time for such dramatic change (ZeF 8:373:2–7). As to the permissive law, Kant states:

These are permissive laws of reason that let the situation of a public law marred by injustice remain for so long as complete transformation of everything either comes by itself or matures through peaceful means because any *juridical* constitution, albeit only in a small degree in accord with right, is better than no constitution at all, which fate (anarchy) would meet with too *rash* of a reform. (ZeF 8:373:27–32)

Notable about both of Kant's discourses on the permissive law in "Toward Perpetual Peace" is that they provide what might be called a justification for a temporary period of time to right a wrongful situation. I say "might be called a justification" because today one normally thinks of a justification as making an exception to a primary norm of conduct, such as "Thou shalt not kill!" to permit killing another human being in self-defense or defense of others. A justification, in other words, is black and white. What is wrong becomes right because of the circumstances and it stays right for all time. Kant would most likely call a justification such as self-defense simply a hindrance to freedom which itself is compatible with freedom (RL 6:231:28–34) and not a prohibited action in need of justification. Kant's permissive law in "Toward Perpetual Peace" seems to be more of an excuse for taking one's time in correcting some injustice. What is wrong stays wrong but tolerable because of the expediencies of the situation.

Regardless of whether the permissive law in "Toward Perpetual Peace" is a justification or an excuse, in the *Doctrine of Right* the permissive law fulfills neither of these functions. As Kant says, "a maxim according to which if it were law an object of choice must become in itself (objectively) masterless (*res nullius*) is wrongful." It is thus not wrongful for me to see an object of my choice as potentially mine or someone else's. Indeed, the very opposite is true, namely it is wrong to regard external objects of choice as necessarily no one's. "[I]t is a legal duty to act toward others such that the external (usable) can become someone's own" (RL 6:252:13–15). I need no justification for fulfilling a legal duty. If I am obligated to treat and regard external objects of choice as potentially someone's then *not* fulfilling the obligation would need justification.

Does it not violate someone else's freedom of choice when I take and use an unowned apple thereby depriving others of doing the same?<sup>6</sup> The

<sup>6</sup> Important to bear in mind is that Kant's postulate permits one to impose an obligation on everyone else only if one was the *first* to take possession of the thing according to the principle "prior in time, stronger in right" (*prior tempore, potior iure*) (RL 6:259:14–17).



answer is no. If I am closer to the apple than anyone else then I have the apple within my power to take and use. The others do not have the apple within their power because they cannot get to the apple before I do. The apple is thus not an object of their choice. They may *wish* they were closer to the apple than I am, but in fact they have no choice with regard to the apple. As Kant states:

The faculty of desire according to concepts, insofar as the ground determining it to act lies within itself and not in the object, is called a faculty *to do or not do as one pleases*. If it [this faculty] is connected with the awareness that one's action can bring about the object [of one's desires] then it is called *choice*. (RL 6:213:14–19)

If it is not so connected it is called “wish.” A right, to the extent it relates to a corresponding obligation, does not mean the relation of choice to another's wish, but instead to another's choice (RL 6:230:11–15). Thus my right to take and use the apple does not depend on another's wish to do the same, and if I am closest to the apple I can take it and use it and regard it as mine without violating anyone else's freedom of choice with respect to the apple.

Another pitfall for misinterpreting Kant seems to lie in Kant's claimed assumption that taking an unowned object does not violate anyone else's freedom of choice. It has been argued that Kant simply assumes what he sets out to prove. The problem for interpretation lies with an incorrect English translation of the German text. After introducing the postulate, Kant makes his *reductio ad absurdum* argument claiming that freedom would rob itself of the use of its choice with respect to objects of choice by placing these objects beyond the possibility of any use, thus making them *res nullius*. The problematic passage follows: *obgleich die Willkür formaliter im Gebrauche der Sachen mit jedermanns äußerer Freiheit nach allgemeinen Gesetzen zusammenstimmte* (RL 6:246:17–19). In the Cambridge Edition of the Works of Immanuel Kant, Mary Gregor translates the passage as: “even *though* in the use of things choice *was* formally consistent with everyone's outer freedom in accordance with universal laws.”<sup>7</sup> The German *zusammenstimmte* can be the past tense of *zusammenstimmen*, as Gregor translates it, but it can also be subjunctive, in which case the clause should read “even *if* ... choice *were* formally consistent ...” Since the rest of the sentence is in the subjunctive, which Gregor also translates into the subjunctive, the correct solution here would have been in favor

<sup>7</sup> M.J. Gregor (ed. and trans.), Kant, *Practical Philosophy* (Cambridge University Press, 1996), 405 (emphasis added).

of the subjunctive. This passage in its incorrect English translation has prompted authors to claim that Kant is assuming what he is supposed to be proving, namely that using external objects of choice *is* consistent with others' external freedom.<sup>8</sup>

Kant's argument is far more subtle. He is claiming that an *absolute prohibition* against using external objects of choice would make their use wrongful for any conceivable object of choice, regardless of whether someone could use the object without violating anyone else's freedom of choice. Certainly there are some objects of choice that people can use without violating anyone else's freedom of choice – such as the unowned and unclaimed apple most proximate to me. If there are, then it cannot be wrongful to use objects of choice, assuming I do not have to violate anyone else's freedom of choice by doing so. Therefore, Kant concludes that no absolute prohibition against using objects of choice can be compatible with freedom. Again, the taking and using of such objects of choice is *not wrongful* and thus in no need of justification.

Another problem with the common interpretation of a permissive law as a justification for engaging in wrongful conduct is that this interpretation would limit the applicability of the permissive law to property claims.<sup>9</sup> Kant, however, discusses the permissive law in connection with any object of choice whatsoever. Kant thus intended the permissive law to apply not only to having external things as one's own, but also to having someone else's choice under a contractual agreement as one's own and to having someone else's status in relation to oneself as one's own. The secondary literature focuses on property claims, arguing that such claims are different and thus need justification through the permissive law and through social approval. Yet, all claims to external objects of choice are covered by the permissive law, meaning that one would need a justification for marrying someone or for entering into a contractual agreement in cases in which neither the marriage nor the agreement are in any way a violation of the parties' or others' freedom of choice.

<sup>8</sup> Kenneth R. Westphal, "Do Kant's Principles Justify Property or Usufruct?" *Jahrbuch für Recht und Ethik* 5 (1997), 141–94, esp. 141, 153–54.

<sup>9</sup> Brandt, "Das Erlaubnisgesetz," 260. Brandt argues that the permissive law distinguishes between two classes of objects of choice: (1) things and (2) someone's choice or status. He claims that only things can be masterless (*herrenlos*). Therefore he concludes that the part of the permissive law "i.e. a maxim according to which if it were law an object of choice must become in itself (objectively) masterless (*res nullius*) is wrongful" applies only to things. Gregor translates *herrenlos* as to "belong to no one," which can apply to any of the three objects of choice. I prefer "masterless" because it corresponds more closely to the German.

That the permissive law applies as well to contractual and to family claims is obvious from Kant's placement of the permissive law in what one can call the "general part" of the external mine and thine in §§1–9 of the *Doctrine of Right*. The general part applies to all three specific parts that follow, namely to property claims (§§10–17), to contractual claims (§§18–21), and to family claims (§§22–30). Furthermore, Kant refers expressly to "objects of choice" in the postulate, and not to "things," and otherwise in the *Doctrine of Right* he never confuses or conflates these two terms. Finally, a contractual claim to someone's choice to perform an act which constitutes performance under a contract indeed does remain masterless if no one can accept that person's offer. One might think that the person's choice remains his own,<sup>10</sup> but without any acceptance of his offer he is not master over his choice to perform the act. Indeed, he would not be permitted to perform an act that is unacceptable to his promisee. Imagine that *A* offers to sell his cobra to *B*, who gingerly refuses the offer. *A* is not master over his choice to sell the cobra to *B*, because *B* has no desire to have the snake and will not pay for it. *A* cannot even give the cobra to *B*, unless *B* is willing to accept it. *A* may *wish* to give the cobra to *B*, but as noted above law is not concerned with wishes, but only with choice. *A* has no choice and thus cannot be master over his choice to give *B* the snake. The same is true of someone's status in relation to another person. A person's status as a husband depends entirely on someone else's acceptance of him as her husband. A man's status as a husband in a marriage relationship remains masterless until someone enters into a marriage with him. Consider *C*, who would dearly love to be *D*'s husband, but *D* is not interested. *C* cannot relate to *D* as a husband. *C* has no choice with respect to being *D*'s husband and therefore is not master over his status as *D*'s husband.

Reinhard Brandt has suggested that the permissive law in the *Doctrine of Right* functions to "justify" the coercion permissibly used under the postulate of public law to force all others to move together with oneself to civil society (*bürgerliche Gesellschaft*).<sup>11</sup> Although Brandt mainly refers to the Vigilantius lecture notes and "Toward Perpetual Peace" to support this theory, one could read §8 of the *Doctrine of Right* as adding another permissive law to the one expressed in the postulate of practical reason:

*Corollary:* If it must be legally possible to have an external object as one's own, then the subject must be permitted to *coerce* everyone with whom dispute arises

<sup>10</sup> Brandt, "Das Erlaubnisgesetz," 260.

<sup>11</sup> Brandt, "Das Problem der Erlaubnisgesetze," 74.

regarding the mine and thine over the object to move with him to a civil constitution. (RL 6:256:14–18)

To avoid confusing the §8 corollary and the §2 postulate, one must consider the differing aims behind the two permissions. The permissive law in §2 permits one to impose an obligation on all others to refrain from using objects of choice that one has first claimed as one's own. The obligation is the foundation for the right to have external objects of choice as one's own because without it one could take and use an object of choice but not be able to demand return of that object from anyone who took it when it was not under one's (physical) control. The right to demand return of one's own external object of choice from any other possessor, however, is the crux of intelligible possession and it is the postulate that enables one to attain this right.<sup>12</sup> Furthermore, one has permission to impose this obligation in the state of nature, and imposing it does not depend on being in, or moving to, a civil state.

In contrast, the permission in §8 allows one to coerce all others to move to a civil constitution. The reasoning behind §8 is that one can obtain *security* in one's intelligible possession of external objects of choice only in a civil state. The permission in §2 is the foundation for the corollary in §8 because without §2 no one has intelligible possession of any object of choice and thus has nothing to secure. Kant's reason for allowing coercion to force the move to civil society is not to establish rights to external objects of choice, but instead to *secure* these rights. Indeed Kant states that without the external mine and thine there would be no duty to leave the state of nature and enter a juridical state (RL 6:313:5–8), and thus no postulate of public law. Without acquisition of (acquirable) rights, a juridical state would be impossible (RL 6:312:34–36). Kant refers to provisional rights to external objects of choice in the state of nature as being preliminarily as opposed to peremptorily secured possession (RL 6:292:28–30). Provisional possession is “comparatively legal possession” (RL 6:257:14–19). Property must first be “determined and specified,” because any “guarantee presumes the thine of someone (for whom it is secured)” (RL 6:256:27–31). These provisional rights become peremptory, meaning secured, in civil society.<sup>13</sup>

Under the postulate of *public law* everyone has a duty to move to a civil constitution: “From private law in the state of nature proceeds the

<sup>12</sup> See the real definition of intelligible possession of external objects of choice in section 4.

<sup>13</sup> My interpretation of the §8 corollary is supported by Kant's discussion immediately thereafter in §9 of the difference between provisional and peremptory possession (RL 6:256:19–257:19).

postulate of public law: in a situation of unavoidable contact, you should leave this state [the state of nature] with all others and move to a juridical state, i.e. the state of distributive justice” (RL 6:307:8–11).<sup>14</sup> This duty is a legal duty owed to others. Because the duty is a legal duty, its fulfillment can be coerced by those to whom the duty is owed. Furthermore, this coercion is permitted under the corollary because the duty is imposed to secure individual rights. These rights are provisional in the state of nature, but moving to a juridical state secures them and makes them peremptory rights. Refusing to move to a juridical state, meaning not fulfilling one’s legal duty under the postulate of public law, is equivalent to waging an attack on those with whom one is in unavoidable contact. This attack is a hindrance to freedom and thus wrong. As such it may be hindered under the principle of contradiction and is right (RL 6:231:30–34). One needs no permissive law to coerce all others to move to a juridical state, because the coercion one exercises is not wrong. The right to exercise this coercion follows from Kant’s proof of the possibility of intelligible possession and is thus contained in the *corollary* in §8.

An alternative interpretation of the permissive law is that it is a power-conferring norm.<sup>15</sup> This interpretation has support in the natural-law tradition surrounding Kant. Both Christian Thomasius (1655–1728) and Gottfried Achenwall (1719–72) use this concept of a permissive law. Thomasius, in discussing Grotius’ question<sup>16</sup> of whether permissive laws exist or are self-contradictory,<sup>17</sup> writes:

A permission is not an act of law, because the person who permits something does not issue a norm. If one understands a “permission” to be the affirmation of a right someone else has or the introduction of such a right, then at first glance that may seem different. Yet even then the permission is not a new act of law, but is already included in a prohibitory norm. *As long as parental power or citizens’ property is allowed by the law, then as a consequence it is prohibited to disturb anyone who is exercising his right.*<sup>18</sup>

<sup>14</sup> For the reasoning behind the postulate of public law, which lies in the presumption of badness, see Byrd and Hruschka, *Kant’s Doctrine of Right*, ch. 9.

<sup>15</sup> Hruschka, “The Permissive Law.”

<sup>16</sup> Hugo Grotius, *De Jure Belli ac Pacis* [1625], B.J.A. de Kanter-van Hettinga Tromp (ed.) (Leiden: Brill, 1939), Lib. 1, Cap. 1, §9, 34 (emphasis added): “A permission is not an act of law but rather the negation of such an act, *unless the law obliges someone other than the person who has been given the permission not to interfere with that person.*” Grotius’ comment following “unless” corresponds precisely with Kant’s idea that acquiring an external object of choice imposes an obligation on all others not to interfere with the acquirer.

<sup>17</sup> Kant considers this question in ZeF 8:347:34–348:10.

<sup>18</sup> Christian Thomasius, *Fundamenta Juris Naturae et Gentium*, 4th edn. [1718] (repr. Aalen: Scientia, 1963), Lib. 1, Cap. v, §VI, 146–47 (emphasis added). Similarly in Johann Georg Walch, *Philosophisches Lexicon*, 4th edn. [1775], Justus Christian Hennings (ed.) (repr.

Thomasius thus recognizes permissions that affirm or introduce certain rights.<sup>19</sup> For Thomasius, as for Kant later, these rights include parental power and citizens' property. Parental power and citizens' property are *introduced* by a permissive law, just as the postulate of practical reason in §2 of the *Doctrine of Right* introduces everyone's capacity to be the owner of things, claimant under a contract, a spouse or parent. Furthermore, Thomasius draws the same conclusion as Kant later does, namely that the permissive law reveals itself in prohibitions against others that they not interfere with the rights acquired under the authorization contained in the permissive law.

Kant's ideas resemble Achenwall's<sup>20</sup> even more closely: "A law is called a 'permissive law' because through such a law the lawgiver *awards the faculty* to commit a certain act as a permitted act."<sup>21</sup> Achenwall expressly calls a permissive law a law that attributes a faculty to persons. Similarly, in §2 of the *Doctrine of Right*, Kant says that the permissive law gives us an authorization (*Befugnis*). Kant defines *Befugnis* as a moral faculty (*facultas moralis*) (RL 6:222:29). Kant also calls the permissive law the "postulate of the faculty" (*Postulat des Vermögens*) (RL 6:268:25).

Interpreting the permissive law as a power-conferring norm has the benefit of making the Kant text comprehensible. As a power-conferring norm, the permissive law gives us the capacity to be the owners of external things, claimants under a contract, and members of a family with rights by virtue of the family relation. In other words, the permissive law as a power-conferring norm covers all three possible objects of choice, not by providing a justification for otherwise wrongful conduct, but by extending our freedom of choice to include the freedom to have external objects of choice as our own. As such, the permissive law is the logical consequence of Kant's *reductio ad absurdum* argument. If we did not have

Hildesheim: Olms, 1968), vol. 11, col. 1708, who in his article "Zulassung" (license) says that one could also "conceive of" the "introduction" of a right, "when, e.g. parental power, citizens' property, etc. are licensed by the law."

<sup>19</sup> Although Thomasius says that the law does not contain a permission, he does indicate that a permissive law can be seen in norms prohibiting others from interfering with someone who is exercising a right he has by virtue of the permission.

<sup>20</sup> That Achenwall strongly influenced Kant is unsurprising. From the summer semester of 1771 to the winter semester of 1789–90 inclusively, Kant announced his lecture entitled *Ius Naturae secundum Achenwall* (or similarly) fourteen times. See M. Oberhausen and R. Pozzo (eds.), *Vorlesungsverzeichnisse der Universität Königsberg (1720–1804)* (Stuttgart-Bad Cannstatt: frommann-holzboog, 1999), vol. 11, 318, 340, 374, 388, 402, 416, 430, 444, 466, 473, 500, 524, 547, 566.

<sup>21</sup> Gottfried Achenwall, *Prolegomena Iuris Naturalis*, 2nd edn. (Göttingen: Bossiegel, 1763) §90, 90.

this empowerment, external usable objects of choice would remain unusable, and freedom would rob itself of freedom even in cases in which our using an external object of choice violated no one else's freedom.

#### 4 INTELLIGIBLE POSSESSION OF OBJECTS OF CHOICE

By virtue of the permissive law I can treat any object of my choice as potentially mine, assuming I have the object within my power to possess and my taking of the object into my possession would not violate anyone else's freedom of choice. Important to realize about Kant's argumentation is that he is considering possession of external objects of choice from an abstract viewpoint. Within the Kantian framework no one yet owns or possesses anything. The situation is similar to Robinson's when Friday first arrives on the island. The question is whether Robinson or Friday can have anything on the island as their own and the answer to this question comes from practical reason's pondering it abstractly. With Robinson and Friday, the only objects of choice available are the land and anything upon it, and the counterpart's free choice to perform some act. Land and anything upon it are the only things one can acquire originally<sup>22</sup> and they are thus the logical starting point for the inquiry.

Kant's sole prerequisite for being permitted to take an external object into one's possession is that no one else's freedom of choice is violated through the taking. No one else's freedom of choice is violated if the object belongs to no one and I am the first to take it into my possession (RL 6:247:6). Taking something into my possession can mean that I hold the thing in the sense of empirical physical possession or that I have the thing under my control in the sense of possession as a pure concept of the understanding. I can pick up and take an apple into my empirical physical possession or I can take a parcel of land under my physical control by putting a fence around it and locking the gate to the fence.

Kant does not define "possession" itself, but does seem to rely on Achenwall's concept of possession: "Possession is a durational status, during which someone has a thing under his control to the exclusion of others, or during which someone has the physical capacity to use the thing to the exclusion of others."<sup>23</sup> Furthermore, Kant is aware of Achenwall's ideas on possessing something "juridically" (*iuridice*). According to Achenwall,

<sup>22</sup> Original acquisition is acquisition that is not derived from someone else's own (RL 6:258:9–11).

<sup>23</sup> Gottfried Achenwall, *Ius Naturae*, 5th edn. (Göttingen: Bossiegel, 1763), §120, 104. Kant does speak of "possession of a thing to the exclusion of others," VRL 23:296:5–6.

one possesses something juridically when one has control over use of the thing to the exclusion of others and also has the intent to possess the thing as one's own.<sup>24</sup> If one combines Kant's ideas on holding or having something as one's own and Achenwall's element of intent, one can conclude that a person can take an external thing and have it under her control for a period of time with the intent that the thing be her own.

Kant then distinguishes the nominal from the real definition of the external mine and thine. The nominal definition states that the external mine is some external object for which hindering me in using it in any way I please would be an injury, namely a hindrance of my freedom which is compatible with everyone's freedom under a universal law (RL 6:248:32–249:3). The real definition states that the external mine is some external object for which hindering me in using it would be an injury even though I am not in possession of it in the sense of physical possession (RL 6:249:5–7). Even though I am not in physical possession of some external object, Kant notes that I must be in some sort of possession of it or someone else's interference would not be an injury to me. This possession is intelligible possession (*possessio noumenon*).

Even without the concept of intelligible possession, Kant has covered a lot of ground. I am physically capable of using external objects of my choice and thus must also be legally capable of using them. To use them I must possess them. I can possess an external object of choice if I have it in my power to take the object into my possession without violating anyone else's freedom of choice. If I take an external object into my possession rightfully, I can possess it over the period of time I keep it under my physical control and I can use it in any way I choose (*beliebiger Gebrauch*, RL 6:246:26). I can use a physical thing by keeping it as an object of art, using it for the purpose for which it was designed, using it in any other imaginative way, giving it away, lending it to someone, selling it, transferring it to my heirs on my deathbed. In short, I have all of the rights associated with an ownership right as long as I keep the thing under my physical control. Furthermore, I can possess the thing with the intent that it be mine.

The capacity I have to possess external objects of choice and treat them as mine under the postulate is supported by a duty the others have not to hinder me in so doing, namely the duty: "to act toward others so that external (usable) objects can become someone's own" (RL 6:252:13–15). Moreover,

<sup>24</sup> Achenwall, *Ius Naturae*, §120, 104. Possession with the intent to have something as one's own does not amount to ownership, because a thief can possess booty with the intent to have it as his own. Nonetheless, the thief does not own the property.



“one who proceeds following a maxim through which it would be impossible to have an object of my choice as mine injures me” (RL 6:256:25–27).

The final step in Kant’s argumentation is showing that I have all of these use rights associated with my possession of the object even though I do not possess it physically. In addition, the concept of intelligible possession of an external object also means that I have a right to demand return of the object if it is under someone else’s physical control. Intelligible possession is a right one has against everyone that they not interfere with what one intelligibly possesses. This right results from my faculty to impose unilaterally an obligation on all others they otherwise would not have not to use an object of my choice because I was the first to take that object into my possession (RL 6:247:4–6).

Originally I can acquire only external things (RL 6:259:12–14). I acquire something originally according to Kant’s “principle of external acquisition”:<sup>25</sup>

What I take under my *control* (according to the laws of external *freedom*) and which I have the capacity to use as an object of my choice (according to the postulate of practical reason), and finally what I *will* (according to the idea of a possible united *will*) it should be mine: that is mine. (RL 6:258:22–27)

I have already discussed the first two “moments” of external acquisition and will therefore concentrate here on the third. It is the third that turns physical possession into intelligible possession by permitting the acquirer to impose the obligation on all others not to interfere with the external thing I have acquired originally. In particular, I argue that the “possible united will” of which Kant speaks is not the will of society or the state and thus that acquisition of external things does not depend on social or state approval.

The possible united will of which Kant speaks is the will of the original community of the earth (*communio fundi originaria*). This community is formed by virtue of the spherical shape of the earth’s surface. Since the earth is not an unlimited plain, we cannot scatter indefinitely on it but instead form a community of earth dwellers. This community is the idea of a community and not any actual community of individuals.<sup>26</sup> Kant

<sup>25</sup> The principle of external acquisition also applies to acquiring other objects of choice, such as someone’s choice to perform an act or someone’s status in relation to me, but here I am using it only to explain original as opposed to derived acquisition.

<sup>26</sup> The Grotius–Pufendorf tradition assumed that a primæval community actually existed and that the very first acquisition enjoyed the actual consensus of humankind, see Achenwall, *Ius Naturae*, §116, 97–100. For a comparison of Grotius and Kant, see Katrin Flikschuh, *Kant and Modern Political Philosophy* (Cambridge University Press, 2000), 152–58.

calls the notion of a primæval community (*communio primæva*), meaning a community of individuals that actually existed at one time on the earth, a “fabrication” (RL 6:251:4–10). Still, Kant has the problem of explaining how original acquisition could take place without violating the freedom of all others in light of the obligation imposed on them not to interfere with what was acquired (RL 6:261:8–10). Kant bases his theory on the idea of the original community of the earth and the things upon it (RL 6:251:1–2), which wills that individual acquisition impose this obligation.

That the individual’s will need be contained in the united will of the original community of the earth for the individual will to impose an obligation on all others to refrain from interfering with what that person acquires as her own might seem to suggest we need social approval of individual acquisition. This idea, however, is incorrect. Kant provides an argument to show that the will of the original community is to divide the land and it is this will to divide that makes the individual’s will to acquire a particular piece of land legislating for all others. The individual will to acquire a piece of land is contained in the original community’s united will to divide the land.

Unsurprisingly, Kant begins his argumentation with the original<sup>27</sup> right to freedom (RL 6:237:29–32). He combines this right with the fact that the land on the face of the earth is scarce because the earth is spherical (RL 6:262:23; 6:311:23–24; 6:352:9–11; 6:352:9–11; VRL 23:322:4) and not an infinite plain. The right to freedom guarantees me a right to be somewhere on the earth’s surface. No one may push me off the earth into the ocean or propel me into the universe without violating my original right to freedom. Kant states: “All human beings are originally (i.e. before any act of choice with legal effect) in rightful possession of the earth, i.e. they have a right to be where nature or fate (without their will) has placed them” (RL 6:262:17–20). Kant calls this right to a place on the earth’s surface a “disjunctively universal right,” meaning a right to be here or there (disjunctive: VRL 23:323:26–30; 23:320:20–23) on an unparticularized or still unspecified piece of land (universal: VRL 23:321:14–16).

The holders of the original right to be somewhere on the earth’s surface form an original community of individuals in their relation to each other. They each have a right against everyone else in this community that they

<sup>27</sup> The term “original” means prior to any act with legal relevance. “Original” is what we must assume for our acts to have legal relevance. Kant contrasts “original” to “adventitious,” meaning the contingent that has in fact happened in reality and has legal relevance. The import of these terms cannot be discussed within the confines of this chapter, but they are fully explicated in Byrd and Hruschka, *Kant’s Doctrine of Right*, ch. 2.

have a place to be on the earth. The original community of the earth is not common ownership of the earth's surface by all human beings. "The state of the community of the mine and thine (*communio*) can never be conceived as original" (RL 6:258:11–13).<sup>28</sup> Furthermore, peoples do not originally exist in "a legal community of possession (*communio*) and thus of use or ownership" of the earth's land (RL 6:352:14–17). Instead, this community is simply the idea of all people on the earth in their potential contact with each other, each with an original right to freedom and thus an original right to be somewhere on the face of the earth.

The idea of an original community calls forth the idea of a universal and united will. Kant speaks of "the innate common possession of the earth's surface and its a priori *corresponding* universal will of permitted private possession of it" (RL 6:250:20–21). Similarly, Kant speaks of the "united choice of everyone in common possession" (RL 6:261:10–11), of "common possession by the human race," "to which an objectively united or to be united will corresponds" (VRL 23:323:31–32).<sup>29</sup> The common will of the original community of the earth is to pursue the community's common goal, which is the good of the community. Without a common goal, one could not conceive of a common will of the community. Since the community is a community of persons who have claims against each other within the community to places on the earth, the goal of this community must initially be division of the earth's surface. I have not only a disjunctively universal right to a piece of the earth, but this right is connected to the idea of *particularization*. Kant expresses this idea as follows:

All human beings are originally in *common possession* of the land on the whole earth (*communio fundi originaria*) with the will (of each) given to them by nature to use the land (*lex iusti*), which, because of the naturally unavoidable confrontation of the choice of each against the choice of the other, would extinguish all use of the land if this will did not simultaneously contain the law for choice, according to which a *particular* possession for each can be determined on the common land (*lex iuridica*). (RL 6:267:4–11)

Returning to the idea expressed in Kant's "principle of external acquisition," when I will that a piece of land be mine, my will is binding for all others, because my will, which gives me the law for my occupation, is "contained in an a priori united ... absolutely commanding will" (RL 6:263:19–23). The originally united will wills division of the available land

<sup>28</sup> Kant is not saying here that there is no original community of the earth, but instead using the word "community" to mean a group of joint land owners.

<sup>29</sup> See too VRL 23:289:13n; 23:289:16–18; 23:290:17n.

to avoid constant confrontation, and with my occupation of a piece of land I execute this will. "What I will to be mine, that is mine," because this act is "in accordance with the idea of a possible united will" (RL 6:258:24-27).

Far from requiring social or state approval of property rights, Kant develops an argument for why such approval is unnecessary. It is unnecessary because my will as an original acquirer of an external thing is contained in the universal a priori united will of all, the will of the original community of the earth with its goal to divide the land and the things upon it to avoid constant conflict. Kant's entire line of argumentation on original acquisition of external things can be applied by analogy to derived acquisition under contract or by virtue of a family relationship. None of the acquirable rights to external objects of choice requires social approval. The role of the state is simply to secure these rights, rights we have in the state of nature.

## CHAPTER 6

# *Punishment, retribution, and the coercive enforcement of right*

*Allen W. Wood*

### I INTRODUCTION

Most philosophers who have considered legal punishment suppose that penal institutions, either as they exist or at least as they would exist if they lived up to some normative concept of them, can be morally justified. This supposition, when we consider it in relation to existing criminal justice systems, has long seemed to me dubious in the extreme. But the moral justifications that are commonly offered fall into two general types:

- (1) Retributivist: criminal acts inherently deserve to be punished simply on account of their wrongness or moral wickedness and in some proportion to their injustice or moral badness.
- (2) Consequentialist: the punishment of criminal acts is morally justifiable as a way of achieving one or more goals that are highly desirable or even socially compelling and indispensable. Chief among these is the prevention of crime through deterrence. Other ends that have been suggested are the moral improvement of the criminal, expression of the public's disapproval of the crime, and the satisfaction of a desire to see offenders harmed on the part of the public generally or else the victims of crimes or their relations. (It is noteworthy that these last two justifications, whatever their merits, are consequentialist, not retributivist justifications.)

### 2 KANT'S RETRIBUTIVISM

Kant is best known as a subscriber to the first of these justifications of punishment. It seems to me there can be no doubt that this common view about Kant is correct. "Punishment by a court (*poena forensis*)," he says, "can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society." Rather, punishment "must

always be inflicted on [the criminal] only because he has committed a crime.” “The law of punishment” (by which Kant appears to mean the principle that crime, moral evil, or wrongdoing must be punished) “is a categorical imperative, and woe to him who crawls through the windings of eudaimonism in order to discover something that releases the criminal from punishment or even reduces the amount by the advantage it promises” (MS 6:331).

Further, Kant holds that the proper measure of punishment – what punishment constitutes justice, and is neither too severe nor too lenient – is also to be determined entirely by the gravity of the crime, and not at all by any other aims that judges or legislators may have in instituting penal statutes or enforcing them. “But what kind and what amount of punishment is it that public justice makes its principle and measure? None other than the principle of equality (in the position of the needle on the scale of justice), to incline no more to one side than the other.” This Kant calls the “law of retribution” (*lex talionis*) (MS 6:332).

Perhaps Kant’s best-known (or even, one might say, his most infamous) expression of these ideas is found in a passage where he conjoins the indispensably retributivist justification of punishment, and the *lex talionis*, with his conviction that it is a consequence of the *lex talionis* that willful murder must be punished with death:

Even if a civil society were to be dissolved by the consent of all its members (e.g. if a people inhabiting an island decided to separate and disperse throughout the world), the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve and blood guilt does not cling to the people for having not insisted on this punishment; for otherwise the people can be regarded as collaborators in this public violation of justice. (MS 6:333)

It is important here to distinguish at least two distinct claims that seem to Kant conjointly to require this conclusion. What is most likely to shock people (or at least what most shocks me) about this passage is Kant’s extreme insistence on the death penalty – which he regards as required by justice in cases of deliberate murder, but which civilized people today regard with horror and revulsion. Another claim, however, is more basic to the point he is making and is surely quite reasonable. This is that the established rules of justice must be administered consistently, or as it is sometimes said, that “like cases be treated alike.” For the public not to treat them alike, to give some criminals the benefit of circumstances entirely irrelevant to their guilt, can be regarded as itself an act of injustice in which the public, as Kant says, would be complicit. There should be

nothing to shock us in that thesis, and a sensible consequentialist ought equally to insist on it – even in the fantastic case that the society is about to dissolve. To reject this would imply not so much that one is a consequentialist as that one has a very shortsighted view about the consequences of social policies, as well as no sense of what people will regard as required by basic justice.

Several other misunderstandings of Kant's retributivist position are also possible that might make it seem either not retributivist at all or else even more unreasonable than it really is.

### 3 A NON-RETRIBUTIVIST KANT?

First, when Kant says that the criminal must be punished only because he has committed a crime, we might take this not as providing a general philosophical justification for punishment, but only making a conceptual point about the nature of the institution of punishment itself, which has consequences for the way it is to be administered by judges and penal officials. That is, we might take him to be making only the point that it belongs to the concept of punishment that for a judge administering punishment, the fact that someone has committed a crime should count as the sole and sufficient reason for punishing him, and that the measure of punishment should be neither more nor less than what the laws provide for as the just punishment of that crime. If this is the point Kant is making, then he is not necessarily offering a retributivist justification of the institution of punishment, but only a point about how the concept of punishment constrains those who administer the institution.<sup>1</sup>

I think there is no doubt that Kant accepts the concept of punishment presupposed by this interpretation. In his lectures, he declares that "punishment in general is the physical evil visited upon a person for moral evil," and then he goes on to distinguish "retributive punishments" – those pronounced "because the evil has occurred" – from "deterrent punishments," or "pragmatic punishments," punishments regarded "as a way of preventing crime" (C 27:286). He undoubtedly agrees with the conceptual point that even pragmatic or deterrent punishments, simply as punishments, must be administered in such a way that they are inflicted

<sup>1</sup> This position has been argued by B. Sharon Byrd, "Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution," *Law and Philosophy* 8 (1989), 151–200; others who have defended a similar idea are Don E. Scheid, "Kant's Retributivism," *Ethics* 93 (1983), 262–82 and Thomas E. Hill, Jr., "Wrongdoing, Desert, and Punishment," in *Human Welfare and Moral Worth: Kantian Perspectives* (Oxford University Press, 2002), 310–39.

because a crime has been committed. A judge must not punish someone who has not been found guilty of a crime, solely as a way of preventing future crimes or achieving some other desirable end; and the lenity or severity of punishment a judge meets out must not deviate from rules of justice for the sake of some other end. These points belong to the very concept of punishment, whatever philosophical justification may be given for having an institution of punishment.

Nevertheless, the claim that this is all there is to Kant's retributivist-sounding assertions is pretty clearly contradicted by these assertions themselves, considered in the context where Kant makes them. For Kant also insists that the link between physical evil and moral badness in punishment is "a direct and necessary one, and physical evil is a direct consequence of moral badness" (V 27:552) and that "all means of punishment, therefore, that aim at protecting the person and property of human beings are but means and signs of punishment itself" (V 27:556). "Punishment itself" here is a moral relation between the criminal act and the physical evil it deserves or "brings upon itself"; Kant holds that legal institutions are justified only because this direct and necessary moral relation obtains, and the punishments they inflict are only "means and signs" of it. Clearly the admonition that we must not "crawl through the windings of eudaimonism to discover something that releases the criminal from punishment" is not directed only to judges in their application of existing laws and institutions but also to legislators, who might enact penal laws that punish crimes more leniently than they inherently deserve in order to achieve some aim extraneous to the doing of justice. Kant would doubtless agree with those who think that punishment as an institution involves in its concept that its administration must be retributive in spirit and practice, but he agrees with this chiefly because he thinks that the institution of punishment itself, as a fundamental matter of morality, is and must be justified chiefly, or even solely, by its retributive intent.

#### 4 KANT'S ACCEPTANCE OF NON-RETRIBUTIVIST REASONS FOR PUNISHING

A second possible misunderstanding, however, might be one that exaggerates what Kant is saying when he declares that punishment must be inflicted on the criminal "only because he has committed a crime." Kant is not opposed to legislators or judges also making use of the institution of punishment to achieve the ends of deterring crime, morally improving the offender, and so forth. He accepts Baumgarten's idea that there can be



pragmatic punishments which are *poenae medicinales*, and distinguishes under this heading *poenae correctivae*, which aim at the improvement of the criminal, from *poenae exemplares*, which aim at warning others against committing similar acts (C 27:286; V 27:551). Kant even says that “All punishments by authority are deterrent, either to deter the transgressor himself or to warn others by his example” (C 27:286).

Thus it appears that Kant thinks the justification for punishment as a legal institution must not only tolerate the consequentialist rationale that the institution of punishment prevents future crimes, but even insist on it as a necessary part of the justification for legal punishment. The force of Kant’s retributivism, therefore, lies in his thesis that the right of civil authorities to use punishment for this necessary end, as well as the general outlines of the institution (including the *lex talionis* requirement that all crimes must be punished by a physical evil equal to their gravity as crimes) is subject to and constrained by the “law of punishment” as a categorical imperative that crimes must be punished solely because the criminal has committed a crime, and because this crime, simply as a moral evil, is inherently deserving of punishment.

This combination of views, and the order of priority among them, is quite explicit in the following remark: “Authority punishes not because a crime has been committed, but so that it shall not be committed. But every crime, in addition to this punishment, has a property of deserving to be punished, because it has taken place. Such punishments, which must necessarily follow upon the actions, are moral in character, and are *poenae vindicativae*” (C 27:286). Thus corrective punishments, which improve the criminal, and exemplary punishments, which serve to deter others from committing similar crimes, are what the institutions of criminal justice should be about. Further, Kant allows that once the claims of justice are served, “the degree and nature of [punishment] ... are decided as prudence and mercy may dictate” (V 27:553). Kant even approves of the sovereign right of clemency to remit or lessen punishments, as long as it involves declining to exercise the rights of the state and does not involve a failure to enforce the rights of individuals against one another (MS 6:337).

Kant, therefore, despite some of his apparently extreme pronouncements, does not reject consequentialist reasoning when it comes to the devising of criminal law and its application in particular cases. He even holds that one kind of consequentialist consideration – the coercive enforcement of right – is essential and indispensable to the imposition of punishment by civil authorities. His basic retributivist claim is only that

no penal institutions with such consequentialist intents could be morally justified if there were not a more basic justification of punishment – one that is purely retributive.

In the idea of our practical reason there is something further that accompanies the transgression of a moral law, namely, its *deserving punishment*. Now becoming a partaker in happiness cannot be combined with the concept of punishment as such. For although he who punishes can at the same time have the kindly intention of directing punishment to this end as well, yet it must first be justified in itself as punishment, that is, as mere harm [*Übel*], so that he who is punished, if it stopped there and he could see no kindness hidden behind this harshness, must himself admit that justice was done to him and that what was allotted to him was perfectly suited to his conduct. In every punishment there must first be justice, and this constitutes what is essential to its concept ... Thus punishment is a physical harm that, even if it is not connected with moral wickedness as a natural consequence, would still have to be connected with it as a consequence in accordance with the principles of moral legislation. (KpV 5:37)

It is no doubt for Kant a conceptual truth about legal punishment that it is first of all a harm or physical evil and that it must be administered only to those who have committed crimes – and that its just measure is specified by the *lex talionis*. But it is apparently also part of the very concept of punishment for Kant not only in its particular administration, but in its very nature as a civil institution, that punishment must have a fundamental moral justification of a certain kind – namely, a retributivist rather than a consequentialist justification. It is then also permissible, even desirable, in Kant's view – and perhaps also necessary from the standpoint of civil authority – that punishment be used to prevent future crimes and, if possible, even to benefit the criminals themselves. But *first* punishment must be justified (not only within the context of its civil administration, but even more fundamentally, from a moral point of view underlying the justification of its use by civil authority) simply as (retributive) justice, before it can be directed even to its vital civil end of preventing crimes and protecting the rights of persons and property.

##### 5 WHAT HUMAN BEINGS MAY PUNISH

One final point deserves mention in clarifying Kant's commitment to retributivism. Although he holds that moral evil of any kind deserves punishment, he does not hold that any human being (and in particular civil authorities) is justified in punishing moral faults or moral wrongs that do not involve a transgressions of *right* – that is, of the rights of individuals

to all external freedom that is consistent with the like freedom of others in accordance with universal laws, and the rules of civil society required coercively to enforce these rights. Civil punishment, according to Kant, is always coercion (V 27:554–55). But the performance of ethical duties, as distinct from duties of right, may not be externally coerced – to do so is itself an act of wrong or injustice (MS 6:220, 239). Merely moral transgressions are therefore not punishable (V 27:548) – except perhaps by nature, or its Author; but such “natural punishments” are necessarily distinct from those that may be administered by a human court (MS 6:331, 439; RGV 6:140–41; VpR 28:1084–88; V 27:552–53, 556). And the *lex talionis* may be justly applied only by a court exercising civil authority, not by any human being’s private judgment (MS 6:332).

## 6 RIGHT AND COERCION

Kant’s insistence on retributivism – that it is a fundamental moral principle or categorical imperative that moral evil deserves punishment – is clear enough. We look in vain, however, for any argument from Kant in favor of retributivism, or any attempt to ground it on more basic principles of his philosophy of right or morality. Kant even seems to confess, or rather to proclaim, that no such argument is capable of being given: the direct and necessary link between moral badness and physical evil, that it “consists in a *malum physicum, quod moraliter necessarium est*, cannot be discerned through reason, nor proved either, and yet it is contained in the concept of punishment that it is an immediately necessary consequence of breaking the law” (V 27:552).

The best way to understand this claim, I think, is to consider it in relation to Kant’s general theory of public right, or the right of a state. “A *state* is a union of a multitude of human beings under laws of right, insofar as these are a priori necessary as laws, that is, insofar as they follow of themselves from concepts of external right as such (are not statutory), its form is the form of a state as such, that is, of *the state in idea*, as it ought to be in accordance with pure principles of right” (MS 6:313). The right of a ruler (*Befehlshaber*) in a state to punish crimes is listed by Kant among the powers belonging, jointly or severally, to the three basic authorities (*Gewalten*) that constitute the “idea of the state” or the “state in idea” that is to serve “as a norm (*norma*) for [a state’s] internal constitution” (MS 6:313). (These authorities are the sovereign, or legislative authority, the ruler, or executive authority, and the judicial authority.) The powers pertaining to the idea of the state include

the sovereign's supreme proprietorship of the land (MS 6:323–24), the right of the ruler to impose taxes (MS 6:325–28) and distribute offices (MS 6:328–30), as well as the right to punish and exercise clemency (MS 6:331–37). I suggest that the “law of punishment” – the categorical imperative that crimes must be punished according to the measure of the *lex talionis* – is to be considered as one of the “laws of right” or “a priori laws” that are supposed to “follow of themselves from concepts of external right as such” that “constitute the idea of the state.” Thus Kant claims that “every murderer must suffer death” because “this is what justice, as the idea of judicial authority, wills in accordance with universal laws that are grounded a priori” (MS 6:334). Yet someone might be forgiven for not finding it self-evident that moral badness must be coupled with harm or physical evil, or especially that the harm must be “equal” to the offense, as provided by the *lex talionis*. Alternative consequentialist theories of why crimes should be punished at all, and what the proper or just punishment for a given crime might consist in, seem in principle entirely defensible. Kant is in a particularly weak position in claiming to rule out such theories, because he himself holds, as we have seen, that the punishments meted out by civil authority must always be deterrent in their aim and effect – they must serve to prevent wrong or crime. Kant himself holds that it lies outside the authority of the state – or indeed any human authority – to attempt to couple moral badness with harm or physical ill except in cases where this serves the purpose of protecting rightful freedom and coercively preventing the violation of right. So he himself insists on this deterrentist (consequentialist) justification for legal punishment as a necessary condition of its rightfulness. It remains unclear – apart from Kant's bare assertions of “universal laws that are grounded a priori” – why a retributivist justification of punishment would be needed in addition to this one.

The deterrentist (consequentialist) justification is, however, connected closely to Kant's general theory of right in quite compelling ways. External freedom, independence from constraint by another's will (as long as your will leaves them externally free in accordance with universal law), is the sole innate right possessed by human beings, and is grounded on their dignity as ends in themselves and self-governing agents (MS 6:237). For Kant, the function of the state is to establish a general condition of right (*Rechtzustand*) – in contrast to a “state of nature” (*Naturzustand*) (MS 6:305–06). A rightful condition is one in which everyone's rightful freedom of action is protected by a coercive authority that limits everyone's external actions to those that are right and prevents people from doing

“wrong” (or injustice) to others – that is, doing acts that infringe on their freedom to do acts that are right.

A rightful condition is imperative because even if all happened voluntarily to limit their actions to what is right, no one would truly enjoy a condition of rightful freedom. This is because rightful freedom would still have no determinate boundaries (there would be no way to settle disputes about it, such as disputes over property) and also no coercive protection or guarantee for it. Even if people’s rights weren’t in fact violated, they would not be protected, as justice requires. Thus a state of nature is already a state of injustice, no matter what people in it might choose to do with their lawless freedom (MS 6:312).

7 IT IS ANALYTIC THAT RIGHT INCLUDES  
THE AUTHORIZATION TO COERCE

It is fundamental to the powers of the state, therefore, that it be able rightfully to coerce its subjects to respect the right of others and limit themselves to actions that are right as determined both by pure principles of right and by the state’s own laws and statutes. Kant derives this right to coerce from something fundamental to externally free (or right) action in general. He takes it to be an analytic proposition that any act of coercion that prevents wrong (*Unrecht*) (i.e. prevents the interference with an action that is right) is consistent with everyone’s freedom according to universal law. Such coercion is therefore right.

I think we can see why Kant considered this proposition to be analytic if we realize that external freedom requires more than merely that others not in fact interfere with your freedom of action. I am not truly free to do something as long as someone else is in a position to interfere with my doing it, even if this person might choose not to exercise his power to interfere. If a mugger is holding a gun on me, I am not free to retain my wallet as long as he is in a position to shoot me, even if he might choose not to do so. Journalists, for instance, are not free to publish their reports of events as long as the state claims the power to suppress publications and imprison journalists, even if in fact the state (perhaps for its own reasons) decides to allow a certain report to be published with impunity. To be free to do something, in relation to the possible interference by other people, you must not only not in fact be interfered with, but you must also be immune from such interference, as by being protected by coercive laws requiring them not to interfere, and backed up by sanctions that can be expected to prevent them from interfering. If this is what it

means to be free to do something, then an action which coercively prevents another from interfering with your action is necessarily consistent with your freedom to do it. More generally, once we have determined which actions are right – which actions would be consistent with the like freedom of others according to universal law – the freedom to do actions that are right cannot be violated by forms of coercion that prevent only actions that are wrong – that is, actions that coercively interfere with right actions. Although all coercion deprives someone of freedom, not all coercion deprives anyone of rightful freedom, and some acts of coercion are necessary to protect rightful freedom. It follows that the concept of right itself involves, analytically, the authorization (*Befugnis*) to use coercion against wrong (MS 6:231). In a rightful condition of civil society, however, the authorization to exercise rightful coercion resides not in individuals but exclusively in the ruler (the state's executive power) acting under the laws made by the sovereign (the united general will of the people).

#### 8 KANT'S BEST JUSTIFICATION OF PUNISHMENT IS NOT RETRIBUTIVIST

This line of reasoning can be used to provide a clear and cogent account of the state's right to punish. Punishment is justified as a form of coercion used to protect right.<sup>2</sup> But this is not in the least a retributivist justification. It justifies certain forms of coercion by arguing that they systematically have certain consequences – the coercive protection of right actions – and

<sup>2</sup> By this I do not mean that this justification of punishment is free from problems. There are, I think, some fundamental difficulties with any attempt to justify threats of harm as a way of coercing people to refrain from wrongful behavior. Such threats can be justified, for example, only insofar as they *successfully* prevent wrong. Where they do not do so, the presumed justification vanishes. Kant is aware of, and quite emphatic about, this point in his discussion of so-called "equivocal rights," those that cannot be coercively enforced because there is no coercive threat that can suffice to prevent the wrong – as in the case of the "right of necessity," where one man in a shipwreck kills another by pushing him off the plank; here even the threat of the death penalty would not prevent him from committing the homicide, since he would die immediately if he did not commit it (MS 6:235–36). The point also plays a role in Kant's discussion of infanticide by an unwed mother and one military officer's killing another in a duel (MS 6:335–37). These, he says, are crimes deserving the death penalty but the act of killing is so bound up with the killer's need to preserve her or his honor that the threat of death cannot prevent the killing. But this line of argument, it seems to me, may be used to pose a quite general challenge to the justifiability of coercion by means of threats. What is to prevent us arguing that no punishment is justified in the case of a terrorist fanatic who will not be deterred by any penalty whatever? Or even more fundamentally, consider any crime that is actually committed despite the law's threat to punish it: if the threats had been sufficient coercively to prevent the crime, it would not have occurred. May we not conclude, by reasoning analogous to that used in the case of the right of necessity, that it is always and necessarily wrong to punish any actual crime whatever? Perhaps these objections will be deemed sophistical, but I think they point to a serious problem involved in morally justifying

do not have others – the coercive interference with right actions. In the context of Kant’s practical philosophy, this seems to be a much better grounded justification of punishment than Kant’s retributivism. The insight into this fact is what I think motivates those, such as Sharon Byrd and Thomas Hill,<sup>3</sup> who want to treat the deterrence of wrong as Kant’s real justification of punishment, and interpret away his retributivism as merely a conceptual analysis of punishment from within the practice.

Punishment as the coercive deterrence of wrong is, of course, clearly not by itself *inconsistent* with retributivism. The problem is that Kant’s theory of right would seem to make the coercive deterrence justification of punishment fundamental, and threatens to make Kant’s retributivism superfluous, or else to turn retributive punishment into merely a way of administering it (as Byrd and Hill want to do). Kant, however, is emphatic, as we have seen, that the relation between the two is just the reverse: “All means of punishment, therefore, which merely aim at protecting the person and property of men are but means and signs of the punishment itself” (V 27:556). Kant clearly intends that punishment as the coercive protection of right would not be justified if the retributivist justification were removed. Pragmatic justifications of punishment – or punishments as *poenae correctivae* and *poenae exemplares* – are, he says, founded on *poenae vindicativae sive morales in sensu stricto*. Only the latter “are truly *poenae justitiae*, because they are immediately necessary according to the principles of justice” (V 27:551).

But it is hard to see how Kant could be right about this. If it is analytic that the coercive prevention of wrong is right, then the state necessarily acts rightly and justly, never wrongly, whenever it coercively interferes with wrong – as it does in every case where it justly threatens the punishment of wrong acts and justly carries out the threat. No retributive justification of punishment is required. Retributivism seems to be an extra wheel, unconnected to the mechanism of Kant’s theory of right and morality. Despite Kant’s claims that the law of retribution is a “universal law of right grounded a priori,” it has no discernible grounding in anything else in Kantian ethics. The correct reaction to Kant’s retributivism would then seem to be the reaction of Kant’s greatest (and most consistent) follower, Johann Gottlieb Fichte, when he concluded that Kant’s

any form of coercive prevention of wrong that operates by means of threats. And that means it identifies a serious problem with Kant’s justification of punishment based on the thesis that right analytically includes the authorization to coerce.

<sup>3</sup> See again Byrd, “Kant’s Theory of Punishment” and Hill, “Wrongdoing.”

retributivism is a moral doctrine without any foundation, “grounded on a categorical imperative which is inscrutable (*unerforschlich*).”<sup>4</sup>

9 IS KANT’S RETRIBUTIVISM EVEN CONSISTENT  
WITH KANTIAN ETHICS?

But that’s not all. It gets even worse. When we look more closely at some of the basic claims of Kantian ethics, we see that it is difficult (or, I think, downright impossible) even to reconcile retributivism with some of them.

In the *Metaphysics of Morals*, Kant argues that our ethical duties are all founded on certain “duties of virtue,” or ends it is our duty to have. The two fundamental such ends are our own perfection and the happiness of others (MS 6:385–88). Kant might also be seen as deriving these ends from the formulas of universal law and humanity, when he argues (in the third and fourth examples) that the maxim of refusing to set them is impermissible under these formulas (G 4:422–23, 430). The fact that the happiness of others is a duty of virtue makes it equally impermissible to adopt any maxim setting a contrary end – namely, one that makes anyone’s unhappiness directly my end – as occurs, for instance, in the maxims characterizing the vices of envy, ingratitude, and malice (or hatred) (MS 6:458–60).

There is in Kant’s texts never any suggestion, much less any argument, that there ought to be an exception to this duty of virtue in the case of wicked people or criminals. Yet Kant’s “law of punishment” tells us that if someone has committed a crime, then the judicial authority is both permitted and required to make its end the infliction on the criminal of a pain or harm that is equal to the crime. Kant asserts the “law of punishment” to be a categorical imperative and a law of reason, but he never derives it in any way from his practical philosophy, and even suggests that it does not admit of any derivation. However, the law of punishment seems to be in direct conflict with the basic ethical principle that the happiness of others is a duty of virtue. For it makes the infliction of harm or ill (unhappiness) a direct end of action, required by justice.

<sup>4</sup> J.G. Fichte, *Fichtes sämtliche Werke*, I.H. Fichte (ed.) (Berlin: Walter de Gruyter, 1970), 3:282–83; *Foundations of Natural Right*, Michael Baur (trans.), F. Neuhouser (ed.) (Cambridge University Press, 2000), 245. Fichte is no retributivist at all. In the theory of punishment, as in many other areas of philosophy, he proves himself to be a better friend to the critical philosophy than Kant ever realized, by drawing conclusions from the Kantian philosophy more consequentially than Kant does.



There is no such problem about the justification of punishment as coercive protection of right. It does not involve making the unhappiness of others an *end*, but merely permits the ruler to inflict unhappiness – in ways that accord with right – as a *means* of coercively protecting the rightful freedom of all. So not only is the justification of punishment as coercive enforcement of right validly derived from Kant's theory of right (as the retributive justification is not) but it is also consistent with his moral theory (while the retributive justification is inconsistent with it).

I suppose one might try to reconcile retributivism with the principle that one must always make the happiness of others – never their unhappiness – one's end, by claiming that when we punish the criminal we do not make his unhappiness directly our end, but only inflict unhappiness on him as a means to something else – namely, doing justice. But the distinction between inflicting unhappiness and doing injustice seems to be a distinction without a difference if we hold that the moral link between crime and punishment – between moral evil and physical evil – is direct and necessary, hence that doing justice to a criminal simply consists in imposing harm or physical evil on him. For if that is accepted, then the end of doing justice for its own sake seems here to be identical with the end of inflicting harm or unhappiness for its own sake.

It is true, of course, that this is a conflict not directly between two duties of right but between an alleged duty of right (the duty to punish retributively, to make someone's unhappiness your end) and an ethical duty (the duty always to include the happiness of others, never their unhappiness, among your ends). It is certainly consistent with Kantian ethics to hold that we are permitted by right to do something that violates an ethical duty – for instance, we have a *right* to act on the two immoral maxims of neglecting our talents and refusing to contribute to the happiness of others that Kant famously discusses in the *Groundwork* (G 4:422–23, 430). No one may rightfully compel us to fulfill these duties, or for that matter, the ethical duties of love that forbid envy, ingratitude and hatred, and vengefulness (MS 6:458–61), or the duties of respect for the human dignity of others that forbid arrogance, defamation, ridicule, or giving scandal (MS 6:465–68) – as long as no violation of their rights is involved in these violations of ethical duty. No one may coerce the fulfillment of these duties, and anyone who tried to do so should be coercively prevented from it. But Kant's retributivism threatens not a conflict between what *we may permissibly do* by right and what we are required to do by ethics, but rather presents us with a kind of action – making the unhappiness of another our end, which we seek for its own sake – that we

are *required to do* (by a juridical duty in the sphere of right) but also *forbidden to do* (by a duty of virtue in the sphere of ethics). Every duty, however, according to Kant, is a categorical imperative (MS 6:219, 222–23). Kant's retributivism therefore threatens us with a direct conflict between two categorical imperatives.

#### IO RETRIBUTIVISM AND DIVINE JUSTICE

Kant famously claims that goodness of will constitutes the indispensable condition of the worthiness to be happy (G 4:393). We clearly need not interpret this as authorizing just anyone to punish any person at all simply for having a bad will. Indeed, we have seen that Kant insists that humanly administered punishment may be carried out only by civil authorities, and only when it serves the purpose of coercively preventing wrong, either by correctively deterring the wrongdoer or deterring others from similar crimes by making him an example for them. And basic principles of Kantian ethics forbid us to read it as instructing us to remove happiness from anyone we think of as unworthy of the happiness they enjoy. For such a reading would again directly contradict the fundamental principle of Kant's theory of duties that the happiness of others must always be an end for us, hence that their unhappiness may never be among our ends. So it is directly contrary to Kantian ethics to suppose that you or I might be entitled to go around removing happiness from wicked people because we judge their conduct has not made them worthy of it. Indeed, Kantian principles regard all such attempts as contrary to fundamental ethical duty, and in case they involve the infringement of others' rights, they ought to be coercively prevented or punished. In Kant's view, it is only for God to apportion happiness according to worthiness, in his providential concern to achieve the highest good (*summum bonum*) or what is best for the world (KrV A804–19/B832–47; KpV 5:110–19; KU 5:442–53).

Kant occasionally tries to present God's providential apportionment of happiness in accordance with worthiness as a case of doing retributive justice (RGV 6:140–41; ED 8:327–29; VpR 28:1084–88; V 27:553). But if we accept what I have just been arguing, that retributivist way of viewing God's providence also contradicts some of the most fundamental tenets of Kantian ethics. It also contradicts, I believe, some of the most basic things Kant says about God's will and the nature of justice as one of God's moral attributes.

Kant sees the moral attributes of God as constituting a triad – like the triads that constitute the four quadrants of his table of categories (KrV A80/

B106, 111), or the three main formulas of the moral law (G 4:436). In each of these triads, namely, the first two members present us with an opposition or contrast, which is then resolved in the third member, in which the first two are combined, reconciled, or synthesized. The category of totality is the synthesis of the categories of unity and plurality; the category of limitation synthesizes reality and negation; the category of community synthesizes substance and causality; the category of necessity synthesizes possibility and actuality; and the formula of autonomy or the realm of ends synthesizes the form of the moral law (as universality) with its matter (humanity as an end in itself). God's moral attributes are holiness, benevolence, and justice (RGV 6:140–41; VpR 28:1073–76). Kant tries to see this triadic conception of God's moral nature confirmed as an ideal of the highest being by the way he finds it exemplified in various conceptions of divinity in religions found throughout history and in all parts of the world:

The religion of Zoroaster had these three divine persons, Ormuzd, Mithra and Ahriman, the Hindu religion had Brahma, Vishnu and Shiva ... The Goths revered their Odin (father of all), their Freya (also Freyer, goodness) and Thor, the judging (punishing) god ... The religion of Egypt had its Ptha, Kneph and Neith, of whom, so far as the obscurity of the reports from those ancient times permit us to surmise, the first was to represent spirit ... as the world *creator*, the second, a generosity which sustains and *rules*, the third a wisdom which limits this generosity, i.e. *justice* ... It is hard to give a reason why so many ancient peoples hit upon this idea, unless it is that the idea lies in human reason universally whenever we want to think of the governance of a people and (on the analogy with this) of world governance. (RGV 6:140–41)

Especially revealing is Kant's account of the way that God's justice combines or synthesizes his other two moral attributes: "God's benevolence is something positive, but *justice* is fundamentally only a negative perfection, because it limits his benevolence in the measure that we have not made ourselves worthy of it. God's justice therefore consists in *the combination of benevolence with holiness*" (VpR 28:1076).

Limited benevolence, however, is still benevolence. It does not involve making anyone's unhappiness an end for its own sake, but rather only making their happiness an end to a limited degree. Divine justice, as the apportioning of happiness to worthiness, cannot on this account be treated as a case of retributive justice. This is further confirmed by a criticism Kant levels at Baumgarten for presenting God's justice in too retributivist a light:

But the expression *poenae vindicativae*, like the expression *iustitia ultrix* [both expressions used by Baumgarten] is really too hard. For vengeance cannot be

thought in God ... So it is better to regard the punishments inflicted by divine justice on sins in general as an *actus of iustitiae distributivae*, that is, a justice limiting the apportionment of benevolence *by the laws of holiness* ... God's justice must limit benevolence so that it distributes good only *according to the subject's worthiness*. (VpR 28:1086–87)

“*Poenae vindicativae*,” however, is precisely Kant's own expression for retributive punishments, in those very passages where he claims that corrective and exemplary punishments depend for their justice on retribution (C 27:286–87; V 27:551–53). In the above passage, therefore, he is directly denying that divine justice takes the form of retributive justice, and replacing it with a conception of divine justice as distributive justice which limits benevolence by holiness, but never ceases to be benevolent. God, like the virtuous moral agent in Kantian ethics, always makes the happiness of human beings, never their unhappiness, an end. God's justice is therefore never retributive in character.

Neither in this world nor beyond it, therefore, can the basic principles of Kantian ethics be made consistent with Kant's retributivism. Kant is famous (or infamous) for advocating an extreme form of retributivism concerning the justification of punishment. He ought, however, to be equally famous for having advocated a conception of moral value and constructing on its basis an ethical theory which is fundamentally incompatible with retributivism about punishment.

## II A FINAL TROUBLING THOUGHT

Someone might think that in arguing that Kant's retributivism is without foundation in his theory of right and even openly contradicts some fundamental tenets of his ethics, I am merely showing myself to be an anti-retributivist malcontent who wants to save Kant from himself and that I hope to do so by attributing to Kant's basic moral theory certain views that I myself hold but Kant obviously rejects. But this would not only be to dismiss, on merely *ad hominem* grounds, all the *arguments* I have provided for my claims that Kant's retributivism is groundless relative to his moral theory and also in direct contradiction to its principles, but it would also badly mistake my own attitude toward retributivism. For I am in certain respects very much drawn to retributivism as a theory of punitive justice, and I regard it as something highly unsatisfactory about Kantian ethics that it cannot justify or even consistently accommodate the retributivist doctrines Kant obviously held. I will conclude this discussion by explaining this last remark.

I think, namely, that retributivism gets right at least one important thing that all alternative theories probably get wrong. By focusing attention almost exclusively on what the criminal deserves, retributivism enables us to set strict rightful limits on what it is permissible to do to criminals in the name of justice – something that not only nonretributivist theories, but also popular moral opinion and actual penal institutions, seem deplorably unable to do.

In nineteenth-century England, pickpockets were publicly hanged, even though the jubilant crowds witnessing these gruesome demonstrations of public bloodthirstiness were at great risk of having their pockets picked while reveling in the grisly festivities. In many places in the USA today, monstrously draconian prison sentences are mandatory for helpless drug addicts. Children are often prosecuted and punished as if they were adults (so that a child, who clearly has not attained to those faculties required for responsible agency, or a mentally incompetent adult, who also obviously possesses those faculties to only a diminished degree, is sometimes punished for murder with life imprisonment, and even in a few cases with execution). “Three-strikes” laws, such as those in California (passed, and then later confirmed by popular referendum) sometimes result in someone’s receiving a life sentence even for a minor theft. Such hateful laws represent only a few of the unspeakably barbarous practices that prevail in the USA and throughout the world that convince me (as I indicated right at the beginning of this paper) that it is implausible in the extreme to think that the actual institution of legal punishment, as it exists (or has ever existed at any time or place in human history) could ever be morally justified. There is no rational standpoint from which the penal laws just described could ever be defended. They show nothing except how mean-spirited and vindictive the public can be, and how shameless politicians are in appealing opportunistically to everything that is worst in their constituents. They are all damning indictments of what Kant called the radical evil in human nature. When you consider that they are acts of the public, exercising power over individuals, they must be regarded as at least as bad as any of the private crimes of lawless individuals to which they are a flagrantly unjust response.

Such laws can certainly be criticized as ineffective deterrents, counter-productive in their effects and involving social costs far greater than any benefits they might promise. (The only benefit, if we may consider it that, is the way they satisfy the public’s most backward and depraved obsessions, sick fears, and irrational cravings for vengeance.) Even in the eighteenth century enlightened philosophers such as Montesquieu

and Beccaria already presented the empirical evidence to show that it is the certainty of punishment, not its severity, that serves as an effective deterrent to crime. Drug addicts, who have only limited control over their impulses when it comes to their addiction, and children or people mentally retarded or defective, whose impulses similarly fall short of self-control, are not going to be deterred by increasing beyond all rational limits the kind and degree of cruelty practiced on them when they violate society's prohibitions. These consequentialist objections to such practices are obviously well taken as far as they go. But to any person of decent moral sensibilities, they ought to seem also strangely cold-hearted. For the first and most obvious thing to say about these penal practices is rather that they are outrageously *unjust* – that the severity of the punishments is grossly disproportionate to the gravity of the crimes – and that this would be true even if they did serve as effective deterrents, or whatever other welcome consequences they might have, which would pale into insignificance next to the violation of human rights that is in question here. Moreover, as far as I can see, it is only a retributive theory of punishment – and especially a reasonable interpretation of the *Ius Talionis* (in the form of the principle that the severity of punishments must be reasonably proportionate to the gravity of the corresponding crimes) – that could adequately justify this natural and immediate reaction.

Kant says: “In every punishment as such there must first be justice” (KpV 5:37). He means that no matter what end we may have in view – and no matter how noble that end, or how vital it may be to the welfare of society – it is not permissible to inflict harm on another person which, but for his commission of some wrong, it would be unjust to inflict on him, unless this conduct toward him, and the kind and degree of the harm, can first be justified simply from the standpoint of justice, and made consistent with his basic rights as a human being. The more basic thought is accordingly closer to the other thought that no end outside justice, no conception of the social good, and no fear on the part of society in the face of threats to it could ever justify punishing anyone more severely than their conduct inherently deserves.

That thought is an essentially, irretrievably, retributivist thought. For this reason, I regard retributivism as indispensable when it comes to assessing the morality of punishment. And I hold it to be a serious deficiency in Kantian ethics that not only can it provide no good justification for retributivism but apparently cannot even consistently accommodate it. In short, I regard the acceptance of some form of

retributivism as indispensable if society is ever to be (in the words of our fashionable political slogans) as “soft” on crime as we would need to be to if we are deal justly with criminals. Whatever opinions Kant himself may have held or expressed in favor of retributivism, Kantian ethics is seriously lacking as long as it cannot justify them or even consistently include them.

*Moral feelings in the Metaphysics of Morals*

Paul Guyer

## I INTRODUCTION

In his notorious illustrations of morally praiseworthy actions from duty in Section I of the *Groundwork for the Metaphysics of Morals* Kant says that “action first has its genuine moral worth” only when it is done “without any inclination, simply from duty” (G 4:398) and that “an action from duty is to put aside entirely the influence of inclination” (G 4:400). A page later, however, he also says that because “an action from duty is to put aside entirely the influence of inclination and with it every object of the will ... there is left for the will nothing that could determine it except objectively the *law* and subjectively *pure respect* for this practical law” (G 400–01), and in his footnote to this passage he equates such respect with a feeling, although not an “obscure” feeling, but one “*self-wrought* by means of a rational concept and therefore specifically different from all feelings ... which can be reduced to inclination or fear” (G 4:401n.) In the *Critique of Practical Reason*, similarly, although Kant maintains that “What is essential to any moral worth of actions is *that the moral law determine the will immediately*,” which in turn seems to mean *not* “by means of a feeling,” this comes within a chapter in which he argues that the feeling of respect is the “incentive” of pure practical reason (KpV 5:71). And a decade later, in the Introduction to the *Doctrine of Virtue* of the *Metaphysics of Morals*, Kant refers not to respect in general but to no fewer than four “aesthetic preconditions of the mind’s receptivity to concepts of duty as such,”<sup>1</sup> which seem to be causally necessary conditions intervening between the moral

<sup>1</sup> “*Ästhetische Vorbegriffe für Empfänglichkeit des Gemüts für Pflichtbegriffe überhaupt*”; Gregor translates this as “Concepts of what is presupposed on the part of feeling by the mind’s receptivity to concepts of duty as such” (CE *Practical Philosophy*, 628). It is by no means obvious how the term “*ästhetische Vorbegriffe*” should be translated. Gregor’s circumlocution is right to avoid translating it simply as “feelings,” since one of the four items that Kant discusses under this rubric, namely conscience, is not, as we will subsequently see, a feeling at all, although it has an effect on feeling; but her circumlocution is nevertheless cumbersome. My more direct translation is intended to leave the meaning of the term to be garnered from Kant’s subsequent discussion.



law itself and the performance of the particular actions that this law calls for in particular circumstances, one of which is a particular feeling of respect that is also designated as “self-esteem” (*Selbstschätzung*).

These facts raise at least two distinct questions. One question is how a single feeling of respect or any greater number of “aesthetic preconditions” are supposed to function as incentives of morally worthy actions when the choices made by the good will are supposed to be determined by nothing but the recognition of the binding validity of the moral law itself, a form of determination that, at least in light of the examples offered in the first section of the *Groundwork* such as that of the formerly sympathetic philanthropist who has fallen into “deadly insensibility” (G 4:398), is apparently supposed to take place without any feeling at all. The other question is how to understand the relation between the unique feeling of respect singled out in the *Groundwork* and second *Critique* and the four different “aesthetic preconditions” described in the *Doctrine of Virtue*.

I will comment only briefly on the first question before focusing on the second. On the transcendental idealist theory of free will and action that Kant developed and refined from the *Critique of Pure Reason* to *Religion within the Boundaries of Mere Reason*, phenomena such as feelings that occur at particular moments in time obviously cannot be causes of the atemporal determination or choice of the noumenal will to make the moral law its fundamental maxim that is supposed to be the ultimate locus of moral worth. But as Kant is at pains to point out in both the first and second *Critique*, the act of noumenal choice, to which ordinary temporal predicates do not apply, cannot then be regarded as the cause of particular actions in time under our ordinary conception of a cause as one event preceding another in time in accordance with a rule (e.g. KrV A144/B183). Rather, the character of an agent’s free noumenal choice, or his “intelligible character,” can be considered only as the non-temporally specific ground of the agent’s entire “empirical character” (KrV A539/B567); “every determination of his existence changing conformably with inner sense, even the whole sequence of his existence as a sensible being 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*” (KpV 5:97–98). This means that the feeling of respect or any other “aesthetic precondition” that might appear to play a role in the mind’s susceptibility to concepts of duty cannot be regarded as a cause of the noumenal determination of the will, only as part of its effect. And this is consistent with Kant’s claim that what he has to demonstrate in the “Incentives” chapter of the second *Critique* is “not the ground from which

the moral law in itself supplies an incentive but rather what it effects (or, to put it better, must effect) in the mind insofar as it is an incentive" (KpV 5:72): the feeling of respect can be regarded only as a phenomenal effect of the noumenal determination of the will by the moral law, not as the cause or any part of the cause of the latter. More precisely, the feeling of respect or any other feelings that seem to play a role in the phenomenal etiology of morally worthy action can be regarded only as part of the effect of the noumenal determination of the will: everything else in the phenomenal process of moral action, from phenomenal consciousness of the moral law, to empirical recognition that one is in a situation that presents some moral choice, to the actual performance of a physical action in the attempt to fulfill the demand of morality (or to subordinate it to the demand of self-love) must also be considered as part of the phenomenal effect of the noumenal determination of the will by the moral law (or the principle of self-love). And this means that the role and indeed the number and kinds of moral feelings involved in the phenomenal etiology of moral action<sup>2</sup> can be decided on empirical grounds and only such: if experience tells us that one or more feelings play a causally indispensable role in the production of moral actions, that can be accepted as the empirical character of moral action – the empirical character of action must be described empirically.<sup>3</sup> Put more harshly, Kant's claims about the noumenal determination of the will become irrelevant to his empirical story about the etiology of moral action, and can safely be ignored in figuring out the latter even if they are accepted as a metaphysical account of freedom of the will.

So that leaves the question of how Kant's account of the four "aesthetic preconditions of the mind's susceptibility to concepts of duty" in the *Doctrine of Virtue* is supposed to relate to his conception of the feeling of respect in his two main works in moral philosophy of the 1780s. That will be my chief focus in what follows, which can thus be taken as an account of Kant's final empirical theory of the etiology of moral action.

## 2 THE FEELING OF RESPECT BEFORE THE *METAPHYSICS OF MORALS*

In his lectures on ethics prior to the *Groundwork*, Kant insisted upon an indispensable causal role for feeling in the production of moral action,

<sup>2</sup> I will henceforth use this as shorthand for "morally worthy action."

<sup>3</sup> See in this regard R #6626, 19:116–17, from 1764–70, where Kant says that "The doctrine of moral feeling is ... a hypothesis to explain the *phenomenon* of approbation" (CE *Notes and Fragments*, 427).

and did not complicate matters by any reference to his eventual transcendental idealist solution to the problem of free will, which he might in any case not yet have developed.<sup>4</sup> At the start of his introductory section on “The Supreme Principle of Morality,” Kant distinguished between the criterion of moral right and wrong on the one hand and the incentive or “principle of execution” for morally right actions on the other, and held that the latter although not the former is moral feeling.<sup>5</sup> Kant maintained that “(1) The principle of appraisal [*Dijudication*] of obligation, and (2) the principle of its performance or execution,” “Guideline and incentive [*Triebfeder*],”<sup>6</sup> must be distinguished (C 27:274). At this time Kant clearly believed that the intellectual apprehension and even approbation of the moral law is necessary but not sufficient to impel an agent to the action that morality requires; for that an additional element, a “feeling” in the “heart” rather than the understanding, is required.<sup>7</sup> The feeling of respect that Kant introduces in the *Groundwork* and second *Critique* clearly takes over the role of this “moral impulse” (C 27:275) in the heart as the proximate cause of moral action, but with the new claim that this feeling is itself the phenomenal effect “self-wrought” by the noumenal determination of

<sup>4</sup> The Collins lectures on moral philosophy are dated “Winter Semester 1784–85,” in other words, the same time as the composition and publication of the *Groundwork*, but they are virtually identical to the notes transcribed by Johann Friedrich Kaehler in the summer semester of 1777 (Immanuel Kant, *Vorlesung zur Moralphilosophie*, Werner Stark [ed.], [Berlin: Walter de Gruyter, 2004]), and may descend from that manuscript or a common source for both rather than reflecting what Kant actually said in the classroom in 1784–85. There is no evidence of Kant’s transcendental idealist theory of free will in any material up to 1777, even though the transcendental ideality of space and time had been introduced in the *Inaugural Dissertation* of 1770. However, the same passage appears also in C.F. Mrongrovius’ notes (M 1 27:1422–23), apparently from the winter semester 1782–83 (see AA 27:1052), which might be thought to increase the likelihood that Kant was still making this claim in his lectures not long before the publication of the *Groundwork*, and that he did not mean to repudiate it even after he had formulated his transcendental idealist theory of free will.

<sup>5</sup> Thus here Kant already rejected his earlier attraction to the moral sense theory of Shaftesbury and Hutcheson, according to which moral feeling is both criterion and incentive, or in Hutcheson’s terms both “justifying reason” and “exciting reason,” for morally appropriate and praiseworthy actions. For this distinction, see Francis Hutcheson, “Illustrations upon the Moral Sense” (1728), Section 1, in Aaron Garrett (ed.), *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*, (Indianapolis, Ind.: Liberty Fund, 2002), 141–46. For Kant’s early attraction to moral sense theory, see the 1764 prize essay *Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality*, Fourth Reflection, §2, AA 2:299–300, and *M. Immanuel Kant’s Announcement of the Programme of his Lectures for the Winter Semester 1765–1766*, AA 2:311–12.

<sup>6</sup> Heath translates “*Triebfeder*” as “motive,” but I am substituting “incentive” for the sake of consistency with the Gregor translations of the *Groundwork* and *Critique of Practical Reason* to be quoted below; see *CE Lectures on Ethics*, 65–66.

<sup>7</sup> See also R #5448, 18:185, from 1776–78, where Kant says that “Moral motives should not have merely *vim objective necessitantem* for the [crossed out: practical] conviction of the understanding, but *vim subjective necessitantem*, i.e., they should be *elateres*” (*CE Notes and Fragments*, 415).

the will by the moral law, previously considered only as the principle of appraisal and not as a cause of any kind at all.

Kant's main discussion of respect in the *Groundwork* is in the footnote in Section 1 to which I have already referred. Here Kant introduces the characterization of the feeling of respect as a mixture of pain and pleasure that he will retain in the second *Critique*:

Respect is properly the representation of a worth that infringes upon my self-love. Hence there is something that is regarded as an object neither of inclination nor of fear, though it has something analogous to both ... As a law we are subject to it without consulting self-love; as imposed upon us by ourselves it is nevertheless a result of our will; and in the first respect it has an analogy with fear, in the second with inclination. (G 4:401n)

But he also writes as if this feeling were merely epiphenomenal: "What I cognize immediately as a law for me I cognize with respect, which signifies merely consciousness of the *subordination* of my will to a law without the mediation of other influences on my sense" (G 4:401n).<sup>8</sup> In the *Critique of Practical Reason*, however, where respect merits an entire chapter rather than a footnote, Kant suggests that this feeling has an indispensable causal rather than merely epiphenomenal role in the phenomenal etiology of moral action.

This is not clear in the chapter's opening assertion that "What is essential to any moral worth of actions is *that the moral law determine the will immediately*" (KpV 5:71) nor from Kant's inference that we must not look for some "other incentive" in order to give the moral law "influence on the will" (5:72), some "*antecedent* feeling in the subject that would be attuned to morality" (5:75), but must instead "determine carefully in what way the moral law becomes the incentive, and, inasmuch as it is, what happens to the human faculty of desire as an effect of" the moral law as the "determining ground upon it" (5:72). However, Kant subsequently offers an account of how the complex feeling of respect, produced by the *thought* of the moral law, *reweights* our preferences before they lead to action, which suggests that the feeling of respect does play a causal role in the motivation of moral actions. Thus Kant writes:

The representation of the moral law deprives self-love of its influence and self-conceit of its illusion, and thereby the hindrance to pure practical reason

<sup>8</sup> Daniel Guevara argues on the basis of text from the *Critique of Practical Reason* that the feeling of respect is an entirely positive feeling, although it does have a "feeling of pain caused by respect for the law" as a consequence (*Kant's Theory of Moral Motivation* [Boulder, Colo.: Westview Press, 2000], 111), but he does not consider this passage from the *Groundwork*, which clearly assigns the painful and pleasurable analogues of fear and inclination to one and the same feeling of respect.

is lessened and the representation of the superiority of its objective law to the impulses of sensibility is produced and hence, by removal of the counterweight, the relative weightiness of the law (with regard to a will affected by impulses) is produced in the judgment of reason. And so respect for the law is not the incentive to morality; instead it is morality itself subjectively considered as an incentive inasmuch as pure practical reason, by rejecting all the claims of self-love in opposition with its own, supplies authority to the law, which now alone has influence . . . This feeling (under the name of moral feeling) is therefore produced solely by reason. It does not serve for judging actions and certainly not for grounding the objective moral law itself, but only as an incentive to make this law its maxim. (KpV 5:75–6)

This passage seems to suggest that the feeling of respect is produced by the mere *representation* of the moral law, not by the antecedent *determination* of the will to act in accordance with that law, and that it is the *reweighting of empirical feelings* or preferences by the feeling of respect that “supplies authority to the law” and leads the agent to make the moral law instead of self-conceit his fundamental maxim. This seems to make the determination of the will to make the moral law its supreme authority follow the feeling of respect rather than precede it, and thereby to suggest that the determination of the will by the moral law is the effect rather than the cause of the feeling of respect.

If Kant were here describing the relation between the noumenal and phenomenal levels of action, his model of free will would be in trouble. But if we take Kant to be describing only the phenomenal etiology of moral action, there need be no problem here: he could be saying that the noumenal determination of the will to respect the moral law expresses itself phenomenally by a sequence in which the phenomenal representation of the moral law leads to the complex *feeling* of respect which, by adding the expectation of pain to our otherwise pleasurable representation of acting in some way contrary to morality and pleasure to our otherwise painful representation of forgoing that course of action for the sake of morality, reweights our empirical incentives and impels us to perform the action required by morality rather than the one contrary to it. This causal model of the representation of the moral law as modifying our prospects for pleasure and pain and thereby removing the hindrance to making the moral law our maxim and giving it the requisite authority would all be the phenomenal manifestation of the noumenal determination of the will by the moral law.

This proposal might not perfectly fit the last passage, however, which concludes by suggesting that the feeling of respect is not the incentive to perform particular actions but rather the motive to make the moral law

itself one's (fundamental) maxim. A few pages later, however, Kant suggests that the empirical effect of the feeling of respect is not the adoption of the moral law itself as one's fundamental maxim but rather the adoption of *more particular maxims of conduct*, which would place the causal role of the feeling of respect between the general adoption of the moral law and the performance of any particular actions. This passage begins in the same way as the previous one –

Whatever diminishes the hindrances to an activity is a furthering of this activity itself. Recognition of the moral law is, however, consciousness of an activity of practical reason from objective grounds, which fails to express its effect in actions only because subjective (pathological) causes hinder it. Therefore respect for the moral law must be regarded as also a positive though indirect effect of the moral law on feeling insofar as the law weakens the hindering influence of the inclinations by humiliating self-conceit, and must therefore be regarded as a subjective ground of activity – that is, as the incentive to compliance with the law – (KpV 5:79)

and thus far seems to suggest the same picture as before, namely that the mere representation (“recognition”) of the moral law produces a change in feelings, thereby creating an empirical incentive to compliance with the moral law itself. However, Kant continues the sentence by describing the “subjective ground of activity” that is “the incentive to compliance with the law” also “as the ground for *maxims* for a course of life in conformity with” the law (KpV 5:79; emphasis added). That is, he now suggests that the feeling of respect is not (or not only) the incentive for the adoption of the moral law itself as one's *fundamental maxim*, but rather the incentive for the adoption of the more *particular maxims* that govern the course of one's life, such as the maxims to cultivate one's talents, be beneficent to others, be respectful of others, and so on.

Kant refers to the feeling of respect “under the name of moral feeling” (KpV 5:76). Thus, however precisely he understands its causal role, in the second *Critique* he recognizes only one causally efficacious moral feeling, the feeling of respect.<sup>9</sup> In the *Doctrine of Virtue*, however, as already noted, Kant recognizes no fewer than four “aesthetic preconditions of the mind's receptivity to concepts of duty, among which “Moral Feeling”

<sup>9</sup> In his discussion of the highest good in the “Dialectic of Pure Practical Reason,” Kant distinguishes the happiness that is to be part of the highest good from a mere feeling of “contentment” at having mastered one's inclinations (KpV 5:119), arguing that this is too easily connected with virtue. This might be regarded as a distinct moral feeling from respect, but it is clear that it is meant only as an effect of moral conduct, not an incentive to it, and so it is not a “precondition of the mind's susceptibility to concepts of duty.”

(*Das moralische Gefühl*) and “Respect” (*Achtung*) are listed separately and “Conscience” (*Gewissen*) and “Love of Human Beings” (*Menschenliebe*) are added. So we must now ask, how do these relate to the single although internally complex feeling of respect recognized in the previous works, and can we correlate the causal roles of these “aesthetic preconditions” with anything hinted at in the previous discussions?

### 3 AESTHETIC PRECONDITIONS OF RECEPTIVITY TO DUTY IN THE DOCTRINE OF VIRTUE

Kant’s discussion of the “Aesthetic Preconditions of the Mind’s Receptivity to Concepts of Duty in General” in Section XII of the Introduction to the *Doctrine of Virtue* (MS 6:399–403) is brief but raises many questions. Questions begin with the title of the section (as already noted). I am using “aesthetic preconditions” as a translation for Kant’s expression “*ästhetische Vorbegriffe*,” but both words in this phrase are problematic. “*Vorbegriffe*” could be more literally translated as “preconcepts,” but that is not English, or “preconceptions,” but that suggests mere prejudices and would thus be misleading; I have chosen “preconditions” in light of the terms “natural predispositions of the mind” (*natürliche Gemütsanlagen*) and “*praedispositio*” that Kant uses in his own explication of what he has in mind. But this still leaves the question of what he means by “aesthetic” here. Gregor paraphrased this term as “on the part of feeling,”<sup>10</sup> but this seems to me at once both too vague and too specific: too vague, because it is not clear what relation or relations are connoted by “on the part of,” and too specific, because it seems to suggest that each of the four instances of “aesthetic predisposition” that Kant will subsequently discuss is itself a feeling, and that is not quite right: moral feeling, love, and self-esteem, three of the four items that Kant discusses, may be feelings, but the fourth, conscience, does not itself seem to be a feeling, although it may have an effect on feeling. But putting the various clues together, it seems clear that what Kant intends to discuss are phenomenal feelings or effects on feeling – “*ästhetisch*” and both “*natürlich*” and “*Gemüts-*” make that clear, especially since Kant uses the last term only to refer to the phenomenal mind or empirical consciousness – that play a causal role – “*-anlagen*” and “*praedisposito*” suggest that – in the actual transition from mere concepts of duty to the performance of action – “receptivity to concepts

<sup>10</sup> Her translation of the title is “Concepts of What is Presupposed on the Part of Feeling by the Mind’s Receptivity to Concepts of Duty as Such” (CE *Practical Philosophy*, 528).

of duty” suggests that. Kant also stresses that these “aesthetic preconditions of receptivity” to duty are *necessary* conditions of being put under obligation – and that too suggests a causal role for these phenomena.

Our questions are, though, *what* causal roles do these distinct “aesthetic preconditions” play and how do they relate to the causal role or roles of the single feeling of respect Kant previously recognized? What I now want to argue is that in his late work Kant has developed a multistage model of the role of feeling in the performance of morally requisite and worthy action. What he now calls “moral feeling” seems to be the same as what he previously called the feeling of respect, and is the first stage in making the moral law effective in the phenomenal etiology of action. But this moral feeling, which itself needs to be cultivated and strengthened in the normal course of human life, works by in turn motivating the cultivation and strengthening of the further “aesthetic preconditions” that Kant mentions, namely conscience, understood as the empirical disposition to hearken to the voice of the moral law when presented with particular situations, and feelings of love toward others or sympathy and of self-respect or self-esteem as more particular kinds of feeling that function as the most proximate causes of particular other- or self-directed moral actions. The first, general form of moral feeling could also be considered to play a causal role in the transition from the recognition of the fundamental principle of morality to the formulation and commitment to particular moral maxims, something that no doubt happens in actual experience only as particular sorts of situations present themselves to the maturing moral agent; Kant does not explicitly suggest this function for moral feeling, but as we saw it was suggested in the second *Critique* and seems entirely compatible with the multistaged empirical model of action offered in the *Doctrine of Virtue*. Even if the agent’s choice to make the moral law his fundamental maxim is supposed to be noumenal and hence atemporal, the encounter with particular moral situations or dilemmas, the selection of particular maxims, and ultimate decisions to perform particular actions must all surely be thought of as things that happen in real time, and for that reason as well as others it makes sense for Kant to give phenomenal occurrences such as feelings a role in the etiology of all of them.

Let us now turn to Kant’s exposition. It begins with the following introductory paragraph, now quoted in full:

There are certain moral qualities such that if one did not possess them there could also be no duty to acquire them. – They are *moral feeling, conscience, love* of one’s neighbor, and *respect* for oneself (*self-esteem*), to have which there is no obligation



because they lie at the basis of morality as *subjective* conditions of receptivity to the concept of duty, not as its objective conditions. They are all *aesthetic* and antecedent but natural predispositions of the mind [*Gemütsanlagen*] (*praedispositio*) for being affected by concepts of duty; to have such predispositions cannot be seen as duty, but every human being has them and it is by means of them that he can be put under obligation. – The consciousness of them is not of empirical origin; rather it can only follow from that of a moral law, as its effect upon the mind [*Gemüt*]. (TL Introduction XII, 6:399; translation modified)

There are several points to be noted here. First, as already noted, Kant's use of the term *Gemüt* in the body of this paragraph as well as in the title of the section makes clear that Kant is talking throughout about empirical conditions for receptivity to or compliance with concepts of duty. Second, Kant says that these factors are affected by the *consciousness* of the moral law, not, as in the opening sentence of the discussion of respect in the second *Critique*, by the *determination* of the will by the moral law; but since the present discussion is entirely empirical, we can assume that Kant means everything he is describing to be phenomenal effects of the noumenal determination of the will while still seeing these factors as causally indispensable factors in the motion of agents from empirical awareness of the moral law to the performance of particular actions in the phenomenal realm. Third, Kant actually shifts between describing these aesthetic preconditions as conditions of receptivity to the *concept* of duty and to *concepts* of duty; this hints at the two-staged model I have described, for the general moral feeling that Kant mentions first might be crucial to the efficacy of the fundamental principle of morality and in turn lead to the choice of particular maxims and cultivation of particular feelings that are the proximate causes of particular actions.<sup>11</sup> Finally, Kant's insistence that we do not have a duty to *have* these empirical feelings and factors because having them is a precondition of being effectively put under obligation at all is complemented in what follows by a repeated insistence that we do have a duty to *cultivate* and *strengthen* these predispositions. On Kant's ordinary account of duty, that would suggest that the cultivation of these predispositions is a product of free choice for which we may be praised,

<sup>11</sup> My argument in this section can thus be interpreted as a response to the charge by Harald Köhl, whose account of the causal role of moral feeling in Kant I have otherwise found very helpful, that Kant's account makes "das Achtungsgefühl die *einzig* und *alleinige* moralische Motivationsquelle," in contrast to a Strawsonian account of moral sentiments, which recognizes many distinct motivational reactive attitudes; see Köhl, *Kants Gesinnungsethik* (Berlin: Walter de Gruyter, 1990), 154–55. My argument is precisely that Kant's theory of the "aesthetic preconditions" in the *Doctrine of Virtue* makes room for an at least phenomenally motivational role for a variety of specific moral feelings, even if he does not list as many of these as a contemporary writer might.

and the failure to cultivate them an omission for which we may be held responsible and blamed, even if it would make no sense to say that we have a free and therefore imputable choice whether or not to have these preconditions.

We can now turn to Kant's separate discussions of the four aesthetic preconditions themselves.

(A) *Moral feeling.* Kant here defines moral feeling as "the receptivity to pleasure or displeasure merely from the consciousness of the correspondence or conflict of our action with the law of duty," and asserts that "All determination of our faculty of choice [*Willkür*] proceeds *from* the representation of the possible action through the feeling of pleasure or displeasure, for taking an interest in it or its effect, *to* the deed." He then adds that if the feeling of pleasure or displeasure that is the linchpin between possible and actual action precedes the *representation* of the moral law, then that feeling is pathological, but if it succeeds the representation of the moral law, then the feeling is moral (TL 6:399). The second of these statements makes it indisputable that moral feeling plays a causal role in the etiology of particular actions, although it leaves open how direct this role is and thus whether moral feeling leads to actions directly or through *further* "aesthetic preconditions." In the first and third of these statements Kant describes moral feeling as a feeling of pleasure or displeasure that is produced by the representation of the moral law. That Kant describes moral feeling as resulting from *either* the correspondence of an action to the moral law or its contradiction of that law might suggest that there are two types of moral feeling, either a pleasurable kind, in the former case, or a painful kind, in the latter case. Kant's opening definition might also make it seem as if the action that he is referring to is an action *already done*, and thus that moral feeling is a retrospective response to an action already performed, in which case it would be either a pleasurable feeling, if the action had been in accord with the moral law, or an unpleasant feeling, if the action had been a violation of moral law; but this would make moral feeling a product of self-judgment, which Kant discusses only later (TL §14), not of mere consciousness of the moral law, and in any case Kant's second statement makes it clear that he is talking about moral feeling in connection with *possible* action, that is, *prospective* action, not retrospective judgment on action already done. Moral feeling arises from the thought of the correspondence or contrariety of a possible, future action with the moral law, not from judgment of prior action. This could still leave open the possibility that there are two separate kinds of moral feeling, pleasurable moral feeling at the thought of a possible

action corresponding with duty and painful moral feeling at the thought of a possible action contrary to duty. However, if we think about Kant's assumptions about human action, we will realize that such feelings do not occur in isolation from one another: Kant's assumption is that we typically consider whether to perform an action in conformity with duty in the face of the possibility of an alternative action suggested by self-love, and thus, that assuming we are moved by the thought of the moral law, then we typically experience *both* displeasure at the thought of one action contrary to duty that is open to us and pleasure at the thought of the alternative action open to us that would correspond to duty. If we interpret this to mean that we feel displeasure at the thought of forgoing an action contrary to duty but (even greater) pleasure at the thought of performing an action in conformity with duty, then moral feeling as Kant describes it here would have the same twofold structure as the feeling of respect as he described it in earlier works, and which he at least once also called moral feeling (KpV 5:76).

But is what Kant is now calling moral feeling the *same* as what he earlier called the feeling of respect? Kant's second statement, that moral feeling is the link between the representation of a possible action and the performance of the action, might suggest that it is not the same as the feeling of respect: thinking back to the *Critique of Practical Reason*, one might suppose that the feeling of respect has already played its role in giving "authority" to the moral law to which the thought of one's possible action has been compared – or more precisely, as we have just seen, to which the alternative courses of action open to one have been compared – and that what is now being called moral feeling is a further feeling leading from the comparison of one's possible action to the moral law to the performance of the action itself. However, several of Kant's remarks in the two further paragraphs of his present discussion of moral feeling suggest that this is not what he has in mind, that what he means is rather that moral feeling is what makes us susceptible to the general idea of acting in accordance with duty, in other words, that it is what gives authority to the idea of the moral law itself in the phenomenal realm and is thus involved early in the causal chain leading to the performance of particular actions in accord with duty, not further down this chain. We will see how this spatial metaphor can be cashed out in the discussion of the three further aesthetic preconditions.

Kant's next paragraph states that:

There cannot be a duty to have or acquire a moral feeling, for this feeling lies at the basis of all consciousness of obligation, in order to become conscious of the necessitation that lies in the concept of duty; rather every human being (as

a moral being) has it in him originally; the obligation with regard to it can only be to *cultivate* it and even strengthen it through the admiration of its inscrutable origin. (TL 6:399–400, translation modified)

And his final paragraph on moral feeling, after taking one more swipe at moral sense theory by insisting that “It is inappropriate to call this feeling a moral *sense*” for it is not a “theoretical capacity for perception” of the content of the moral law, then adds “We have, rather, a susceptibility on the part of free choice to be moved by pure practical reason (and its law), and this is what we call moral feeling” (TL 6:400). Kant’s statements here that “the concept of duty” and “pure practical reason (and its law)” are what activate our susceptibility to the moral feeling make it clear that this feeling is aroused by the moral law in general, that its role in the etiology of particular moral actions is thus that of making the moral law in general effective in us, and therefore suggest that what Kant is now calling moral feeling is in fact the same as what he previously called the feeling of respect, not something distinct from it. Kant’s theory, I suggest, is then that this feeling, once strengthened through cultivation, leads to further events, including the cultivation of the more particular aesthetic preconditions of love of others and self-respect or self-esteem, that ultimately lead more directly to the performance of particular actions.

That a well-developed moral feeling or feeling of respect could lead us to strengthen dispositions to more particular feelings is certainly coherent. But is there a circularity in Kant’s idea that the general moral feeling must be cultivated in order to make the moral law itself efficacious, namely that we need to have strong moral feeling in order to make our general consciousness of the moral law efficacious but we need to have an already strong commitment to the moral law in order to take the steps necessary to cultivate and strengthen this moral feeling (whatever they might be – Kant does not here spell out his obscure suggestion that this feeling can be strengthened through admiration of its inscrutable source). This circularity need not be vicious if we think empirically, as we should be doing here: we may suppose that we have a positive feeling in behalf of the moral law as soon as we are aware of it, but one that would not be strong enough to be relied upon in any moment of moral crisis yet that is strong enough to get us to undertake the exercises in moments of moral repose (with the encouragement of our parents and teachers, in real life) that would then make this feeling strong enough to play its role when called upon. This would seem to be a plausible model of what it would be to cultivate a natural predisposition.

(B) *Conscience*. Kant states that

Conscience is practical reason holding the human being's duty before him for his acquittal or condemnation in every case that comes under a law. Thus it is not directed to an object but merely to the subject (to affect moral feeling by its act), and so it is not incumbent on one, a duty, but rather an unavoidable fact ... The duty here is only to cultivate one's conscience, to sharpen one's attentiveness to the voice of the inner judge and to use every means to obtain a hearing for it (hence the duty is only indirect). (TL 6:400–01)

There are several issues here. First, although conscience is being included among the aesthetic preconditions of receptivity to duty, it is not itself a feeling. However, it causes or stimulates – “affects” – some moral feeling that is, presumably, a trigger to an action. But insofar as it is itself also an aesthetic precondition of receptiveness to duty, even though not a feeling, conscience must be an empirical phenomenon, an empirical awareness of moral law. Thus conscience is not an element in the noumenal determination of the will, but an event in the actual experience of agents, and the empirical etiology of moral action must be at least two-staged: first there is empirically hearkening to the voice of conscience with regard to a particular choice, and then there is the feeling which that “affects” and which will itself be an empirical impulse to action.

Further, since cultivation is an activity that takes place in time, that we have, not a duty to acquire a conscience – for it “is not something that can be acquired ... rather, every human being, as a moral being, *has* a conscience within him originally” (TL 6:400) – but a duty to cultivate it also implies that conscience is itself an empirical phenomenon. Thus “[t]he duty here is only to cultivate one's conscience, to sharpen one's attentiveness to the voice of the inner judge and to use every means to obtain a hearing for it” (TL 6:401).

So conscience is a cultivable empirical awareness of moral law, but of *what* moral law? If conscience were empirical consciousness of the fundamental principle of morality, then it would seem that it would have to be strengthened before moral feeling could be strengthened – we would have to strengthen our awareness of the general law of morality in order to strengthen conscience. And then it would seem that we would have both expositional and substantive problems: Kant should have discussed conscience before he discussed moral feeling, and then we would have to ask what feeling would dispose us to strengthen conscience? However, Kant's juridical metaphor for conscience strongly suggests that conscience is our tendency, inescapable but in need of strengthening, to listen the moral law *in particular cases*, that is, to seek and listen to the kind of moral law

that applies to particular cases, namely, particular maxims. We can readily suppose that we can be impelled to undertake whatever sorts of exercises it is that cultivates this tendency out of the impetus of the general moral feeling or feeling of respect that has already been strengthened in the name of the fundamental principle of morality. If we think of conscience as what leads us to seek out the appropriate maxims for particular situations, we can think of conscience as at least in part the disposition to search for maxims which, Kant suggested in the *Critique of Practical Reason*, could be strengthened out of the general feeling of respect. The cultivation of conscience could thus be thought of as the first effect of the cultivation of general moral feeling, and the etiology of moral action can now be thought of as comprising at least three stages, three predispositions each of which can be strengthened – the cultivation of moral feeling leading to the cultivation of conscience leading to the cultivation of particular moral feelings as the proximate impulses to particular actions.

The interpretation of conscience as the empirical awareness of particular maxims is confirmed by Kant's further discussion of conscience under the rubric of "The Human Being's Duty to Himself merely as a Moral Being" in the body of the *Doctrine of Virtue* (TL §9, 6:430), where he elaborates his concept with theological as well as juridical imagery. Under the more specific title of "The Human Being's Duty to Himself as His Own Innate Judge," Kant here describes conscience as "Consciousness of an *internal court* in the human being ('before which his thoughts accuse or excuse one another')," in which sits an "internal judge," an "authority watching over the law in him" and keeping him "in awe (respect coupled with fear)," which authority "is not something that he himself (voluntarily) *makes*, but something incorporated in his being." Kant argues that we must think of this judge as like God and think of conscience "as the subjective principle of being accountable to God for all of one's deeds" because it would be "an absurd way of representing a court" to "think of a human being who is *accused* by his conscience as *one and the same person* as the judge,"<sup>12</sup> but then turns around and says that the human being is not entitled, "through the idea to which his conscience unavoidably guides him, to *assume* that such a supreme being *actually exists* outside himself ... For the idea is not given to him *objectively*, by theoretical reason, but only *subjectively*, by practical reason, putting itself under obligation to act

<sup>12</sup> Kant thus glosses over Adam Smith's suggestion that we could think of an internal spectator that is arrived at simply by learning to consider how *other people* would judge our own actions if they knew everything we did about our own motivations and intentions.

in keeping with this idea" (§13, 6:438–39). But what is important in all of this is that we conceive of conscience as an "ideal person" who "impose[s] all obligations" as well as "a scrutinizer of hearts." The point is just that conscience can be thought of as the empirical voice that informs us of our *specific obligations*, and our natural tendency to listen to this voice can itself be cultivated with the impetus provided by general moral feeling, and will in turn provide us with the impetus to strengthen particular moral feelings that can be proximate causes for the actions called for by our particular obligations.

We can now turn to the last two of Kant's "aesthetic preconditions," love of others and (self-)respect. We can think of these as the most immediate impulses to the particular actions called for by the particular maxims entailed by the moral law, impulses that can be strengthened with the impetus given by general moral feeling, perhaps through conscience, although they may also need on occasion to be checked by reflection on our other obligations, which we can certainly think of as being enabled by an adequately cultivated conscience.

(C) *Love of human beings.* Kant discusses the third aesthetic precondition of susceptibility to concepts of duty under the general rubric of "love of human beings," but the ensuing discussion makes it clear that he is referring to love of *others*, in connection with the duty of beneficence. His actual discussion here is limited and disappointing. He starts by saying that "*Love* is a matter of *feeling*, not of willing, and I cannot love because I *will* to, still less because I *ought* to ... so a *duty to love* is an absurdity" (TL Introduction XII, 6:401). But then he adds that if someone practices the duty of beneficence "often and succeeds in realizing his beneficent intention," then "he eventually comes actually to love the person he has helped" (TL 6:402). In other words, Kant actually claims that feelings of love towards (specific) others are a consequence or effect of beneficence, not an aesthetic precondition or cause thereof.

However, Kant returns to the issue of feelings towards others when he discusses the duties of love in the body of the *Doctrine of Virtue*, and there he describes specific feelings towards others that are aesthetic preconditions of susceptibility to duty, thus do play a causal role in the initiation of beneficent actions, and which, as Kant makes explicit, can and must be actively cultivated. He divides the duties of love towards others into beneficence, gratitude, and sympathy (TL §28, 6:452). The first of these is simply the duty to perform actions that directly "promote according to one's means the happiness of others in need, without hoping for something

in return" (§30, 6:453), and is not directly a duty to have any feelings, a fortiori not directly a duty to cultivate any feelings. The second is the duty of "*honoring* a person because of a benefit he has rendered us" (§31, 6:454), which is connected to a specific feeling of respect: "The feeling that is connected with this judging is that of respect toward the ... benefactor" but also requires some act of expression, "*active* as well as merely *affectional* [*affektionellen*] gratitude" (TL 6:455). Kant does not explicitly say that the disposition to have such a feeling or the disposition to perform acts that express them needs to be cultivated, but he certainly could argue that. However, Kant does explicitly assert that the third duty of love, the duty of sympathy, requires the cultivation of feelings, specifically the cultivation of naturally occurring feelings of sympathy towards others – not specific beneficiaries of past acts, but any who might need help – as a means to the performance of beneficent acts in their behalf, as required by the first duty of love, the duty of beneficence. He says that "Nature has ... implanted in human beings receptivity" to the feelings of "*Sympathetic joy* and *sadness*,"<sup>13</sup> that these can be used as "means to promoting active and rational benevolence" (§34, 6:456), and that it is therefore "an indirect duty to cultivate the compassionate natural (aesthetic) feelings in us" in order to be able to "make use of them as so many means to sympathy based on moral principles and the feeling appropriate to them" (§35, 6:457). The cultivation of the feelings of sympathetic joy and sadness to which we are naturally predisposed can be accomplished by undertaking various actions to which we might not otherwise have a direct duty, such as visiting sickrooms or debtors' prisons, which will supposedly strengthen our natural disposition to sympathy. Kant infers that we thus have a duty to perform such actions in order to strengthen "the impulses that nature has implanted in us to do what the representation of duty would not accomplish by itself" (*für sich*

<sup>13</sup> So Gregor translates Kant's pair of terms "*Mitfreude und Mitleid*"; in her sentence, the adjective "sympathetic" translates the "*Mit-*" in each of these terms, so "sympathetic joy and sympathetic sadness" would be closer to Kant's meaning. After this phrase, Kant inserts the parenthetical phrase "*sympathia moralis*," and both "*Mitfreude*" and "*Mitleid*" are to be understood as species of the genus "sympathy," so "sympathy" cannot be used as the translation for "*Mitleid*." By itself, "*Mitleid*" would most naturally be translated as "compassion," and Gregor translates the adjective "*mitleidig*" in the next phrase quoted as "compassionate." But having translated "*Mitleid*" as "compassion" in the first place would have lost the parallel with "*Mitfreude*." This would not have been a terrible loss, however, since it is "*Mitleid*" only and not "*Mitfreude*" which is the proximate cause of beneficent actions: we do not need to help those who are in a state of joy that we can sympathetically share, but only those who are in a state of suffering that we can both sympathetically share and take steps to remedy (in the last two paragraphs of TL §34, Kant endorses the Stoic point that there cannot be a duty to feel the pain of another when there is nothing one can do about it, because that would only add to the amount of evil in the world rather than reduce it; see TL 6:457).



*allein nicht ausrichten würde*) (§35, 6:457). Kant's statement that these feelings are the means that nature has implanted in us to the performance of beneficent actions suggests that these feelings, when sufficiently strong, are the proximate causes of beneficent actions, and thus are the penultimate stage in the phenomenal etiology of (these) particular moral actions. That these feelings need to be strengthened and can be strengthened suggests that there must be something that impels us to perform the actions that will result in strengthening them, for example, visiting sickrooms. It seems plausible to suppose that general moral feeling, directed by a cultivated conscience, can impel us to strengthen our natural disposition to sympathy by performing actions that are not too demanding, such as visiting sickrooms, which will make that disposition strong enough for us to perform beneficent actions that are more demanding than the actions that strengthen the disposition.

Several points should be noted here. First, there is nothing in the claims leading to Kant's last point that suggests that sympathetic feelings should be cultivated only in case the "representation of duty" is too weak by itself to move one to beneficent action; his previous statements that sympathetic feelings are the means that nature has implanted in us to promote benevolence rather suggest that sympathetic feelings are the means *through* which the general representation of duty naturally and ordinarily works to move us to beneficent actions – they are not safeguards in case the representation of duty fails to work, but the means through which the representation of duty normally works. This in turn gives us a way of interpreting Kant's remark that we must cultivate and use the feelings of sympathy as "so many means to sympathy based on moral principles and the feeling appropriate to them": the latter, singular feeling could be general moral feeling or the feeling of respect, and this general moral feeling could be the cause of the cultivation of the more specific feelings, namely the feelings of sympathy, that are in turn the aesthetic preconditions, the proximate causes, of the beneficent actions – particular actions, though within the limits of determinacy for imperfect duties – that the moral law in general and the particular maxims that it entails for such needy rational beings as human beings require of us.<sup>14</sup>

<sup>14</sup> Both Marcia Baron and Nancy Sherman hold that the role of sympathetic feelings is strictly *epistemic*, that is, they are supposed to alert us to occasions when our help is needed; see Marcia Baron, *Kantian Ethics Almost without Apology* (Ithaca, NY: Cornell University Press, 1995), 220 ("they draw our attention to human need and to ways in which we might help"), and Nancy Sherman, *Making a Necessity of Virtue: Aristotle and Kant on Virtue* (Cambridge University Press, 1997), 146 ("Given a practical interest in the moral law ... we still require further information

But second, of course, and as Kant notes, the duty to use sympathetic feelings as a means to action is only a “conditional” duty, because even well-cultivated, strong feelings of sympathy might sometimes prompt one to actions that are morally *impermissible* – as in Barbara Herman’s trenchant example, one has a duty *not* to help someone struggling with a heavy package, which one might otherwise be moved to do by one’s well-cultivated feelings of sympathy, if that person is in the middle of committing a robbery<sup>15</sup> – and in such cases the impulse of sympathy will have to be checked by reflection on all the implications of the moral law, that is, all relevant maxims, and by one’s general commitment to that law and all of its implications. One’s recognition of all of the other requirements of morality and their use to check particular aesthetic preconditions such as feelings of sympathy can both be ascribed to conscience, which has itself, like the more particular moral feelings, been strengthened with the impetus provided by general moral feeling. But none of this means that the duty to *cultivate* sympathetic feelings is a conditional duty; only the duty to *act* upon them is conditional, that is, restricted to appropriate circumstances as determined by the moral law through the application of all relevant more particular maxims (only all of that can determine that “imperfect duties always succumb to perfect ones, just as several imperfect duties outweigh a single one” [V 27:537]). The duty to cultivate these feelings is indirect but unconditional, since it is only by means of them that we are ever capable of performing the beneficent actions that are called for by the maxim of beneficence in appropriate circumstances.<sup>16</sup>

about when and where and how to deploy our practical interest. And such information is often provided through the emotions”). The motivation for such an interpretation is to avoid the implication that emotions such as sympathetic feelings are supposed to be causes of sympathetic actions *instead* of or in *addition* to duty itself; thus Baron says that “One way that cannot be what Kant has in mind is this: the sympathetic impulses join forces with the motive of duty so that their combined strength is more able to combat competing forces than the motive of duty alone” (219). But there is no hint of such a purely epistemic role in Kant’s discussion of the feelings of sympathy in TL §§34 and 35, and Kant’s word “impulses” (*Antriebe*) does suggest that these feelings function as proximate causes rather than as mere information. The way to interpret Kant’s statement that these feelings are impulses that nature has implanted in us to do what the representation of duty alone may not accomplish is, as has been argued here, to interpret these feelings as part of the phenomenal causal process through which the noumenal commitment to duty is expressed; there is no conflict between their causal role in this phenomenal process and the noumenal fact of determination of the will by the moral law alone since the cultivated and causally efficacious state of these feelings is, as has Kant stressed since the *Groundwork*, “self-wrought” by the moral law and its determination of the will at the noumenal level.

<sup>15</sup> See Barbara Herman, *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), 4–5.

<sup>16</sup> This conclusion of course implies that Kant’s example of the philanthropist in whom all feelings of sympathy have been extinguished (G 4:398) is not on his own view a realistic account of moral motivation, but a thought experiment intended only to elucidate the content of the moral law.

The discussion of sympathetic feeling in §§34 and 35 thus provides the positive account of the role of such particular moral feelings in lieu of the merely negative account of love provided in Section XII of the Introduction to the *Doctrine of Virtue*. Of course, to reconcile these two passages, some distinction between the love of others that only follows upon the practice of beneficence and the feelings of sympathy that precede and cause specific beneficent actions would have to be made. Kant does not address this issue, and neither will I. Instead, I will now turn to the last of Kant's aesthetic preconditions, the feeling of (self-)respect.

(D) *Self-respect*. Kant's discussion of the last of the four aesthetic preconditions is very brief and its interpretation necessarily somewhat speculative. The section is entitled with Kant's most general term for moral feeling, "respect" (*Achtung*), and its first three sentences make familiar points about what seems to be that general feeling, which Kant also parenthetically calls *reverentia*: respect is a subjective feeling which it cannot be a duty to have because it is rather the subjective state through which duty is represented. Thus Kant seems to be discussing the same moral feeling as in the first subsection of Section XII. But then he inserts a dash and apparently changes the subject to a more specific feeling of *self-esteem* (*Selbstschätzung*): "it is not correct to say that a human being has a *duty of self-esteem*; it must rather be said that the law within him unavoidably forces from him *respect* for his own being, and this feeling (which is of a special kind) is the basis of certain duties, i.e., of certain actions that are consistent with his duty toward himself." Once again Kant says that we should not simply say that one "*has* a duty of respect toward oneself, for one must have respect for the law within himself in order even to think of any duty whatsoever" (TL Introduction XII, 6:402–03). This is presumably meant to imply what was said explicitly in the previous cases, namely that although one cannot have a duty to *have* a feeling that is the condition of being susceptible to duty, one can have a duty to cultivate and strengthen one's natural disposition to such a feeling in order to increase one's susceptibility to (fulfilling one's) duty. In the final clause of the sentence and section Kant seems to revert to discussing the general feeling of respect rather than the specific feeling of self-respect that he introduced in the previous sentence. Perhaps Kant makes this last leap because he is here assuming, as he elsewhere often argues, that fulfillment of one's duties to oneself is a condition of one's fulfillment of one's duties to others: unless one has perfected one's own moral being, one will neither adequately know one's duties to others nor be disposed to fulfill

them, and unless one has worked toward the (of course only imperfectly attainable) goal of perfecting one's physical and non-moral mental talents one will not be in a position to successfully fulfill one's duties to others even if one does know them and is disposed to fulfill them. In any case, one could ascribe the following model to Kant, analogous to the model suggested in the case of feelings of sympathy towards others: the general feeling of respect prompts one to the cultivation of one's natural disposition to feelings of self-respect which are themselves means afforded by nature to the fulfillment of one's duties to oneself, and the fulfillment of one's duties to oneself, including the perfection of one's moral as well as non-moral abilities, will in turn be the means (or among the means) to the fulfillment of one's duties toward others, which itself involves, at least in part, the cultivation of one's disposition to feelings of sympathy toward others as the natural means to the fulfillment of the duty of benevolence.

Kant does not add any detail here on either the precise effects of feelings of self-esteem or the exercises necessary to cultivate it.<sup>17</sup> But even though detail is lacking, Kant still seems to suggest a model on which both the general moral feeling of respect as well as the specific feeling of self-esteem or self-respect play causal roles in the fulfillment of duties to oneself and to others, and at least sometimes in the fulfillment of duties to others through the intermediate stage of fulfillment of duties to self. Feelings of self-respect thus play a role as proximate causes of particular actions, sometimes actions in direct fulfillment of duties to self and sometimes actions that are themselves means to the further end of fulfilling duties to others.

#### 4 CONCLUSION

As early as 1769–70, Kant had noted that “The *praxis* of moral philosophy ... consists in that formation of the inclinations and of taste which makes us capable of uniting the actions that lead to our gratification with moral principles.”<sup>18</sup> Three decades later he suggested a complex model of the roles of various kinds of moral feeling in the phenomenal etiology of moral

<sup>17</sup> He has a bit more to say about this topic in the *Vigilantius* lectures. Here Kant notes, for example, that by not only inwardly maintaining but even outwardly expressing, for example through proper dress, an appropriate degree of self-respect, one not only avoids making oneself “contemptible in the eyes of others” but even gains influence over others and thereby – again assuming that one's self-respect has not degenerated into self-conceit – “promotes the dissemination of virtue, to which we are obligated in any case; he acquires for himself, and for moral perfection itself, an influence upon others that resides in the feeling of taste” (V 27:635).

<sup>18</sup> R #6619, 19:113; *CE Notes and Fragments*, 425.

action. The empirical consciousness of the moral law produces a general feeling of respect that causes or at least strengthens our commitment to acting in accordance with it and that needs to be and can be cultivated. Strengthening this general feeling can lead us to cultivate conscience, which leads us to discover and hearken to particular moral maxims, and also leads us to strengthen our natural dispositions to particular feelings, such as sympathy and respect for both others (gratitude) and self (self-esteem), which are in turn the final impulses prompting us to fulfill our specific duties towards ourselves and others, although it is proper for us to act on these impulses only where a check of our ensuing dispositions to actions back against the moral law itself authorizes our proposed actions. This is moral *praxis* at the phenomenal level. I have not questioned here Kant's transcendental idealist position that this complex empirical causal process must all be understood only as the phenomenal manifestation of the determination of the noumenal will to abide by the moral law, and is not supposed to be the primary and self-sufficient explanation of how we go from the mere idea of the moral law to the actual determination of the will to make the moral law its fundamental maxim. But if we were willing or even happy to leave Kant's transcendental idealist theory of the freedom of the will in the dustbin of history, then we might be left with an empirical theory of the role of feelings in the general commitment to the moral law, the commitment to particular maxims of duty, and the initiation of particular actions in light of those maxims that seems interesting and promising, although some of its details certainly remain fuzzy.

## CHAPTER 8

# *What is the enemy of virtue?*

Jeanine Grenberg

### I INTRODUCTION

What is the enemy of virtue? What is it that explains the empirically undeniable fact that, for human beings, becoming virtuous is a struggle, something accomplished in the face of some opposing force?

Answering this question is crucial for understanding what virtue itself is: if we do not understand what virtue has overcome, we don't really understand the state that results from the struggle. Yet Immanuel Kant is notoriously contradictory, or at best ambiguous, in his answer to this question. In the *Doctrine of Virtue*, the apparently unambiguous enemy of virtue is the inclinations, understood as a natural force.<sup>1</sup> But in *Religion within the Boundaries of Mere Reason*, Kant suggests that inclinations "bear no direct relation to evil" (RGV 6:35), and thus cannot be opposed to virtue. Instead, we must look to our free capacity to reason and choose as somehow being the internal, conflicted source of opposition. There is, thus, a major conflict in Kant's moral works. Having just (in 1793) identified our freely chosen propensity to evil as the enemy of virtue, he turns around four years later and claims that virtue involves the constraint not of freedom but of our unruly sensible natures. Which account should we prefer, and why?

In this chapter, I investigate this question. First, I compare the two scenarios: inclinations, then a particular state of a free, rational being, as the enemy of virtue. Such comparison will reveal problems for both approaches, but will ultimately show the *Religion* position to be preferable because more genuinely practical. This needn't mean, however, that Kant is left in a state of irresolvable contradiction between his *Religion* and *Doctrine of Virtue* positions, since further investigation of the latter reveals that something approaching the *Religion* position is in fact implicit

<sup>1</sup> See, for example, MS 6:380, 384, and 394.

in it. Most importantly, the central connection Kant makes between virtue as strength and inner freedom can make sense only if we reject natural opponents of virtue and admit that the battle for virtue takes place on the territory of reason and freedom. Through appeal to this internal, rational enemy of virtue, we can, finally, distinguish Kantian and Aristotelian virtue more clearly.

## 2 INCLINATION AS THE ENEMY

Let us consider, then, whether inclination as a sensible, natural force is the enemy of virtue. Most every mention Kant makes of virtue as strength refers to the inclinations as those obstacles against which it fights. Virtue is “the capacity and considered resolve to withstand a strong but unjust opponent ... with respect to what opposes the moral disposition within us” (MS 6:380). And that strong and unjust opponent is most often described as our “inclinations” or, alternatively, our sensible nature.<sup>2</sup>

There are, however, several problems with understanding the moral struggle as a battle between our rational and sensuous natures. First, we have opponents whose weapons cannot touch each other. Sensible things like inclinations operate via natural forces, in quantities that can be measured. But the power of virtue is in reason. Reason is not a physical or mechanical – and thus not a quantifiable or measurable – force. We thus have an enemy external to reason, and, what is worse, the wrong weapons for the battle! To restrain a sensible force, we need another, stronger sensible force. Instead, we have the strength of reasons, a strength to which the forces of nature do not respond.

There is a deeper way of making a similar point: inclinations considered as natural impulses have no standing from the practical point of view, the only point of view from which a discussion of virtue can make sense. Rather, one analyzes natural, mechanical things like inclinations from a theoretical or scientific point of view, as objects in the empirical world.

But reasons operate according to rational laws based in our capacity for free, autonomous choice. This is a non-natural capacity, inaccessible theoretically.<sup>3</sup> To understand free choice, Kant needed, in the *Critique of Practical Reason*, to set aside theoretical reflection and institute a new,

<sup>2</sup> See MS 6:380, 384, and 394 for unambiguous assertions that natural forces are the opponents of virtue.

<sup>3</sup> “[W]e cannot present *theoretically* freedom as a *noumenon*, that is, freedom regarded as the capacity of man merely as an intelligence, and show how it can *exercise constraint* upon his sensible choice” (MS 6:226).

practical way of thinking. This new, practical mode of thinking is re-emphasized as Kant opens the *Doctrine of Virtue*. There, he asserts that were we to “explain the phenomenon” in which one “shows more propensity to listen to his inclinations than to the law,” we would need to explain it as an event in nature, and not as a freely chosen act: “we can explain what happens only by deriving it from a cause in accordance with laws of nature, and in so doing we would not be thinking of choice as free” (MS 6:379n). Our discussion of virtue needs, then, to avoid such natural “explanations” and instead enter that point of view from which we *do* think of choice as free, the point of view of the practical. This involves taking the very same event that a scientist explains as a natural phenomenon – inclination – and considering it instead as an aspect of ourselves as actors and producers instead of as explainers. Or, as Kant later suggests, “a practical philosopher is one who makes *the final end of reason* the principle of his actions and joins with this such knowledge as is necessary for it” (MS 6:375n). In practical philosophy, we are concerned with action and how reason influences action. This new end of philosophizing requires a different mode of philosophizing than what would be appropriate simply for explaining or knowing something as an object in nature. Practically, then, we are interested in “knowledge” only to the extent that it helps us to act and, especially, to move toward actions concerned with “the final end of reason,” itself a practical, not theoretical, concern.

If one refused this new perspective, freedom – a central concern of the *Doctrine of Virtue* – could never make sense. There would be no room for it in a world guided by natural causality. The claim of the Second Analogy, that every event has a cause, would be the rule of the day.<sup>4</sup> Kant thus speaks in a nearly derisive tone of those who refuse this practical point of view, those “who are accustomed merely to explanations by natural sciences,” and who “band together in a general call to arms, as it were, to defend the omnipotence of theoretical reason” (MS 6:378). Such persons simply will never understand morality and virtue: “People who are accustomed merely to explanations by natural sciences will not get into their heads the categorical imperative from which these laws proceed dictatorially, even though they feel themselves compelled irresistibly by it” (MS 6:378). Feeling oneself categorically obligated to laws of reason can hit them over the head like a baseball bat, but they will turn such a practical encounter into a “proud claim” of “speculative” or “theoretical”

<sup>4</sup> Other than to suggest, as Kant does in the mostly theoretically minded argument of the Third Antinomy of the *Critique of Pure Reason*, that we can admit the bare possibility of freedom.



reason, misunderstanding its import entirely. For the seeds of our practical lives to bear fruit, we must therefore think of ourselves not as knowers or explainers, but as rational actors.

It is thus curious, having just insisted that we need to operate practically, that Kant goes on to appeal to inclinations in a way that throws him back into a natural, theoretical mode of operation. Inclinations are “impulses of nature” (MS 6:380), operating according to the laws of the natural world. “Lasting inclinations” are even described as operating according to “a mechanism of sense rather than a principle of thought” (MS 6:480). But analysis of such things would require exactly the natural, theoretical perspective that needs to be avoided. In appealing to inclinations as impulses of nature, Kant thus undermines his own efforts to consider virtue rationally and practically, not naturally and theoretically.

### 3 REASON AS THE ENEMY

To remain in the practical realm, we need an opponent that can be spoken of in rational and practical instead of natural and theoretical terms and, in the *Religion*, Kant identifies just such an opponent. Inclinations considered as sensible impulses are *not* the enemy of virtue. Indeed, *nothing* sensuous could be the true opponent of virtue, since sensuous nature lacks freedom. One cannot therefore be held responsible for the state of one’s sensible nature (RGV 6:35). But whatever it is in us that opposes virtue *is* something for which we can be held responsible. As such, the ground begins to shift: perhaps the struggle is not so much between foreign forces of sensibility and reason. The enemy is instead our propensity to evil: our tendency to freely choose to place inclinations and their satisfaction at the ground of maxims for action in a way that values the self illicitly over other objects and persons of moral value (RGV 6:36). This propensity is a characteristic of our free reasoning capacity itself, not an enemy external to reason.<sup>5</sup>

The obvious reason to prefer this account is that, unlike direct appeal to inclinations as natural forces, appeal to the propensity to evil occurs on the ground of the practical and thus in the realm of freedom. We still

<sup>5</sup> See especially RGV 6:57–59. Kant does sometimes (RGV 6:35) speak of this propensity to evil as a “natural” propensity. But here Kant is attributing this propensity to all persons as part of their nature, not reducing it to a physical state operating according to natural causality. There are, of course, problems here about how to affirm a choice that is also a nature, since what hangs in the balance is the ultimate ground of our free choice. For more on this, see J.M. Grenberg, “Replies,” *Philosophy and Phenomenological Research* 75 (3) (2007), 640–54.

appeal to inclinations, but now not as natural forces. Instead, we consider them as grounds of reasons for action. Furthermore, we can reconcile at least some of Kant's appeals to inclination in the *Doctrine of Virtue* with this more practically minded claim of evil.

First, although Kant regularly appeals to inclinations as natural, sensible forces in the *Doctrine of Virtue*, sometimes his appeals suggest he is implicitly thinking from a more practical point of view. For example, he considers inclinations as integrated into the ends humans set:

[S]ince the sensible inclinations of human beings tempt them to ends ... that can be contrary to duty, lawgiving reason can in turn check their influence only by a moral end set up against the ends of inclination. (MS 6:380–81)

The setting of ends is identified by Kant as the ultimate act of freedom.<sup>6</sup> So, if inclinations are involved in end-setting, they are considered from the practical point of view, that is, from the point of view in which one considers oneself as a free, choosing being. Perhaps inclinations aren't so thoroughly natural and sensible as we first thought. The opposition between these two accounts thus begins to disintegrate. The assertion of a connection between inclinations and ends provides room for us to consider whether, even in the *Metaphysics of Morals*, there is another, deeper, underlying enemy of virtue of which the inclinations are only an indication or hint.

That deeper enemy is what Kant calls "vice." When Kant first introduces the notion, he describes it as the "real opposite" of virtue, even though, since virtue is "strength," its "logical opposite" would be a lack of that strength, or "moral weakness" (MS 6:384). Kant makes a similar point in the *Religion*:

In us ... the law is incentive, = a. Hence the lack of the agreement of the power of choice with it (= o) is possible only as the consequence of a real and opposite determination of the power of choice, i.e., of a resistance on its part, = -a; or again, it is only possible through an evil power of choice. (RGV 6:22n)

The point implied by Kant's *Doctrine of Virtue* discussion is here made explicit: because we possess an incentive toward acting in accordance with the moral law, we need some positive, "real" force to explain why we do not *always* act on it. This "resistance," this "real opposite" of acting morally (and, though he does not make the point here, the reason that

<sup>6</sup> "[S]ince no one can have an end without *himself* making the object of his choice into an end, to have any end of action whatsoever is an act of *freedom* on the part of the acting subject, not an effect of *nature*" (MS 6:385).

we describe any successfully “moral” action not simply as moral but as “virtuous,” that is, as strong actions in the face of an unjust opponent) which was described in the *Doctrine of Virtue* as “vice” is, in the *Religion*, described instead as “an evil power of choice.” A closer look at Kant’s discussion of each will reveal, however, that the two are near equivalents.

In the *Doctrine of Virtue*, Kant draws connections between vice, inclinations, and freedom. Vice is a state in which one’s inclinations against the moral law become established as the ground of maxims guiding one’s actions and disposition: “[i]t is when an intentional transgression has become a principle that it is properly called a *vice*” (MS 6:390). On this model, inclinations are obstacles to virtue only because “it is man himself who *puts* these obstacles in the way of his maxims” (MS 6:394, emphasis added), that is, when one freely *chooses* to make them the ground of maxims, thus establishing inclinations as “obstacles” to proper, free choice. Indeed, Kant argues that every vice “has its distinctive maxim” (MS 6:404). If vice, the “real opposite” of virtue, has its own “maxim,” then the true opponent of virtue is not one’s sensible nature or one’s inclinations as such, but instead a state of one’s capacity for choice in which one makes inclinations opposed to the moral law the basis of one’s actions. That is why Kant is able to say, even as he asserts that “obstacles” to virtue are the “inclinations,” that, in fact, it is more appropriate to say that “[t]he *vices*, the brood of *dispositions* opposing the law, are the monsters he has to fight” (MS 6:405, emphases added).

There is one point at which Kant even associates vice with the *Religion* language of “evil” (*Böse*). In discussing the distinction between affects and passions, he suggests the latter are more problematic than the former:

[A] propensity to an affect (e.g., *anger*) does not enter into kinship with vice so readily as does a passion. A *passion* is a sensible *desire* that has become a lasting inclination (e.g., *hatred*, as opposed to anger). The calm with which one gives oneself up to it permits reflection and allows the mind to form principles upon it and so, if inclination lights upon something contrary to the law, to brood upon it, to get it rooted deeply, and so to take up what is evil [*Böse*] (as something pre-meditated) into its maxim. And the evil is then *properly evil*, that is, a true *vice*. (MS 6:408)

Vice is thus associated with a feeling that becomes “calm[er]” precisely because it has been reflected upon and chosen freely. Once again, Kant succeeds in taking the practical perspective on feelings and inclinations: as passions at least, feelings are things that become incorporated into one’s maxims through free choice. Yet here he takes the further move of describing such rational incorporation of affects contrary to the law as

“evil,” equating such evil with “true vice.” We are not responsible for feelings that we merely “light ... upon” in the course of actions. Anger can come and go; this is just part of what it means to be a sensibly affected being in a natural world. The problem for virtue is when such feelings are “take[n] up” by free choice, and made “calm[er]” and more reflective, thus establishing a “deep” disposition toward immoral actions. *This* freely chosen state – and not feelings or inclinations as such – is the true enemy of virtue; indeed, it is “true vice” and is “properly evil.”

But we are not yet fully at the *Religion* position. For there, in addition to asserting that our propensity to evil is a maxim that incorporates inclinations to form a disposition (RGV 6:36), he asserts further that all such individual incorporations are guided by a meta-maxim that takes one’s inclinations generally and their satisfaction as the guide of *all* action, subordinating the incentive to morality as a secondary concern:

[the evil person] makes the incentives of self-love and their inclinations the condition of compliance with the moral law – whereas it is this latter that, as *the supreme condition* of the satisfaction of the former, should have been incorporated into the universal maxim of the power of choice as the sole incentive. (RGV 6:36)

It is this tendency to subordinate our concern for morality to the concerns of self-love that is the propensity to evil itself. In accepting it as the ultimate enemy of virtue, we move from an image of combat to one of collaboration: one does not so much reject morality as one finds for it an agreeable, lesser place within its own domain of reason-giving. Kant does not take this position in the *Doctrine of Virtue*; yet, his discussion of vice, with its emphasis on principles, not inclinations *simpliciter* as the enemy of virtue, is entirely consistent with this more precise claim. The enemies of virtue as articulated in the *Doctrine of Virtue* and the *Religion* are thus not quite as different as they at first seemed.

#### 4 IS THE ENEMY WITHIN REASON UNCONQUERABLE?

One might worry at this point that we have found an enemy whose strength is more than virtue can vanquish. Kant had considered this worry in the *Religion*: while sensibility provides “too little” to explain our opposition to virtue (since it cannot bear responsibility), making reason the demon would bring “too much,” since this would destroy that within us which is the source of morality and its laws, viz., reason (RGV 6:35). To accept reason as the enemy of virtue would thus not answer at all

the question of what it is within us that fights against virtue. Instead, it would be to give up the fight entirely, since there would not be two worthy combatants. We would not be struggling for virtue, but would instead be merely “diabolical.”

We have not, however, at this point, succumbed completely to this “diabolical” worry: we have not lost the incentive for the moral law entirely, only subordinated it to concerns of self. Yet a concern remains: this new enemy is found within reason itself, and this was just the weapon we meant to use to combat vice. Do we have a mutiny on our hands? Has reason abandoned its role not of legislating the moral law to us but of making it a genuinely irresistible incentive? If freedom is fighting itself in its search for virtue, is it necessarily self-undermining? Can reason really provide the tools to counteract our propensity to evil if this enemy is itself?

Some would say that the only place to turn at this point is to God’s grace.<sup>7</sup> Only by appeal to something apparently external to our corrupted capacity for freedom – viz., God – could we hope to achieve virtue.

Kant himself, though, is not so willing to abandon autonomous action in the face of evil. Even if divine assistance is needed, achievement of a moral cast of mind requires us to make ourselves worthy of such assistance. It is “the lazy and timid cast of mind ... [that] waits for external help” (RGV 6:57). And such lazy refusal of one’s own capacity for autonomous control of his life “renders [the human] ... unworthy” of help from God, or anyone else.

What, though, can one do when the enemy is within? Ignoring the fact that we have an internal enemy of virtue seems unwise. Yet, there is another option: take the proper attitude toward that inextirpable fact of an internal enemy. Through appeal to an already established pattern of moral activity, we can claim a “reason[able] hope” about our moral abilities (RGV 6:68), and a “faith” in our own moral capacities, since “only faith in the practical validity of the idea that lies in *our* reason has moral worth” (RGV 6:63, emphasis added).

So, recognizing rationally chosen vice as the enemy of virtue does make the pursuit of virtue more complicated. After all, hopefulness can become blind optimism if not regularly informed and balanced by the awareness that one can always go wrong. Furthermore, though one can have faith in oneself, one must abandon any certainty about one’s moral status.<sup>8</sup> But

<sup>7</sup> John Hare is the clearest defender of this position. See J.E. Hare, *The Moral Gap* (Oxford: Clarendon Press, 1996).

<sup>8</sup> For further reflections on this point, see J.M. Grenberg, *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue* (Cambridge University Press, 2005), 217–41.

hopefulness and faith in oneself tempered by such vigilance are indeed the proper attitudes with which to respond to this challenge. Vigilance in the use of one's freedom is required to prevent it from undermining itself.

Such a more hopeful, faithful yet vigilant attitude toward one's own capacities is, in fact, one way of describing the attitude that Kant himself makes central to virtue as strength in the *Doctrine of Virtue*. There, he identifies virtue as the most successful realization of our freedom, a claim that could not make sense without first admitting vice or evil as the enemy of our virtue in the guise of a false realization of our freedom. Let us turn to that discussion.

##### 5 VIRTUE AS STRENGTH AND INNER FREEDOM

The relationship of virtue to freedom can be best understood by exploring Kant's idea of "inner freedom." Kant suggests that the constraint accomplished by virtue is not just any constraint, but more precisely, self-constraint of one's end-setting.<sup>9</sup> This capacity for constraining oneself to adopt certain ends and not others is, further, accomplished "not by means of other inclinations but by pure practical reason" (MS 6:396). The struggle to set virtuous ends is not a struggle between inclinations, with the inclination toward morality being the strongest one. Rather, it is an effort of "pure practical reason" to guide the end-setting of free, choosing agents toward those ends that one ought to have instead of those grounded only in concerns of self-love. The end-setting activity of self-constraint is thus best described not just as an act of freedom, but of "*inner* freedom." It is a task that can be accomplished only in oneself and by oneself (MS 6:381).

Stephen Engstrom has thus recently suggested that an act (or, as he describes it, the "strength") of inner freedom is equivalent to virtue as strength.<sup>10</sup> When an agent actually engages in an act of inner freedom to claim the moral law as an end, she uses the "strength" of pure practical reason to constrain her temptation to act in ways that would instead make self-serving inclinations the basis of her ends. As such, the very act or "realization" of one's inner freedom is simultaneously an act of virtue as strength operating to constrain the opponent of virtue within oneself.

<sup>9</sup> Virtue involves "not only the concept of self-constraint but that of an end, not an end that we have but one that we ought to have, one that pure practical reason therefore has within itself" (MS 6:396).

<sup>10</sup> S. Engstrom, "The Inner Freedom of Virtue," in Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 289–315, 305.

But a problem arises at this point: we have appealed to the self-constraint of end-setting to explain the strength of virtue. Yet it is clear that Kant also asserts that *any* setting of an end, not just the setting of a virtuous end, is an act of freedom accomplished within and by oneself: “to have *any* end of action whatsoever is an act of freedom on the part of the acting subject” (MS 6:385, emphasis added). Indeed, in accord with what we have already suggested, freedom of *some* sort – and apparently an “inner,” end-setting form of freedom – is exercised even when one sets immoral, vicious, or evil ends. In the *Doctrine of Virtue*, Kant is of course most concerned with moral inner acts of freedom. But inner freedom could “realize” itself in the opposite direction: one could rely on the strength of one’s rational end-setting to constrain the incentive to morality within oneself and instead set ends grounded in the satisfaction of one’s inclinations. In fact, since we have accepted that the enemy of virtue is not a non-rational, natural force, but rather a state of choice taken on freely, it seems we have to accept also this characterization of an evil choice of ends as an act of inner freedom. Each of these states seems a possible realization of our capacity for freedom.<sup>11</sup>

Admitting this other possible realization of virtue as strength raises further problems, though. Kant says there is nothing “strong” about vice at all. The very thought that the claiming of vice would be an inner act of freedom recognizing its strength is, rather, anathema to Kant (MS 6:384). Furthermore, Kant refuses to accept that freedom itself could be defined as a capacity equally able to realize itself in either the virtuous or vicious direction: “Only freedom in relation to the internal lawgiving of reason is really a capacity; the possibility of deviating from it is an incapacity. How can that capacity be defined by this incapacity?” (MS 6:226–27).

It seems, then, to make no sense to speak of vicious acts as strong realizations of one’s freedom, even less to suggest that such vicious acts could be understood as acts definitional of the capacity of freedom itself. Yet, on the model we have constructed, it seems we need to admit vicious acts as genuine acts of freedom in *some* sense. How to resolve this conundrum?

Kant’s own resolution of the matter doesn’t really help us here, at least not in just the form we find it, since this is a moment in the *Doctrine of Virtue* when he most vociferously claims the “inclinations as natural forces” model of explanation for the enemy of virtue:

<sup>11</sup> Although I’m not sure whether Engstrom would agree with this picture, his discussion of inner freedom reveals his awareness that more needs to be said about the equation of inner freedom and virtue, especially vis-à-vis the question of whether and how this capacity can be lost or misused. See Engstrom, “Inner Freedom,” esp. 304 n. 18.

[S]ince the basis of great crimes is merely the force of inclinations that weaken reason, which proves no strength of soul, the above question [of whether great crimes might require great strength of soul] would be tantamount to whether someone could show more strength during an attack of sickness than when he is healthy. (MS 6:384)

Kant suggests here that vicious acts (or “great crimes”) occur when the strength of one’s natural inclinations overwhelms the strength of one’s inner freedom. So, a free, rational person doing evil things is akin to a healthy body being taken over by a foreign illness, hence Kant’s parallel of the vicious person to someone having “an attack of sickness.” But, as we know, we cannot accept this explanation. Vice is a traitor in the City of Reason; we thus need to make sense of vice as an act of freedom, not as an act or power of a foreign invader.

We need somehow to show that, while a vicious or evil act does emerge from one’s capacity for freedom, such an evil realization of freedom is, nonetheless, not as genuine, definitional, or strong a realization of that freedom as a virtuous one. But how do we accomplish this? We need to look more deeply at what goes on when one utilizes reason to claim a vicious end: the vicious realization of freedom requires a moment of rationalization, or even self-deception, in order to work; and this is a perversion – really, a “weakening,” not a strengthening – of the very capacity for free rational choice upon which the vicious person relies in setting vicious ends.

We can appreciate this point best by considering first the opposing, virtuous, and strong appreciation of the strength of moral reasons as categorical. The Gallows example from the second *Critique* provides a helpful example.<sup>12</sup> There, Kant introduces a man who learns the categorical nature of moral reasons by realizing that, even in the face of death, the moral requirement not to tell a lie still stands. Faced with this difficult choice pitting his commitment to morality against his love of life, this man realizes that the demands of morality present themselves in a categorical fashion, whereas the demands of the satisfaction of inclinations (here, his “love of life”) present themselves with only hypothetical necessity. He *must* do the right thing, even at the cost of his life; but he need not ensure the maintenance of his own existence at all costs. He thus recognizes the irresistible, untrumpable quality of moral demands; and because of this recognition, he also recognizes himself as a free being, capable of fulfilling these demands (whether in fact he does so or not).

<sup>12</sup> See KpV 5:30.



One might, however, react to this presentation of categorical demands differently. In the *Groundwork*, Kant describes a person tempted, through a process of rationalization, perversion, or corruption, to avert his eyes from the categorical demands of morality (G 4:405). This human being encounters the same challenge as the man in the Gallows example, namely, a conflict between happiness and morality. But instead of realizing his capacity for freedom via an honest admission of the categorical nature of his obligations, he instead hijacks his capacity for freedom toward the satisfaction of his inclinations and happiness, that is, toward vicious ends. Freedom has thus been realized in that one successfully sets an end; but this realization is accomplished only at the cost of the perversion of the very reasoning capacity itself, “weakening” it in the sense of lying to oneself about what it genuinely demands. This agent thus *utilizes* the power or capacity of freedom, but only in the sense of *hijacking* it to perverse, not fully rational ends. Reason presents demands as categorical. But one can freely choose to hide that fact from oneself in the name of realizing those ends that satisfy one’s inclinations.

Ironically, this refusal to recognize the quality of one’s experience of obligation was exactly what Kant accused philosophers who insisted on turning the practical task of philosophy into a theoretical one of doing. He is thus accusing them, when they cling always and only to theoretical philosophizing, of a moral, and not just a philosophical, failure. We can thus say that, while acting against the law is not a complete act of freedom the ability for which would be definitional of that very capacity for freedom, it is nonetheless an act that utilizes the power of freedom, albeit in a perverse way.

Virtue, on the other hand, involves an ability to look reason’s categorical demands straight in the eye, unflinchingly. As a result, one also has a clearer sense of one’s capacity for freedom as well as its tendency to warp the demands of reason. The man in the Gallows example, by recognizing a categorical demand as genuinely categorical, has recognized his inner temptation to pervert the demands of reason, and refused it. He has thus constrained (at least for the time being) his propensity to prefer self to moral law. Instead of shirking the categorical nature of some demands, he accepts them as categorical, taking such recognition as a buttressing of his hope that he has sufficient power (that is, freedom) to act as he must. Acts based on such a clear-eyed acceptance of the nature of our moral obligations are thus the truer, more complete realizations of freedom.

Admitting evil as that other possible, but weaker and perverse, realization of freedom thus provides insight into what a true realization of inner

freedom must be: it must be one that avoids this corruption of reason. The import of avoiding the perversion of one's capacity for freedom sheds light, furthermore, on the importance Kant attributes to *contemplation* of the moral law in becoming a virtuous person: one must "enhance the moral *incentive* (the thought of the law) ... by contemplating the dignity of the pure rational law in us" (MS 6:397). Contemplation is similar to what Kant had earlier, and for more moral epistemic than moral developmental purposes, called "attention" or "attentiveness" (KpV 5:30).

To appreciate contemplation or attentiveness, think of this need to dwell upon the dignity of the law from the perspective of the human being from *Groundwork* 1: if we have an internal tendency to pervert, corrupt, or rationalize the powers of the moral law to serve one's own ends, then it makes perfect sense we would be morally required to reflect regularly on the truly categorical nature of moral demands, and to admire the dignity of this fact instead of dwelling on the annoyances one might be tempted to attach to it. We require contemplation precisely because we have a tendency to distract ourselves from the categorical nature of moral reasons.

Furthermore, such attentiveness to or contemplation of the moral law occurs via the moral feeling of respect:

[A]ny consciousness of obligation depends upon moral feeling to make us aware of the constraint present in the thought of duty ... Obligation with regard to moral feeling can be only to cultivate it and to strengthen it through wonder at its inscrutable source. (MS 6:399–400)

Appeal to moral feeling here reveals that, at the heart of this felt attentive state, is an awareness of obligation, that is, an awareness of "the *constraint* present in the thought of duty," expressed affectively in the dual poles of moral feeling.<sup>13</sup> This constrained mode is always the mode under which humans appreciate the dignity of the moral law (that is, the categorical nature of its demands). This fact distinguishes human virtue from that "strength" of "a holy (superhuman) being, in whom no hindering impulses would impede the law of its will" (RGV 6:40).

Contemplation or attentiveness is thus that means for counteracting our tendency to rationalize one's most basic experience of obligation. It allows one to keep an eye on one's own internal conversation about the meaning of one's experience of constraint. Those not willing to engage in such contemplation (like the over-excited theoretical philosophers in the Introduction to the *Doctrine of Virtue*) will lose their understanding of

<sup>13</sup> See, for example, KpV 5:75 for a description of these contrasting sides of the feeling.

themselves as constrained and obligated beings. But the contemplative, attentive person understands her experience of constraint as what it in fact is: an affirmation of the untrumpable strength of moral reasons. Such possession of the moral incentive is, in fact, simply a further articulation of what the constraining strength of virtue is.<sup>14</sup>

We can, then, accept the equation of virtue as strength with the realization of inner freedom, even as we accept the potential for warped, vicious realizations of that freedom. Virtue as strength realizes freedom through an attentiveness that rejects efforts to minimize the categorical force of moral reasons. A vicious realization of freedom requires rationalization and self-deception about those reasons, since only in this way is one able to accept reasons that are not as strong as those presented categorically. A vicious realization of freedom thus realizes freedom only as one undermines its grounding reasons. No strong realization of freedom is accomplished here: strength is realized only in being able to look at the reasons of morality unflinchingly, recognizing them as irresistible. A true realization of freedom is a strong one; a perverse realization of freedom is a mere imposter of this strength. Indeed, vicious realizations of freedom are weak, sneaky, and conniving ones that reveal one's *lack* of strength in claiming the categorical nature of moral reasons. We can thus accept vicious acts as acts of inner freedom, even as we agree with Kant that vice is not a strength and that freedom as such is not properly defined by these vicious acts. Vicious acts are perverse realizations of freedom, a turning of the *strength* of freedom on its head as *weakness*. Vice *is* a state or mode of freedom; it is, however, a perverse and weakened state of it.

## 6 KANT AND ARISTOTLE: CONTINENCE, VIRTUE, HOLINESS

Engstrom has also emphasized the import of contemplation (or, in my language, attentiveness), drawing a picture of the virtuous person who

<sup>14</sup> See, for example, MS 6:391 and 392. It is important to note here that I am attributing to moral feeling a role in attitude, not directly in moral motivation *per se*. I am not suggesting that moral feeling is a natural causal force in producing moral action, only that it is the affective center of the proper attitude or state of mind of the virtuous person. It just so happens, though, that this state of mind, when taken up attentively through one's capacity for free reflection upon one's feelings, functions as an incentive for the virtuous person, since, through it, she sees her status as an obligated being more clearly. Such awareness does not guarantee she will act morally, but it does remove one obstacle to so doing (*viz.*, the obstacle of self-deception). Once moral feeling enters into this broader epistemic state, it has been taken up practically. Its influence is thus not as a natural, causal force, but as an aspect of one's state as a free being.

engages in such contemplation as one very similar to Aristotle's magnanimous and fully virtuous (instead of merely continent or self-controlled) person. Contemplation is neither "theoretical cognition" nor pure "intellectual intuition," but rather "a kind of reflection ... and consideration of ... how attending to the pure rational law in us elevates the moral motive." It is ultimately "the self-consciousness of a pure practical reason"<sup>15</sup> or "inner freedom ... as freedom become self-conscious," a state Engstrom equates with Aristotelian "magnanimity."<sup>16</sup> Further, Engstrom asserts that the Kantian virtuous person, often characterized as nothing more than the Aristotelian continent person, is better characterized as approaching full virtue in the Aristotelian sense (a state of peace, tranquility, and harmony within oneself instead of battle, conflict, and suffering).<sup>17</sup>

There is much in this characterization of virtuous contemplation with which I am in deep sympathy: Engstrom rightly rejects any theoretical or non-practical characterization of this activity; he is right, too, to emphasize the role that contemplation via moral feeling plays in allowing freedom to become conscious of itself in an essentially peaceful, "apathetic," and healthy state of readiness to act morally.

Drawing parallels with Aristotelian magnanimity does, however, cloud the waters a little, for to do so is to fail to remember that, for Kant, even in this calm state of moral apathy, the moral thing to do is always in one meaningful sense done reluctantly: even if my inclinations are in order (that is, have been developed in a way consistent with the end-setting in which I engage through my strong, un-self-deceived realization of my inner freedom), my *propensity* to rationalize the moral law in my favor always exists. I am never at a point at which I could stop contemplating attentively in full confidence that the threat of such rationalization no longer existed.

This does not mean Engstrom is wrong to claim ease or readiness of virtue. Indeed, such ease is realized largely because one *can* train one's inclinations to be in agreement with the ends of virtue. Yet this state cannot be equated with Aristotelian virtue or magnanimity, since Aristotle, repeatedly using the language of "*complete* excellence," insists that the magnanimous and virtuous person *does* get to this exalted state in which *nothing* is opposed to virtue.<sup>18</sup>

<sup>15</sup> Engstrom, "Inner Freedom," 311.   <sup>16</sup> *Ibid.*, 313.   <sup>17</sup> *Ibid.*, 307–08.

<sup>18</sup> For the magnanimous man, "there could be no honour worthy of complete excellence" (NE1124a9). Further, it is impossible to be magnanimous "without complete excellence" (NE1124a29). All quotations from Aristotle are from *Nicomachean Ethics*, S. Broadie (ed.), C. Rowe (trans.) (Oxford University Press, 2002).

Further, Aristotle describes the magnanimous man in a way that makes clear the character of such a person is in serious tension with any would-be Kantian model of virtue. Whereas Kantian virtue demands an abiding awareness of one's tendency toward rationalization, a belief that all humans share this tendency, a resulting avoidance of comparison of oneself with others, and thus a belief that all finite human virtue is to be distinguished from holiness (which encounters no such conflicts), the magnanimous man, in his complete excellence, would never stoop to such vigilance, equality of persons, or distinction between the human and the holy. Quite to the contrary, it is crucial for him to recognize himself as so "superior" (NE1124b11) to other persons that he is, in fact, worthy of honor approaching what "the gods" deserve,<sup>19</sup> not simply that appropriate to other, lesser persons.

The deeper point to appreciate here, though, is that Aristotle and Kant have essentially different concerns in their accounts of virtue. Aristotle is concerned to order that part of the soul meant to be obedient to reason (one's feelings, desires, and appetites).<sup>20</sup> Aristotle thus distinguishes between the virtuous and the continent (or self-controlled) person: the continent recognize right reasons for action, but their "natural" feelings and desires are not in conformity with those reasons. So, the continent person "does nothing contrary to the prescriptions" of reason, yet "does it while having bad appetites." She thus "is such as to feel [a bad appetite] but not be led by it" (NE1152a2–4). The virtuous person, on the other hand, is one who does have her natural state in such conformity: she both does the right thing and has appetites in conformity with that thing.

In contrast, Kant is concerned most centrally to combat vice or evil. This is, as we have emphasized, a corruption of freedom and reason itself. In Aristotle's language, this would be a corruption of that part of the reasoning capacity that orders or gives reasons, not of that part meant merely to obey them. But the corruption of reason is not seriously entertained by Aristotle. The closest he comes is when, as he reflects on this non-rational appetitive part of the soul, he wonders whether this part of

<sup>19</sup> "Worth is stated in terms of external goods, and greatest of these we would suppose to be the one we mete out to the gods ... for this is in fact greatest of the external goods ... [I]t is of honour, especially, that the great think themselves worthy, and worthy they are" (NE1123b16–1123b24). Later, he even speaks with some approbation of the common belief that "men become gods because of an excess of excellence" (NE1145a24). He admits that "it is a rare thing for a man to be godlike" (NE1145a29), but it is apparently not impossible.

<sup>20</sup> "There remains a practical sort of life of what possesses reason; and of this, one element 'possesses reason' in so far as it is obedient to reason, while the other possesses it in so far as it actually has it, and itself thinks" (NE1098a4–6).

the soul opposed to reason “seems to participate in reason,” at least in the sense of being “persuaded” by it (NE1102b14–1103a1). Yet this opponent is never recognized as reason itself in its role of giving “an account of things.” It is, rather, a non-rational opponent that is “rational” only in the weaker sense of being something that can be made tractable to, or be persuaded by, reason.

It is precisely because, for Kant, freedom and reason *do* have this internal opponent that virtue itself needs to be defined differently for him than for Aristotle. Although the virtuous person does train her feelings so as to conform to right reasons, virtue itself is the strength (and, really, the proper orientation or state) of the free capacity of reason itself, not simply the state of one’s feelings.

So, although Engstrom suggests that Kantian virtue is more similar to Aristotle’s virtue than continence, things aren’t so straightforward. There is a parallel to be drawn between the apathy of the Kantian virtuous person and the unconflicted state of the Aristotelian virtuous person. But there is also a similarity to be admitted between Kantian virtue and Aristotelian continence: for both, something in the person is in tension with the “right” way to go, and thus needs to be constrained. For Kant, the thing in need of constraint is not (as it is for Aristotle) natural inclinations themselves, but rather something at the heart of reason itself (not just that part of the rational soul meant to obey reason, but that part of the rational soul that gives reasons). The state of conflict for the Kantian virtuous person is thus somewhat different from that of the Aristotelian continent person. Consider: the person who admits an inextirpable conflict between reason operating properly and reason tempted toward rationalization could, at the very same time, have developed a state of her feelings in accord with reason operating properly. She could, that is, be beyond Aristotelian continence. Her conflict is not one of reason versus inclinations; it is a conflict within reason itself. Often, though not always, one’s feelings will express the unruly side of one’s freedom. But when one has trained one’s feelings toward virtue, then the Kantian virtuous person has gone beyond continence even as she admits something in her that could tempt her astray. Her house is in order, but a potential threat remains.

She is thus not the person of “complete excellence” in the Aristotelian sense, since this would be a state in which one has no fear of internal, mutinous factions. But Kant cannot accept this ideal: humans are not “finite holy beings (who could never be tempted to violate duty)” (MS 6:383). Indeed, were such beings to exist, “there would be [for them]

no doctrine of virtue but only a doctrine of morals.” We humans can only take holiness as a regulative “ideal (to which one must continually approximate),” not as an attainable goal. All of this is, however, just as it should be. This state of strength in the face of an inextirpable enemy is that state of virtue most appropriate to *human* beings, not the Aristotelian dream of a divine (Kant would say “holy”) state.

It should not be supposed, however, that the state of our feelings and inclinations is of no import to Kant; indeed, I have argued elsewhere that the cultivation of one’s feelings is part and parcel of Kantian virtue.<sup>21</sup> But feelings cannot find their proper ordering unless and until freedom itself gets its house in order. All the concerns of Aristotle’s virtue (e.g. acquisition of habitual states of feeling) thus remain for Kant, though now as a result of one’s reorientation and strengthening *of* freedom and reason. But, in all cases, such ordering follows from the strength (attentiveness) of one’s will constraining itself. The “absolute health” of which Kant speaks (MS 6:384) is this fuller integration of a well-oriented and strengthened freedom with one’s person overall.

By identifying vice or evil, instead of natural inclinations, as the true enemy of Kantian virtue and thus admitting a mutinous element within reason itself, we are thus in a better position to understand both virtue as strength as a state in which our inner freedom is authentically realized, and, further, the similarities and differences between Kant’s and Aristotle’s conceptions of virtue.

<sup>21</sup> See Grenberg, *Kant and the Ethics of Humility*.

## CHAPTER 9

# *Freedom, primacy, and perfect duties to oneself*

Lara Denis

### I INTRODUCTION

Kant attributes primacy to duties to oneself in general and perfect duties to oneself (PDS) in particular. He calls PDS the “highest duties of all,” above even our duties to perfect ourselves or to comply with others’ rights (V 27:604). Kant suggests not only that PDS are somehow theoretically or conceptually basic, but also that the implications of their violation are devastating – rendering us incapable of fulfilling other duties and unjustified in demanding respect from others (C 27:343–44; MS 6:437). Why and in what ways do PDS have primacy in Kant’s ethics?

By examining Kant’s accounts of perfect duties to oneself in the *Doctrine of Virtue* and two sets of lecture notes, I show that PDS relate to freedom in uniquely direct, vital, and fundamental ways, in virtue of which they have several sorts of primacy.<sup>1</sup> Section 2 of this chapter (“Freedom and humanity”) highlights freedom’s centrality within Kant’s moral theory, Kant’s identification of humanity as an end in itself with the human being regarded as free, and the grounding of all duties in autonomy and self-constraint. Section 3 (“Perfect duties to oneself”) sketches three accounts of PDS and elucidates their special relations to freedom.

Thanks to Marcia Baron, Anne Margaret Baxley, Katrin Flikschuh, Jeanine Grenberg, Oliver Sensen, Jennifer Uleman, Roger Wertheimer, and Allen Wood for comments on various earlier drafts, and to Manfred Kuehn for correspondence on the right of humanity in our own person.

<sup>1</sup> My analysis has affinities with Mary J. Gregor’s treatment of PDS (to which I owe much of my initial interest in the right of humanity in our own person) in *Laws of Freedom: A Study of Kant’s Method of Applying the Categorical Imperative in the Metaphysik der Sitten* (Oxford: Basil Blackwell, 1963), 46–48, 113–27 and Jens Timmermann’s “autonomy interpretation” of the primacy of duties to oneself in “Kantian Duties to the Self, Explained and Defended,” *Philosophy* 81 (2006), 505–29, esp. 508–15. Distinguishing features of my analysis include its focus on *perfect* duties to oneself (shared with the cited parts of *Laws of Freedom*); its reconstruction of three distinct accounts of PDS; and its explicit delineation of multiple construals of the primacy of PDS in relation to those accounts, and ultimately in light of their special relations to freedom as revealed by those accounts.



Drawing on that section, section 4 (“Primacy”) suggests a variety of ways to construe PDS as primary.

## 2 FREEDOM AND HUMANITY

The main points of this section are that freedom is central to Kant’s conceptions of morality and moral duties, and that much of Kant’s talk about humanity (e.g. as a source of moral constraint) can be interpreted as talk about the human being regarded as free or the freedom embodied by the human being.

Freedom, especially autonomy, lies at the heart of Kant’s ethics. Kant describes morality as “based on the conception of the human being as one who is free but who also, just because of that, binds himself through reason to unconditional laws” (RGV 6:3). The *Critique of Pure Reason* makes room for freedom, and thus for morality, by showing that rational beings are warranted in regarding our wills as undetermined by the laws of nature and capable of complete spontaneity, i.e. as transcendently free (KrV Bxxix–xxx, A533–34/B561–62, A557/B585). In order to show that the moral law is no mere phantom of the brain, the *Groundwork* argues that we may – indeed, must – regard ourselves as autonomous (G 4:445, 452). Only if the moral law is laid upon each agent by her reason can it hold for and motivate all rational beings independently of their empirical natures and desires (G 4:440–41; KpV 5:39–41). Autonomy is not only a feature we attribute to human agents, but also a limiting condition on our choice: we must repudiate “all maxims ... that are inconsistent with the will’s own giving of universal law” (G 4:431; see also 440, 454). Furthermore, the realization of autonomy through morality is the highest vocation of each individual human agent (KpV 5:87–88; RGV 6:183) and of the whole human species (ApH 7:324–25; TP 8:308–09).

Kant’s conception of humanity as an end in itself reflects the centrality of freedom in his ethics. In the contexts in which he sets it forth as an end in itself, Kant identifies *humanity* mainly with the human being as a kind of being who necessarily sees himself, as a result of the freedom of his will, as elevated above the rest of nature (e.g. G 4:428; F 27:1319–22).<sup>2</sup> Freedom

<sup>2</sup> On humanity, freedom, and dignity, see Paul Guyer, *Kant on Freedom, Law, and Happiness* (Cambridge University Press, 2000), esp. 148–71; and Oliver Sensen, “Dignity and the Formula of Humanity,” in Jens Timmermann (ed.), *Kant’s Groundwork of the Metaphysics of Morals: A Critical Guide* (Cambridge University Press, 2009), 102–18.

distinguishes persons from things (G 4:428; MS 6:223; F 27:1335); “freedom and freedom alone warrants that we are ends in ourselves” (F 27:1322).<sup>3</sup> By “freedom,” Kant has in mind independence of causal determination and the self-legislation to which it points (G 4:431–33; MS 6:223; F 27:1322). He identifies the freedom of human choice with its ability to conform to the dictates of law-giving reason (MS 6:226–27). Independence from causal laws or natural instinct is not sufficient to make us ends in ourselves (F 27:1320). Autonomy and freedom of choice are essential to our standing as such (G 4:447; KpV 5:87). While Kant often treats “human being” (*der Mensch*) as interchangeable with “humanity” (*die Menschheit*), referring to each as an end in itself, he sometimes contrasts them (G 4:428–29; MS 6:434–35; V 27:609).<sup>4</sup> In some passages, Kant reserves “humanity” for the human being thought of purely in terms of his supersensible, noumenal freedom, and “human being” for the embodied moral agent or *homo phenomenon* (MS 6:379–80, 462; V 27:579, 593). Sometimes Kant writes as though the humanity “in” a person consists in the rational capacities for setting and systematizing ends that mark him as free (KU 5:431; MS 6:387, 392). Often, however, Kant suggests that the humanity in a person is that person considered as a member of an intelligible world, or *homo noumenon* (MS 6:239, 295; V 27:579, 593).

All duties presuppose the freedom and humanity of the obligated subject. Only autonomous beings are subject to the moral law. For beings like us, whose wills do not thoroughly accord with reason, that law is not merely *necessary* but *necessitating*; it is a categorical imperative and the acts it dictates are duties (G 4:413–14; MS 6:222–23, 379). As beings who are both autonomous and unholy, we are not only self-legislating but also self-constraining. The capacity for rational constraint over our sensible choice, whereby we bring it into conformity with the moral law, is integral to human freedom (MS 6:226–27; V 27:520). Kant identifies ethical obligation, the “internal,” rational self-constraint that binds us to all duties (including juridical ones [V 27:584; cf. MS 6:219–20]), with constraint *by the humanity in our own person* on our sensible agency. In conceiving ourselves as both binding and bound, “we set our intelligible self, i.e., humanity in our own person, over against our sensible self, i.e., the human being in our own person, and thus contrast the human being as the agent with humanity as the law-giving party” (V 27:579). So Kant

<sup>3</sup> Translations from *Kant's Naturrecht Feyerabend* are by Frederick Rauscher and used with his permission.

<sup>4</sup> Throughout, I render “*Mensch*” as “human being” and “*Menschheit*” as “humanity.”

portrays humanity in our person (and its capacity to obligate us) as essential to our standing as subjects of duty.

All duties in Kant's system concern freedom and humanity. Early in the *Metaphysics of Morals*, Kant says: "In the doctrine of duties a human being can and should be represented in terms of his capacity for freedom, which is wholly supersensible, and so too merely in terms of his *humanity*" (MS 6:239). The *Rechtslehre* and *Tugendlehre* lay down principles pertaining (respectively) to external and internal freedom (MS 6:230–31, 396, 406–07). Elsewhere, Kant claims that "all rights are based on the concept of freedom, and are the result of preventing damage to freedom in accordance with law" and "[a]ll obligation is the restriction of freedom to the conditions of its universal agreement with itself" (V 27:587; R #7250, 19:294). The *Groundwork* suggests that all duties can be understood in terms of 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 simply as a means" (G 4:429). According to this principle, one must practically recognize each agent's humanity as a limiting condition on one's choice of ends and means, acting only on maxims that the affected subject could rationally will (G 4:436, 438; KpV 5:87). Kant's illustrations of this principle depict duties as "to" (*gegen*) oneself or another depending on whether their violation fails to treat humanity in oneself or in another person as an end in itself.<sup>5</sup> Thus Kant portrays all duties as requiring that we treat persons in keeping with their elevated status as free beings.

Yet different types of duties relate to freedom differently – expressing, respecting, limiting, protecting, honoring, or fostering it in distinctive ways. Juridical duties, for example, prohibit infringing on the rightful external freedom of others, while imperfect duties of natural and moral perfection require (respectively) cultivating our talents as means for expressing our rational freedom and striving toward its purest, fullest realization. For the relationship between freedom and perfect duties to oneself to explain the primacy of PDS, *especially* tight, deep, or otherwise significant relations must hold between them.

<sup>5</sup> The formula of humanity is the interpretive starting point for my "Kant's Ethics and Duties to Oneself," *Pacific Philosophical Quarterly* 78 (1997), 321–48 and *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (New York and London: Garland, 2001), as well as for Samuel Kerstein's "Treating Oneself Merely as a Means," in Monika Betzler (ed.), *Kant's Ethics of Virtue* (Berlin: Walter de Gruyter, 2008), 201–18.

## 3 PERFECT DUTIES TO ONESELF

Kant generally characterizes perfect duties as negative and owed; imperfect duties as affirmative and meritorious. Perfect duties prohibit performing certain actions, acting on certain maxims, or instantiating certain vices; imperfect duties require promoting certain rationally obligatory ends (our own perfection and others' happiness). Conduct that does not conform to a perfect duty is wrong and wrongs someone; these duties are "unremitting" (G 4:424). Acts that fail to fulfill imperfect duties are not for that reason wrong. Non-fulfillment of an imperfect duty on a particular occasion does not as such wrong anyone; particular acts advancing obligatory ends are "optional" or "meritorious" (G 4:429–30; V 27:578, 600–01; MS 6:390). Kant also describes perfect duties as "strict" and imperfect ones as "wide." This distinction sometimes contrasts duties specified directly in terms of external actions and those specified directly in terms of maxims (MS 6:389–90, 411; V 27:579). Even in the *Tugendlehre*, however, where no duties are specified directly in terms of external actions, perfect duties determine choice more precisely than imperfect ones do. Perfect duties designate vices one's maxims must never instantiate, requiring judgment for determining which maxims do so; imperfect duties set forth general maxims for promoting morally necessary ends, specifying little about degree or manner of promotion (MS 6:390–93, 445–47, 452–58).

Among the acts, maxims, or vices prohibited by perfect duties to oneself are enslaving oneself, servility, suicide, prostitution, begging, lying, and gluttony. These duties are ethical, enforceable only through self-constraint. Even when Kant portrays PDS as directly forbidding certain external actions, he holds that, since violation of these duties does not impinge on the rightful external freedom of others, external compulsion of their fulfillment (*as* PDS) is unjustified (V 27:581, 584; M II 29:619).

In what follows, I explicate three accounts of perfect duties to oneself. Within all three, Kant equates violating PDS with degrading oneself, subordinating one's rational freedom to other objects, impairing one's agency, and treating oneself as a thing or mere means. These accounts reveal uniquely basic, deep, and direct relations between PDS and freedom.

3.1 *Collins*

3.1.1 The Collins lecture notes (reflecting lectures as early as 1774–75) characterize perfect duties to oneself as prohibiting acting against *necessary*

conditions of one's greatest, self-consistent use of freedom. Kant here calls freedom not only "the highest degree of life" and "the property that is a necessary condition underlying all perfections" but also – if unregulated – the "basis for the most dreadful vices" (C 27:372, 344, 347). Kant describes duties to oneself as pertaining to one's self-regarding use of freedom (C 27:340): "The self-regarding duties are negative, and restrict our freedom in regard to the inclinations ... Just as the precepts of the law restrict our freedom in our dealings with others, so the self-regarding duties restrict our freedom with respect to ourselves" (C 27:347). According to Kant, "[t]he prime rule whereby I am to restrict freedom is conformity of free behavior to the essential ends of humanity [*die wesentliche Zwecke der Menschheit*]" (C 27:345). The essential ends of humanity consist in those "conditions under which alone the greatest use of freedom is possible and under which it can be self-consistent" (C 27:346). Chief among these conditions is reason's governance of choice: "Anyone who allows his person to be governed by his inclinations is acting contrary to the essential end of humanity, for as a free agent he must not be subject to his inclinations, but should determine them through freedom"; anyone who subordinates reason to inclination "contravenes the essential ends of humanity in his own person, and is acting against himself" (C 27:345). Such a person degrades herself, evinces disrespect for her humanity, and opposes love of honor (*Ehrliche*) (C 27:341, 347; cf. HN 23:403; MS 6:420, 465).

Kant does not employ the perfect–imperfect distinction for classifying self-regarding duties in these lectures. Nevertheless, many of the duties he explicates here in terms of essential conditions for one's greatest, self-consistent use of freedom are negative duties that he classifies elsewhere as PDS. For instance, he describes both killing oneself and excessive drinking as rendering oneself incapable of rationally employing one's freedom and powers (C 27:346). The supreme rule of self-regarding duties generally is, "[i]n all self-regarding actions, so to behave that any use of powers is compatible with the greatest use of them" (C 27:346; see also P 27:190; M I 27:1483–84). So on the Collins account, PDS prohibit acting against necessary conditions for one's greatest, self-consistent use of freedom.

3.1.2 According to the Collins notes, perfect duties to oneself preclude actions in which the agent's freedom directly conflicts with itself. All duties limit the agent's use of freedom. Some limit it in relation to the freedom of others; some by reference to ends she is rationally bound to promote. PDS, however, prohibit acts through which she uses her freedom to destroy itself, to undermine the necessary conditions of its fullest expression, or to subordinate itself relative to other objects. They thus

limit the agent's use of freedom in the most basic, straightforward way possible: by forbidding its immediate opposition to itself.

### 3.2 *Doctrine of Virtue*

3.2.1 The *Tugendlehre* characterizes perfect duties to oneself as prohibiting opposition to one's *moral self-preservation and moral health*. Kant here analyzes perfect as well as imperfect duties in terms of maxims, attitudes, or ends. The supreme principle of the *Doctrine of Virtue* is:

[A]ct in accordance with a maxim of *ends* that it can be a universal law for everyone to have. – In accordance with this principle a human being is an end for himself as well as for others, and it is not enough that he is not authorized to use either himself or others merely as means (since he could then still be indifferent to them); it is in itself his duty to make the human being as such his end. (MS 6:395)

Duties to oneself on this account are grounded in the requirement to make oneself, as a human being, one's end (MS 6:410). Part of what this requires is adopting the end of our own perfection. Duties to promote that end are imperfect; they concern our "moral prosperity" (MS 6:419). By contrast, PDS are "*limiting*" and "negative"; they are duties "of omission" that "*forbid* a human being to act contrary to the end of his nature and so have to do merely with his moral *self-preservation*." These duties concern "the moral health (*ad esse*) of a human being ... the preservation of his nature in its perfection (as *receptivity*)." Kant associates them with the Stoic dictum "live in conformity with nature ... that is, *preserve* yourself in the perfection of your nature." So part of making oneself one's end is taking one's moral self-preservation and moral health as ends not to act against (HN 23:387, 398). The *Tugendlehre* presents PDS as requiring us to reject maxims of action displaying disregard for or opposition to our moral health and moral self-preservation.

Kant's account of virtue's preconditions elucidates his notion of moral health. For example, he describes virtue as presupposing not merely the *autonomy* of practical reason, but also its *autocracy*, "the *capacity* to master one's inclinations when they rebel against the law" (MS 6:383; see also C 27:360–69).<sup>6</sup> Virtue both depends on and concerns the realization of inner freedom, understood as the capacity for rational self-constraint (MS 6:406–07). Inner freedom consists in sufficient resistance to the influence of affects and passions, and sufficient command over one's inclinations

<sup>6</sup> See Anne Margaret Baxley, "Autocracy and Autonomy," *Kant-Studien* 94 (2003), 1–23.

and impulses, so as to be receptive and responsive to reason's demands (MS 6:407–09).

The *Tugendlehre* divides PDS into duties to oneself as an animal and moral being, and duties to oneself as a moral being only, explicating them primarily by reference to the vices opposed to them (MS 6:421; and see 390, 404–05). Vices contrary to duties to oneself as an animal and moral being are hostile to our existence and proper functioning as humanly embodied rational beings. These vices include “self-disembodiment” (*Selbstenleibung*) through suicide or self-mutilation, sexual self-defilement (particularly through “unnatural” lust),<sup>7</sup> and self-stupefaction through immoderate use of food and drink. Self-mutilation “can ... be either *material*, *depriving* oneself of certain integral, organic parts ... or *formal*, *depriving* oneself (permanently or temporarily) of one's *capacity* for the natural (and so indirectly for the moral) *use* of one's powers” (MS 6:421). Duties to oneself as a moral being only consist “in what is *formal* in the consistency of the maxims of his will with the *dignity* of humanity in his person”; they forbid “depriving himself of the *prerogative* of a moral being, that of acting in accordance with principles, that is, inner freedom, and so making himself a plaything of the mere inclinations and hence a thing” (MS 6:420). Vices of lying, avarice, and servility express disregard for one's dignity and subjection of one's choice to empirical desire. Kant's identification of the duty to “*know* (scrutinize, fathom) *yourself*” ... in terms of your moral perfection in relation to your duty” as “the **First Command** of All Duties to Oneself” suggests that self-deceit regarding the purity of one's maxims is especially pernicious (MS 6:441, 430–31).<sup>8</sup>

3.2.2 The *Doctrine of Virtue* account vividly illustrates that the preservation, expression, and furtherance of each agent's freedom depends more fundamentally on her compliance with perfect duties to herself than with her compliance with other duties or on others' compliance with their duties toward her. PDS concern the protection of basic conditions of agency to an extent unrivaled by other duties.<sup>9</sup> Basic conditions of one's

<sup>7</sup> Kant thinks that our conception of ourselves as animal beings involves a teleological understanding of our impulses and organs. See my “Kant on the Wrongness of ‘Unnatural’ Sex,” *History of Philosophy Quarterly* 16 (2) (1999), 225–48.

<sup>8</sup> See Nelson Potter, “Duties to Oneself, Motivational Internalism, and Self-Deception in Kant's Ethics,” in Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 371–89.

<sup>9</sup> I regard the moral self-preservation and moral health (or their constituents) discussed in the *Tugendlehre* to be chief among the essential ends of humanity referred to in Collins.

agency constitute the subject matter of PDS. Some of these conditions are impervious to the conduct of others and all of them are constantly vulnerable to one's own misuse. Just as the agent's relationship to her own choice allows only her to perfect herself as a person (MS 6:386), it allows her to renounce, to impede, to oppose, or to cripple her freedom in ways others cannot. The agent is in a unique position to impair her inner freedom and autocracy; to undermine her sense of herself as a being with dignity and a subject of the moral law; to forfeit her standing as an honorable human being in the eyes of herself and others; and to imperil the grounds of her virtue. PDS are the only duties that forbid actions, vices, or maxims that threaten freedom in these profound and distinctive ways.

### 3.3 *Vigilantius*

3.3.1 Kant sometimes characterizes perfect duties to oneself in relation to – as directly expressing, correlative to, immediately grounded in, or constitutive of – *the right of humanity in our own person* (*das Recht der Menschheit in unserer eigenen Person*) (V 27:603). Kant invokes this notion occasionally in the *Metaphysics of Morals* (e.g. MS 6:236, 240) and discusses it in his handwritten preparatory notes on that work. He employs this notion most systematically in relation to PDS, however, in the *Vigilantius* lecture notes on the metaphysics of morals (1793–94); I draw mainly from these.<sup>10</sup>

*Recht* (like *Jus*) can mean a variety of things, including right, law, system of laws, and justice. Kant defines “a right” (*ein Recht*) as *a moral capacity* or power (*Vermögen*) *for putting others under obligation* (MS 6:237, 239; HN 23:262). An important notion of the right of humanity in my person is its authority to obligate me – to constrain, necessitate, compel, or coerce me – as a human being; so construed, the right of humanity in our own person is presupposed by, integral to, and exercised in all ethical obligation (V 27:579–80; MS 6:417–18). Humanity, the obligating party, is *homo noumenon*, the intelligible self, the ideal of a free, rational being who personifies the legislation of reason; the human being, the obligated party, is *homo phenomenon*, humanity in appearance, an agent who is “affected by the feelings of pleasure and pain, and must be coerced by the noumenon into the performance of duty” (V 27:593, 510, 579, 584). We may thus conceive of PDS as duties to which we are bound solely, directly, and strictly by humanity in our person. The “individual rights”

<sup>10</sup> These, like Collins, are student notes, so may contain confusions or misrepresentations.



of humanity in our person correlate to individual PDS, construed as specific restrictions on our conduct imposed by humanity in our person or entailed by its authority, or as the determinate actions or omissions thus necessitated. Insofar as we understand the right of humanity in our own person as a *law* immediately grounding PDS, it is “the law of the noumenon,” which subordinates the human being to humanity in our person and limits our action to conduct consistent with our rational freedom, respect for humanity, and the idea of our personality (V 27:597, 601–02; HN 23:390). Finally, Kant’s definition of “right” (*das Recht*) as the *totality of laws* through which what is right or wrong can be determined (HN 23:262) suggests that “the right of humanity in our own person” may sometimes refer to the whole system of laws governing one’s use of one’s person, of which PDS are constituents or to which they give determinate expression.

The *Vigilantius* notes divide all duties into two main categories. “[T]he strict duties of right” comprise “the right of humanity in our own person, and the right of human beings in regard to others”; “the broad duties of virtue” comprise “[d]uties stemming from the end of humanity in our own person, namely our own perfection, and from the end of other human beings, namely their happiness” (V 27:583; see also 542–43, 586; cf. HN 23:387). Perfect duties to oneself are *like* the other strict duties of right (and *unlike* imperfect duties to oneself) in that they “are derived from the concept of freedom through the law of non-contradiction, and thus analytically; and they therefore are such that they carry with them a necessity which also determines the act of duty itself” (V 27:587). PDS are *unlike* other-regarding duties of right (and *like* imperfect duties to oneself) in that they are not duties an agent can legitimately be externally coerced to fulfill. By labeling them “inner duties of right,” Kant identifies them not only as duties to oneself, but also as duties we can be brought to perform only through “inner necessitation” or “self-coercion,” i.e. constraint by humanity in our person (V 27:587, 592, 600). No appeal to the freedom of others is needed to ground these duties; none depends on external legislation.

Kant’s systematic catalogues of perfect duties to oneself in *Vigilantius* present them as limitations on the right of a human being to use his own person. Laying the ground for this classification, Kant proclaims that “the first right and duty of the human being in his own person [is:] The human being can never treat himself as a thing” and notes that this principle rests on the even more basic one that the human being as a sensible being “belongs to his own humanity as an intellectual being” (V 27:601).

Kant employs the metaphysical categories of relation – substance, causality, and community – to specify the objects of the right of the human being in his person (V 27:593–94, 601–02). PDS regarding one’s substance preclude (e.g.) killing, maiming, and sexually degrading oneself, for if a human being were entitled to dispose over his body, “he would then himself be master over his very personality, his *inner freedom*, or humanity in his own person” (V 27:601). PDS limiting the right of the human being regarding his causality forbid, among other things, a human being’s granting another person “unlimited disposition” over his powers through voluntary enslavement; his powers of external freedom belong to humanity in his person (V 27:594, 601; see also ZcF 8:345). Given the respect one’s humanity is owed from oneself and others, PDS regarding community forbid a human being’s damaging his own reputation as an honorable person or allowing others to damage it (V 27:594; see also MS 6:236). The “individual rights of humanity” are not exhausted by those presented within this tripartite schema. Kant enumerates many more PDS “in a fragmentary way” (V 27:604). Most of these prohibit lowering oneself in one’s own or others’ eyes, e.g. by lying, incurring debts, begging, or allowing oneself to be insulted without reprisal (V 27:604–07).

3.3.2 The *Vigilantius* account reveals perfect duties to oneself to be tightly bound up with freedom in several interrelated ways. Only PDS have as their sole and sufficient ground the freedom upon which their fulfillment depends: that of the subject of duty. Her own free, rational self directly determines PDS without appeal to the rights or dignity of others, or to the obligatory ends. PDS most fully and purely engage the capacity for self-constraint integral to human freedom: the power of our free, rational self to necessitate our sensibly affected self to fulfill our reason’s requirements. And because the necessitation to these duties involves only the humanity in the subject and carries through to determinate actions or omissions, PDS most purely and fully embody this self-constraint. Furthermore, since we realize our freedom most fully when we comply with the moral law from rational self-constraint (MS 6:382n; C 27:268–69, 464), fulfillment of PDS most clearly manifests our freedom: constraint to them is strict and solely self-imposed. Finally, only these duties reflect unsupplemented, unmediated constraint to action by the free, rational being to whom they are owed. Because I am autonomous, all duties follow from a law I give myself; others can bind me only through my own legislation (MS 6:417–18).<sup>11</sup> All obligation

<sup>11</sup> Timmermann notes this in “Kantian Duties to the Self,” 513.

is constraint by humanity in my person, so humanity in my person binds me even to the strictest duties to others. Indeed, ethical constraint by others amounts to no more than their prompting the humanity in me to compel my compliance with my own reason's legislation (V 27:521; C 27:269).

### 3.4 Summary

The preceding accounts of perfect duties to oneself differ. The Collins account portrays PDS as prohibiting uses of freedom antagonistic to necessary conditions for one's fullest rational expression of it. One degrades oneself, treats oneself as a mere means, and brings one's freedom into immediate conflict with itself, when one acts in opposition to essential conditions of one's greatest use of freedom. The *Tugendlehre* account portrays PDS as forbidding opposition to one's moral self-preservation and moral health – negative ends that one who holds her humanity as a supreme, objective, negative end must also hold. Action on maxims instantiating vices contrary to PDS undermines essential formal and material conditions for moral agency and the perfection possible through it. Acting on such maxims involves self-degradation because we thereby pursue satisfaction of empirical desires at the expense of our rational efficacy or dignity. The *Vigilantius* account portrays PDS as duties to which we are immediately and inflexibly bound by our ideally free, rational self. By violating these constraints, we repudiate an essential condition of our moral agency: the authority of humanity in our person to necessitate us to action. While the *Tugendlehre* explicates PDS in terms of ends not to act against and vices to avoid, the *Vigilantius* notes explicate them in terms of forbidden actions (e.g. acts of suicide rather than the vice of suicide).<sup>12</sup>

Yet despite differences among these three accounts, and in keeping with my thesis, they are united in their depiction of PDS as bearing tighter, more urgent, or more fundamental relations to freedom than other types of duties. Kant's precise explication of the relation between freedom and PDS varies from one account to another. On all three accounts, however, PDS prohibit directly acting against our rational freedom – humanity in our person – upon which our moral obligation, virtue, and dignity depend.

<sup>12</sup> On reconciling tensions between depictions of PDS as inner duties of right and as duties of virtue, see Gregor, *Laws of Freedom*, 113–27. See also Yvonne Unna, "Kant's Answers to the Casuistical Questions Concerning Self-Disembodiment," *Kant-Studien* 94 (2003), 454–73, esp. 455–58.

## 4 PRIMACY

The uniquely close, crucial connections perfect duties to oneself bear to freedom suggest multiple construals of their primacy. My catalogue of them is not exhaustive; and I barely begin to explore connections and tensions among them. Nevertheless, I hope that distinguishing among various sorts of primacy will facilitate a fuller appreciation of the position PDS occupy in Kant's moral theory, as well as a better understanding of some of Kant's own claims about them. I organize my discussion of these somewhat arbitrarily labeled, interrelated (sometimes overlapping) kinds of primacy by reference to the accounts of PDS that best elucidate them.

*4.1 Essential conditions for one's greatest use of freedom*

Perfect duties to oneself have a *rational* primacy. This is suggested especially by the Collins account – and the *Groundwork* account of these duties in relation to the formula of the universal law of nature (G 4:421–22, 424) – which depicts these duties as following most immediately from the basic rational requirement that one's willing not contradict itself. Imperfect duties follow from our inability rationally to endorse universal indifference to others' happiness and to renounce cultivation of our own natural and moral abilities. Perfect duties to others follow from the rational requirement of consistency among all persons' expressions of freedom. But PDS preclude actions in which the agent's own freedom immediately opposes itself; they thus rule out the most flagrant sort of practical contradiction: intrapersonal (corresponding to a self-regarding duty) and within maxims of action (corresponding to a perfect duty). This explains why PDS consistently hold the first place in Kant's taxonomies: they are the first string of requirements issued by reason; they prohibit the actions and maxims that reason most immediately rejects.

*4.2 Moral self-preservation and moral health*

Perfect duties to oneself have a *practical* primacy stemming from the extreme, negative consequences that violating these duties has for the agent's subsequent uses of freedom in fulfilling her duties, cultivating her virtue, and promoting her permissible, discretionary ends. All accounts suggest this sort of primacy, though the terminology of moral self-preservation and moral health found in the *Doctrine of Virtue* does so most vividly. Intoxicating or mutilating oneself, for example, undermines material

conditions of one's agency. Miserly avarice subordinates one's freedom to the hoarding of material goods; servility degrades one's worth in relation to others (MS 6:432–36). Because these vices conflict so directly with one's dignity, they undermine one's recognition of oneself as a bearer of humanity and subject of duty more profoundly, pervasively, and easily than do disrespect and maltreatment from others.

### 4.3 *The right of humanity in our own person*

The Vigilantius account of perfect duties to oneself in relation to the right of humanity in our own person (here construed mainly as its moral power to constrain us) most clearly suggests and best illuminates the remaining sorts of primacy.

4.3.1 Perfect duties to oneself share in three sorts of primacy attributable to duties to oneself as such. First, duties to oneself have a *hierarchical* primacy, for the most fundamental moral requirements can be thought of as duties to oneself. A duty is “to” the being who puts the subject under obligation (MS 6:412, 442). Because we are autonomous (but unholy), we constrain ourselves to comply with all moral principles – and indeed with moral law itself. So the matter of all moral requirements can be thought of as duties to oneself in a broad or formal sense: that is, as ethically legislated obligatory acts, which our humanity compels us to perform.<sup>13</sup> Second, duties to oneself have a *logical* primacy: if our own humanity did not obligate us – if there were no duties to oneself – we would not be bound to any duty. In the *Tugendlehre*, Kant explains that if there were no duties to oneself, there would be no duties to others either:

For I can recognize that I am under obligation to others only insofar as I at the same time put myself under obligation, since the law by virtue of which I regard myself as being under obligation proceeds in every case from my own practical reason; and in being constrained by my own reason, I am also the one constraining myself. (MS 6:417–18; see also V 27:521, 579)<sup>14</sup>

<sup>13</sup> My suggestion here is that all acts to which one obligates oneself are (in that way and for that reason) duties *to* oneself. Though Kant usually restricts the category “duties to oneself” to a subset of ethical duties, he sometimes identifies ethical or “internal” duties as duties to oneself in terms of “form” (e.g. V 27:587) or discusses duties to oneself in ways that imply reference to all ethically legislated duties (e.g. MS 6:417–18). Only section 4.3.1 considers duties to oneself in this wider sense.

<sup>14</sup> In the case of duties falling within Kant's standard, narrower classification of duties to oneself, one's humanity binds not only as the legislator or enforcer of the moral law but also as the being whom the particular duty is “to.” (See Andrews Reath, “Self-Legislation and Duties to Oneself,” in *Agency and Autonomy in Kant's Moral Theory: Selected Essays* [Oxford University Press, 2006], 231–49, esp. 245–46.) One might therefore doubt that the necessary, pervasive presupposition

Kantian moral obligations to particular acts as duties to others presuppose rational self-legislation of, and rational self-constraint through, the moral law. Constraint by humanity in our own person (duty to oneself in what I am calling “a broad or formal sense”) is thus a necessary condition even of duties to others. Third, duties to oneself have a *constitutional* primacy.<sup>15</sup> Because fundamental, general moral requirements constitute an essential link in the normative chain culminating in particular duties, duties to oneself (in a broad or formal sense) help constitute all duties. Moreover, since one’s own humanity compels one’s compliance with duties to others, any duty to another person seems to be also (in a broad or formal sense) a duty to oneself.<sup>16</sup>

These sorts of primacy suggest the primacy of PDS in particular. First, many of the most fundamental moral requirements may best be considered a species of *perfect* duties to oneself. They are strict, negative requirements, concerning determinate (albeit inner) actions – e.g. to reject all non-universalizable maxims. They do not presuppose the obligatory ends. We act wrongly anytime we fail to comply with them. Furthermore, Kant ascribes greater stringency and priority to perfect duties than to imperfect ones (V 27:537, 600–01). So insofar as any moral requirements or self-imposed constraints qualify as PDS, they have primacy relative to those better viewed as imperfect.

Second, insofar as we attribute logical primacy to duties to oneself as constraints imposed by humanity in one’s person, we should recognize the logical primacy of *perfect* duties to oneself in particular. PDS immediately follow from the right of humanity in our person and purely instantiate the moral self-constraint presupposed by all duties. Indeed, PDS have a logical primacy relative to imperfect duties to oneself, for the former follow immediately from a presupposition of the latter. The right of humanity in our person is the sole and sufficient ground of PDS. PDS are thus independent of the end of our own perfection, which is a necessary and

of ethical self-constraint resolves concern about the logical possibility of duties of *this* type. Yet Kant seems correct in thinking that *self-constraint as such* is the basis upon which *any* notion of a duty to oneself might appear conceptually vulnerable. The problem is whether one person can be simultaneously *obligans* and *obligatus* (V 27:508–10, 579, 593; MS 6:381–82, 417–18). There is no comparable, special puzzle about how one who legislates or compels compliance with a law can be owed specific treatment under that law. If duties to oneself, however narrowly construed, were impossible *because* one could not constrain oneself, all duties would be impossible.

<sup>15</sup> This also is suggested by the above-quoted MS 6:417–18 passage.

<sup>16</sup> Timmermann defends the identification of all duties as duties to oneself in “Kantian Duties to the Self,” 512–15. Similarly, Potter claims that moral self-constraint renders all duties “partially” duties to oneself in “Duties to Oneself,” 376. Gregor suggests on a different basis that all duties are “indirectly” duties to oneself in *Laws of Freedom*, 128, esp. n. 1.

proximate ground of imperfect duties to oneself. But imperfect duties to oneself depend on the right of humanity in our own person, for our own perfection is an obligatory end precisely because our humanity compels us to adopt it.

Third, *perfect* duties to oneself have a special constitutional primacy because they alone are constituted as duties immediately and solely through humanity's authority.

[R]ights and duties ... are determined by humanity ... In this sense the rights of humanity in our own person, or rights and duties to oneself, can be thought of no otherwise than as the highest, since they are *directly dictated by humanity itself*, whereas the rights of a human being towards other human beings depend *only indirectly* on that. (V 27:579–80, emphasis added)

PDS are direct, determinate constraints of humanity on the human being, strict constraints on action that issue immediately from humanity in our person. Although they fundamentally presuppose humanity in human beings, juridical duties (as such) are constraints human beings impose on one another; indeed, they may be externally legislated and enforced. Imperfect duties to oneself are determined broadly and indirectly by humanity, through the imposition of an end. For Kant, PDS are “the highest duties of all” because they are strictly and directly determined by the humanity that at least indirectly determines all duties (V 27:604).

4.3.2 Perfect duties to oneself have a *conceptual* primacy: they most clearly exemplify the self-legislation and self-constraint that is integral to and characteristic of Kantian morality. It is PDS that Kant explicates directly in terms of the right of humanity in our person. PDS emerge when an agent lays upon herself an obligation of the strictest necessity, that only she can compel herself to fulfill, and that has its source only in her legislative reason. They thus most purely and fully embody Kant's distinctive conception of ethical obligation – which extends to all duties – as constraint by humanity in one's own person. They also best reflect his notion of morality as based on the conception of the human being as a free being who, because of his freedom, “binds himself through reason to unconditional laws” (RGV 6:3).

4.3.3 Because only they are necessitated solely and immediately by humanity in the agent's person, and because they concern the respect strictly owed to oneself as a free, rational being, perfect duties to oneself have a *motivational* primacy. One has the deepest and closest practical

relationship to humanity in oneself. It seems unreasonable to expect someone who lacks moral motivation adequate to fulfill PDS, in which the preservation, authority, and dignity of humanity in one's person are at issue, nevertheless to have moral motivation adequate to fulfill other sorts of duties.

4.3.4 More importantly, perfect duties to oneself have a *normative* primacy. For one thing, they have the strongest obligating grounds (*rationes obligandi*). The right of humanity in our own person "is the supreme condition of all laws of duty, since the subject otherwise would stop being a subject of duty (person) and would have to count as a thing" (HN 23:390; see also V 27:547, 604).<sup>17</sup> Our moral personality presupposes our humanity's capacity to necessitate our sensible choice; a being incapable of rational self-constraint can have no duties. One could never be obligated to act against the right of humanity in one's person, for that would be an obligation to repudiate a fundamental condition of obligation in general and of one's own obligations in particular. The specific obligating grounds Kant associates with particular PDS (self-respect, independence, honor, etc.) are inextricably linked with the right of humanity in our person. So when, in cases of apparent conflicts of duty, obligating grounds of PDS and other duties conflict, those of PDS prevail (V 27:493; MS 6:224; V 27:537, 600–01).

Another aspect of the normative primacy of perfect duties to oneself is this: there is something uniquely devastating about their violation (even beyond the forward-looking, practical considerations of section 4.2), making compliance with these duties more urgent and these duties more "binding" than others (C 27:433). Because PDS directly reflect and immediately follow from the right of humanity in our person, violations of PDS amount to or express a rejection of our humanity's authority over us; and since this authority is a necessary condition of our moral agency, violations of PDS constitute or imply a renunciation of it (V 27:604). We can understand many of Kant's provocative claims about PDS along these lines. Kant's statement that vices contrary to PDS as a moral being only "make it one's basic principle to have no basic principle and hence no character, that is, to throw oneself away and make oneself an object of contempt" seems to reflect both the practical and normative primacy of PDS (MS 6:420). Kant also says that one who violates a perfect duty to oneself "turns oneself into a thing" and is someone

<sup>17</sup> "Es ist aber die oberste Bedingung aller Pflichtgesetze weil das Subject sonst aufhören würde ein Subject der Pflichten (Person) zu seyn und zu Sachen gezählt werden müßte." See Gregor's discussion, *Laws of Freedom*, 46–48.



of whom nothing moral can be demanded (C 27:341, 346, 373; V 27:604). By violating PDS one makes oneself into a thing practically, by using oneself as a mere means in one's maxims. Nothing can be morally demanded of someone who does not recognize his humanity's authority to obligate him. Because one practically "throws away" one's dignity and freedom, one makes oneself an object of contempt – something a non-human animal or an inanimate object never could be.

4.3.5 Finally, perfect duties to oneself have a *structural* primacy within Kant's system: they set parameters for other rights and duties. The right of humanity in our person is presupposed by all duties, so none may oppose it. Other duties' consistency with the right of humanity in our person requires their accommodation to PDS, for PDS alone express what agreement with the right of humanity in our person inflexibly and fundamentally requires; indeed, they are those requirements. Consequently, PDS hold a foundational place in the structure of Kant's system of duties and shape the remainder of that system; other duties must conform to them.<sup>18</sup>

One manifestation of the structural primacy of perfect duties to oneself is their limitation of what juridical duties to others we can acquire, and what juridical rights others can acquire against us, even with our consent.<sup>19</sup> For instance, conduct contrary to PDS cannot constitute the matter of a rightful contract; it is impossible to give others juridical rights to treat one as a thing by contract. Regardless of how voluntarily they are entered, slavery contracts and contracts for prostitution and concubinage,<sup>20</sup> for example, are unenforceable, devoid of the power to bind (MS 6:278–79; F 27:1335–36). Kant explains the legal impossibility of such contracts and their attendant rights and duties in many ways. The Vigilantius notes portray "inner right," the proper relation of the human being to humanity in her person, as a constraint upon "outer right," rightful relations among human beings. The right of the human being to use her own person is limited by the right of humanity in her person, which prohibits her from disposing over herself as a thing that she owns – including by "ceding,"

<sup>18</sup> The structural primacy of PDS can be seen as reflecting their rational primacy (section 4.1) as well as their logical and constitutional primacy (section 4.3.1).

<sup>19</sup> This can also be understood in light of normative primacy (section 4.3.4). On the right of humanity in our own person as a constraint on acquired rights or positive law, see Leslie A. Mulholland, *Kant's System of Rights* (New York: Columbia University Press, 1990), esp. 262–65; and Sven Arntzen, "Kant on Duty to Oneself and Resistance to Political Authority," *Journal of the History of Philosophy* 34 (3) (1996), 409–24.

<sup>20</sup> Herein, one party makes herself into "a consumable thing ... by contract" (MS 6:360; see also V 27:638, 640–41).

“granting,” or “handing over” the substance of her body or the totality of its forces to others for them to use however they like (V 27:593–94, 601–02; see also F 27:1379). More straightforwardly, Kant sees a contradiction between, for example, slavery contracts’ content and their presupposition of the moral personality of the parties bound by them: “a contract by which one party would completely renounce its freedom for the other’s advantage would be self-contradictory, that is, null and void, since by it one party would cease to be a person and so would have no duty to keep the contract” (MS 6:283). All duties, including juridical duties, are directed to persons; PDS demarcate the boundary of conduct consistent with one’s moral personality; so only those purported juridical duties that can accommodate their subjects’ compliance with PDS are legally possible. The content of rightful contracts and acquired juridical rights and duties is thus limited by consistency with PDS, which may thus be viewed as structurally prior.

4.3.6 There are two kinds of objection one might make to the dramatic claims about the singularly dire implications of violating perfect duties to oneself (sections 4.3.4–4.3.5). First, one might object that because of the role of the right of humanity in our own person in grounding, constituting, and compelling compliance with all duties (section 4.3.1), violation of *any* duty must practically deny our humanity’s binding power, so there is no *special* problem with violating PDS.

We do evince a lack of respect for our humanity’s authority if we consistently and pervasively violate our duties; and we indicate greater respect for it the more conscientiously we fulfill them. But not every violation of duty equally constitutes rejection of our humanity’s authority any more than every violation of a sovereign’s laws constitutes treason, every act of insubordination constitutes mutiny, or every deviation from religious mandates constitutes apostasy. Perfect duties to oneself have the agent’s humanity as the sole *obligans*, have the right of humanity in our person as their immediate ground, and are strictly required as a matter of respect for her humanity. Other duties are about other things and respond (also and more directly and fully) to other normative considerations. There is a longer normative chain leading to these other duties; and in rationally imperfect, radically evil beings like us, many more opportunities for it to break. The expressive meaning of fulfillment or violation of PDS is different both because of the content of these duties and the normative path leading to them. When we do not properly fulfill imperfect duties or when we violate juridical duties to others, there are many ways to understand

these moral failures that are more plausible – because more directly relevant to the acts in question – than as rejections of our humanity’s moral power to constrain us. This is not the case with PDS.

Second, one might object that some of these dramatic claims imply that in violating perfect duties to oneself one loses one’s humanity, its dignity, or all claims to respect from others. In discussing servility, for instance, Kant proclaims, “one who makes himself a worm cannot complain afterwards if people step on him” (MS 6:437). Similarly, he says that if someone sells a limb, “he turns himself into a thing, and then anyone may treat him as they please, because he has thrown his person away” (C 27:346, see also 372–73; V 27:604). Such statements are disturbing in themselves and apparently conflict with passages suggesting it is impossible to forfeit one’s humanity, dignity, or claim to others’ respect – e.g. “even though somebody may be a bad man, the humanity in his person is entitled to respect” (C 27:373; see also MS 6:463–64).

We can interpret the “one who makes himself a worm cannot complain” sort of statement as drawing out a logical consequence of acting on maxims through which one subordinates one’s freedom to other objects. There is a contradiction between a maxim of treating oneself as a thing and a maxim of complaining about others’ treating one as a thing. If you practically endorse – by acting on – a maxim in which you serve as a mere means, you must rationally accept others’ treating you that way as well. It would be inconsistent of you to complain of this, given how you treat yourself (M I 27:1428). On the other hand, if you insist that it is wrong for others to treat you as a thing, then you must hold that you are not a thing. But then you cannot consistently treat yourself as though you were. This argument calls us on our tendency to make exceptions for ourselves. Kant’s point is neither that we can do whatever we choose to ourselves as long as we do not demand that others treat us with respect, nor that we morally ought to allow others to do whatever they like to us if we have once violated our perfect duties to ourselves. His point is rather that since we cannot rationally renounce constraints on others’ treatment of us and our title to demand adherence to those constraints, we must recognize similarly grounded constraints on how we treat ourselves.

One might counter the preceding argument by asserting that your relationship to yourself is different from your relationship to others, such that there is no contradiction between your disposing over yourself as a thing and your objecting to others’ doing so. An obvious way to argue this point is to claim that your relationship to yourself is one of ownership. When it comes to things that you own, you have rights of disposal

and exclusion: you can dispose over things that you own, but others may not (without your consent) (MS 6:245, 249). So if you own yourself, you can complain about others' use of you without you yourself being similarly constrained. According to Kant, however, a human being "is not his own property – that would be a contradiction; for so far as he is a person, he is a subject, who can have ownership of other things. But now were he something owned by himself, he would be a thing over which he can have ownership" (C 27:387; see also MS 6:359). Since the right to own property is enjoyed only by free beings, I cannot simultaneously regard myself as an owner of property and as a thing; so I cannot rationally regard myself as authorized to dispose over myself as a thing that I own. There is a contradiction within a maxim of disposing over oneself as one's own property; this is unsurprising, since it violates a perfect duty to oneself (G 4:424).

Humanity is not an empirical property that can be acquired or lost in time, but an attribute we ascribe to persons from the practical standpoint (G 4:448; MS 6:418). We "throw away" our humanity every time we treat ourselves in ways it is proper to treat only things. But we do not as a result subsequently lack humanity. However often we "cast off" our humanity by violating our perfect duties to oneself, our continuing consciousness of obligation forces us to recognize that our freedom, humanity, and its dignity are still with us. Similarly, regardless of how brazenly or frequently we degrade ourselves, others who engage with us practically must attribute freedom, humanity, and its dignity to us as well (MS 6:463; see also 283; C 27:418).<sup>21</sup> While we may, for example, justifiably criticize as hypocritical a prostitute who complains of her customers' degrading her (by paying for sex that she consensually provides for a fee), her violation of the right of humanity in her person does not render their use of her ethically permissible. Humanity in our person constrains what others may do to us even when we fail properly to respond to it.

## 5 CONCLUSION

Kant's lectures on ethics and the *Doctrine of Virtue* provide distinct accounts of perfect duties to oneself. The Collins account portrays PDS as limiting our use of freedom to agreement with essential conditions for its fullest expression; violations of PDS bring our freedom directly into conflict with itself. The *Tugendlehre* account in relation to moral self-

<sup>21</sup> Kant's claims about how the humanity of a criminal limits the ways we may punish him support this reading (MS 6:331, 362–63, 463; V 27:556).

preservation and moral health vividly depicts ways in which our physical and moral integrity, rational efficacy, and virtue depend on our fulfillment of PDS. The Vigilantius account portrays PDS as strictly and immediately entailed by our humanity's moral capacity to obligate us as human beings, such that their violation repudiates this authority and implies renunciation of our moral personality.

Together these accounts show that PDS have a unique relationship with freedom. No other class of duties is so immediately grounded in rational freedom or so directly responds to the dignity of humanity; so fully and purely embodies the nature of ethical obligation or autonomously legislated moral duties; or so directly concerns autocracy, self-respect, and other essential formal conditions for the expression of a human being's rational freedom. The exceptionally direct, vital, and fundamental relations PDS bear to freedom generate several sorts of primacy.

I cannot here explore whether any one account of perfect duties to oneself is primary. Each informs and enriches the others – and our sense of what treating one's humanity as an end in itself entails. Perhaps the Collins account most perspicuously presents the rational necessity of these duties, requiring only a few moves beyond the *Groundwork* III premise that, as a practically rational being, I must regard myself as free (G 4:448). Since freedom is the will's property of being a law unto itself, independent from determination by alien causes, I must regard my will as self-legislating: neither as lawless, nor as determined by causal laws, objects external to my will, or the legislation of other beings. What rules could I legislate to myself as a free being? Certainly one demanding that my freedom never conflict with itself – that it neither destroy itself, nor undermine the conditions of its expression, nor subordinate itself relative to other objects. This is what PDS require, on the Collins account. Meanwhile, the Vigilantius account best reveals PDS as the *sine qua non* of all duties. For on this account, PDS follow solely, determinately, and immediately from a fundamental condition of all other duties: the right of humanity in our own person, upon which our very standing as subjects of duty depends. Despite the richness of the *Doctrine of Virtue* account, we should not overlook these others if we wish to understand the nature and primacy of PDS.

*Duties to and regarding others*

Robert N. Johnson

## I INTRODUCTION

Kant holds the views that

- (1) we have an ethical obligation to others to adopt the happiness of others as our end, yet
- (2) normally no particular other person has a claim on our assistance in advancing her happiness.

These two claims are at the heart of Kant's conception of our obligatory end regarding others. Although failing to help a particular person in certain situations can be evidence – perhaps conclusive if the circumstance be dire enough – of my having failed to make the happiness of others my end, in general not helping this or that person is not evidence of this. This is not to say that there is no room for debate over how demanding this obligatory end is. Kant's readers differ over whether, for instance, one may forgo efforts to improve the lives of others only if one is perfecting oneself instead (thus pursuing one's other obligatory end) or fulfilling some strict duty of greater importance.<sup>1</sup> But they appear to agree that making the happiness of others one's end does not imply any particular person has a right to one's help *tout court*, and that this is because this ethical obligation is a wide and imperfect one.

Unfortunately, this together with a further feature of duties generates a puzzle. As we will see, Kant himself recognized and made much of the fact that the person *toward* whom one owes a duty is not necessarily the person or thing *with regard to which* the duty is concerned. If I promise to keep your elderly parents company while you are away, my duty is to *you*, not them, and I am obliged to keep *them*, not you, company. The duty is

<sup>1</sup> See, for instance, the contrast between Marcia Baron's position, in *Kantian Ethics Almost without Apology* (Ithaca, NY: Cornell University Press, 1995) and Thomas E. Hill, Jr., in, for instance, "Kant on Imperfect Duty and Supererogation," *Kant-Studien* 72 (1971), 55–76.

to you, not them, but *regarding* them, not you. This raises the question: of what does this distinction between a duty “to” someone as opposed to merely “regarding” her consist? Luckily, there is a ready answer, one I will argue that Kant himself would be aware of, namely,

- (3) you have a duty *to* someone to do something if and only if that person has some claim or right that you do it.

What distinguishes your elderly parents from you is that you have a right to my keeping them company and they don't. And that is all there is to my owing this duty *to* you and not *to* them. So shouldn't Kant simply make use of this easy answer to the question?

On the standard reading of Kant's views, the answer would be “No.” Though I have a duty to others to make their happiness my end (as (1) states) *no one* has a right to my aid in pursuing their happiness (as (2) states). So it cannot be that, as a result of my having a duty *to* others, and not merely regarding them, others must thereby have a right to my adopting it as my end (as (3) states). The problem with the standard reading is the lingering issue of what the “to” as opposed to the merely “regarding” relation consists of. If (3) does not explain the difference, what does? This is the puzzle I discuss in what follows.

In my view, the standard understanding of Kant's views I describe above should be abandoned and Kantians should accept all three claims. Indeed, I think exploring this solution throws light on the general question of what it is to have a duty “to” anyone at all. My contention will be that a certain understanding of collective rights turns out to show that others in fact do have a right that we adopt the happiness of others as our end, even if this does not imply that any individual has a right that we make her happiness our end.

My plan will no doubt seem quixotic to some simply because they hold the standard understanding of Kant's view. On that view, duties are *to* particular people at particular times, yet at the very same time, these duties do not generate a corresponding right on the part of that particular person toward whom we have those duties. But, again, those who accept this standard picture are left with only vague metaphors alluding to some difference between having a duty *to* rather than merely regarding others. My hope is that they will consider the following in the spirit in which I offer it, as a solution to the puzzle of the directionality of obligation in Kant.

In what follows, I begin by discussing Kant's views on the nature of the beings to whom we can owe duties. As it turns out (notoriously) we can only have duties to human beings because duties *to* a person arise because

we use or intend to make use of a rational will. Next, I discuss views held by Andrews Reath and Allen Wood about what it means to have a duty *to* a person. Although I do not entirely disagree with their views, I add what I think are some needed details to explain the metaphors used to depict the directionality of obligation. I then turn to the relationship between duties to a person and the claims she may have on us as a result of such duties. My account is far from exhaustive, since I only aim to defend the view that duties and rights are reciprocal to the extent that this reciprocity is what distinguishes those *to whom* we have duties from those only *with regard to whom we have duties*.

## 2 KANT ON THE OBJECTS OF OBLIGATIONS

In the *Doctrine of Virtue*, Kant offers a principle that he regards as the foundation of the division between obligations toward (*gegen*) oneself and obligations toward others. This principle categorizes “the *beings* in relation to whom ethical obligation can be thought” (MS 6:412). Think of this principle as analogous to a constitutive rule of a game, the game of “obligation.” Obligation has at least two playing positions, one active, the other passive, the person who has the obligation and the person to whom it is owed. In Kant’s version of the game, one can play this game alone, but because one is playing in two distinct positions, it is not like playing solitaire. Just as one can play against oneself or others in chess or backgammon, but not at solitaire, so one can play “against” oneself or others in the game of obligation. These two playing positions are the “being that is under obligation and the being that puts him under obligation” (MS 6:412).<sup>2</sup> Thus, the principle creates a division based on differences

<sup>2</sup> “das eine, welches das subjective Verhältnis der Verpflichteten zu dem Verpflichtenden der Materie nach.” Moral principles are laws fully rational beings would jointly will as legislators for a realm of ends. The rules of the game of ethical obligation – the moral principles that are to guide conduct – are thus themselves not *created* by the player in his role as the being who has put one under obligation. It is only by making a move within the game, not in “legislating” its rules, that one is put under obligation. Thus the players, insofar as they are rational, adopt a set of rules for the game of ethical obligation. But insofar as they are playing the game, they are bound by those rules. The being who “puts one under obligation” is the being who is, because of the circumstances, identified by the rules of the game as the source of a claim on your conduct. Think of, for instance, promising as a game with rules that constitute who is the promise giver and who the promise receiver, when a promise has been made, how one can be released from a promise, what can be promised and what cannot, and so on. One way of understanding Kant’s view of this game is that the rules are “valid” or “binding” on its players just in case those who play the game, independently of their roles *within* that game, under the right conditions, could rationally choose its rules as binding on them within the game for certain kinds of social interactions. One such rule is that the promise giver is “put under obligation” by the promise receiver by making the promise under the right circumstances. For each like part of the game (e.g. gratitude, benevolence) there



among beings that are eligible as players in the game of obligation. And the beings that can play these roles and put one under obligation are oneself and others.

Kant argues that the category “others” as far as we know includes only humans, although non-human rational agents, if there were any, could also obligate us. If possessing a rational will is necessary for one being to put another under obligation, however, and only human agents possess such a will, then only such beings can obligate us. Kant recognizes that many think that we have all sorts of duties to non-rational non-agents. However, he regards this thought as an error: someone who thinks this is “mistaking his duty *with regard* to [*in Ansehung*] other beings for a duty *to* those beings” (MS 6:442).<sup>3</sup>

These notorious positions have been widely discussed and I have nothing directly to add here.<sup>4</sup> I am interested primarily in what light this sheds on what it means to have a duty *to* a person. Kant is pointing out that you can have duties *regarding* yourself, others, animals, or divine entities without having a duty strictly speaking *to* any of these entities. In his view, many duties which are *in regard to* or *concerning* others (or the environment, animals, etc.) may strictly speaking be duties *to* oneself or some other person. This, in his view, turns out to be the case for all duties regarding non-humans – the (non-human) natural world and God. In his defense, Kant thought that we do have duties regarding animals and the environment, even if some may think his view abhorrent that these are not duties we owed *them* but rather owed to ourselves. No doubt it would seem more acceptable for a duty regarding animals or the environment, if not a duty *to* these things, to at least be a duty *to others*. But, setting aside

exists a rule that establishes when, to whom, and under what conditions one player puts another under obligation of some sort. One comes to have a duty *toward* a being by having that being *put one under obligation* by that being. And one is *put under obligation* by some being by making a move within the game of ethical obligation. See Andrews Reath’s discussion of the different “positions” in moral obligation in “Self-Legislation and Duties to Oneself,” in Mark Timmons (ed.), *Kant’s Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 349–70.

<sup>3</sup> “zu welchem Mißverstande er dadurch verleitet wird, daß er seine Pflicht in Ansehung anderer Wesen mit einer Pflicht gegen diese Wesen verwechselt.”

<sup>4</sup> See, for example, Lara Denis, “Kant’s Conception of Duties Regarding Animals: Reconstruction and Reconsideration,” *History of Philosophy Quarterly* 17 (2000), 405–23; Paul Guyer, “Duties Regarding Nature,” in *Kant and the Experience of Freedom* (Cambridge University Press, 1993), 303–34; Elizabeth Pybus and Alexander Broadie, “Kant’s Treatment of Animals,” *Philosophy* 49 (1974), 375–83; Jens Timmermann, “When the Tail Wags the Dog: Animal Welfare and Indirect Duty in Kantian Ethics,” *Kantian Review* 10 (2005), 129–49; Allen Wood, “Kant on Duties Regarding Nonrational Nature 1,” *Proceedings of the Aristotelian Society, Supplementary Volumes* 72 (1998), 189–210.

the merits (or lack thereof) of Kant's views on non-humans, it illuminates two features of his general picture of obligation.

The first is the explanation of how one comes to be obligated *to* someone. To have an obligation *to* someone is to be under "moral constraint by that subject's will" (MS 6:421):<sup>5</sup>

the constraining (binding) subject must, first, be a person; and this person must, *secondly*, be given as an object of experience since the human being is to strive for the end of this person's will and this can happen only in a relation to each other of two beings that exist (since a mere thought-entity cannot be the *cause* of any result in terms of ends). (MS 6:442)

This appears to state that obligations arise because of a relationship between wills. However, I read it against the backdrop of formulations of the categorical imperative that focus on the humanity in persons. These formulations forbid me to "use another person [if he] cannot possibly agree with my way treating him, and so cannot himself share the end of the action" (G 4:429–30). Using a person in a way to which he could not (rationally) agree fails to respect his capacity to set and pursue ends. It treats his will as if it were nothing more than an instrument for achieving your own ends rather than as a rational capacity to set its own ends. By ruling out such acts, the rational will of a person that you use *puts you under moral constraint*. Thus, the first condition above, that we have duties *only* to persons, follows from the fact that it is only when in some way or other we treat *rational wills* as means that Kant thought a moral constraint is generated. A moral constraint is generated by the status of the rational capacity to set and pursue ends.

"Means" is a causal notion. And for Kant causality is a feature of the world only insofar as it is a possible object of our experience. So one cannot rationally intend to use anything that is not a possible object of experience. This is the source of the second condition set out above on the objects of obligation. We have duties only toward beings that are (possible) objects of experience. Only such beings, as far as we can know, bring about effects through the causality of their wills, and so only such beings are the causes of ends that we can be obligated to help them further, as

<sup>5</sup> This voluntaristic conception of moral obligation arises also at the beginning of Kant's discussion of duties to oneself (MS 6:417–18). He puzzles over how the human being that imposes an obligation can be the very same being who is put under that obligation. One who imposes the obligation can release the person who is under that obligation whenever he wants. But one who can release himself from an obligation whenever he wants is under no obligation at all. Kant's proposed "solution" is to distinguish the *noumenal* from the *phenomenal* person. See Reath, "Self-Legislation."

well as be “made use of” in our own actions to further our own ends. This, then, gives us both the circumstances under which others put us under obligation as well as the restriction that this introduces. Others put us under obligation when we intend to make use of their wills, and our use of their wills is restricted by principles to which they could rationally consent.

The phrases “making use of someone’s will” and “treating as means” suggest something far too narrow to be the origin of all of our moral obligations, even those Kant acknowledged. I am obligated to make the ends of others my own ends yet that would be so even if the others whose ends I make my own are totally unrelated to my projects, and so not so much as imagined to be used as means. So it seems Kant’s own view is that I am obligated to make the ends of others my own, but not because I intend to make use of the wills of any of the particular persons whose ends I end up furthering under that heading. A defense of Kant’s position, then, would have to appeal not to intending to make use of the particular wills of particular persons but also to our intentions with regard to a collective: our intention to make use of *the rational wills* of someone or other, whoever he may turn out to be.

Do we necessarily intend to make use of the rational agency in others? Kant has a compelling case that in fact we do. Kant holds that:

there is ... *one* end that can be presupposed as actual in the case of all rational beings (insofar as imperatives apply to them, namely as dependent beings), and therefore one purpose that they not merely *could* have but that we can safely presuppose they all actually *do have* by a natural necessity, and that purpose is *happiness*. (G 4:415–16)

It is unavoidable for human nature to wish for and seek happiness (MS 6:387). Yet individually, each person’s ability to achieve this end is so limited as to be for all practical purposes nonexistent. Therefore, insofar as I am rational, I must recognize that “many cases could occur in which [I] would need the love and sympathy of others” (G 4:423). By “love” and “sympathy,” Kant could not mean simply these emotions. The finitude of our individual powers leaves each requiring the actual assistance of another if we are to think our happiness is even possible. Assistance by some rational agent *at some point* is a necessary means of its achievement.

Kant evidently thought that the necessity of willing our happiness together with our individual inability to achieve it implied the proposition that we cannot rationally will that everyone refuse to adopt anyone else’s happiness as his end. The latter is the proposition he would need in order to gain the conclusion that the happiness of others is an obligatory

*end*. But I cannot see how it does establish that it is an obligatory end. For it is only *acts of assistance* from others at some time that is necessary for our happiness, *whatever their ends may be* in performing those acts. It makes no difference to achieving our own ends whether others have our happiness as their end. What we need is their assistance. Our inability to rationally will that no one ever help anyone does not provide a sufficient premise to conclude that we must therefore adopt as our end the happiness of others.<sup>6</sup>

It is nonetheless significant that we must (if indeed we must) have our own happiness as our end and that we cannot achieve anything close to this individually – that we are interdependent creatures. Given these two facts, it does seem that we could not rationally will a world governed by a law that no one shall help anyone ever. In this broader sense, then, even in having our own happiness as our end, we cannot rationally avoid intending to make use of the rational agency of some person at some one or other time. So while promises and contracts are the most obvious cases, we intend to make use of the rational agency of others – though not necessarily anyone’s rational agency in particular – simply in pursuit of our own happiness. That is sufficient for “others” to put us under obligation.

We can, then, have duties only to persons on Kant’s view, those whose rational will(s) we intend to use. Note, however, that this does not imply that non-persons have no value.<sup>7</sup> To have value is to be the object of a rational will. An object of the will is whatever we intend to will or voluntarily bring about, broadly construed (KpV 5:57–66).<sup>8</sup> Something can be the object of our choice even if it is not something to which one owes a duty. A person or thing *with regard to which* one has a duty, for instance, is an object of one’s will, in the sense that one’s will aims to affect her or it.

If value is grounded in rational willing, it does not follow that non-humans have a value only as a means to some human end. If a duty is also *regarding* a person, then we must of course treat the humanity in that person, the person with regard to whom we have the duty, at the same time as an end in himself, whatever else the duty requires that we do regarding him. But if a duty is toward someone with regard to, say, his property, the

<sup>6</sup> I argue this more fully in my “Self-Improvement: An Essay in Kantian Ethics” (unpublished manuscript).

<sup>7</sup> I believe Wood assumes just this in his attack on what he calls “the personification principle” in “Kant on Duties.”

<sup>8</sup> See also my “Value and Autonomy in Kantian Ethics,” in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, vol. 11 (Oxford University Press, 2007), 133–48.

fact that the property's value is *derivative* does not imply that it is merely *instrumental*. Whatever practical (i.e. non-aesthetic) value the property has, whether instrumental or non-instrumental, it has it because it is an object of a rational will.<sup>9</sup>

If duties toward a rational agent arise because we are treating her agency as having instrumental value, then duties to oneself would presumably arise in this same way. We treat our own agency itself as having instrumental value, since we must inevitably make use of it (MS 6:418). The exercise of our own will is therefore also the ground of our having duties to ourselves. The distinction between duties to oneself and others is therefore just the distinction between obligations arising because we treat our own wills as means as opposed to those arising because we treat the wills of others as means.<sup>10</sup> By making use of yourself as a rational will, your will puts itself under moral constraint toward itself.<sup>11</sup> Just as we use the rational agency of a cab driver to get across town, we put our own rational agency to use in carrying out our plans.

The person *to whom* one owes a duty (e.g. the person whose will our actions make use of in some instance) need not be featured in any description of what is required by that duty. The game of ethical obligation, so to speak, can involve a third person (or thing), though this third role is not, or not directly, as it were, a player. For instance, suppose I promise my neighbor that I will help to take care of his children if he will help to take care of mine. My duty would then be *to my neighbor*, *not to his children*, even if the duty *concerns* doing something with, to, or about his children but not with, to, or about him. If I fail to help to take care of his children, I will have failed in a duty I owed *to him*, not to his children. It is his will, if you like, that I have made use of in promising, and hence constrains my own. In this case, what our duty *concerned* was set by what we (rationally) agreed *to*: the care of each other's children.

In the case of some duties to oneself a single agent occupies all three positions in the obligation game: the person who has the duty, the person to whom it is owed, as well as the thing or person it concerns. Notice that Kant himself explicitly employs the other possible combinations as well: I

<sup>9</sup> I believe part of the confusion many have about Kant's views are based on an inadequate understanding of his conception of value.

<sup>10</sup> Cf. Allen Wood, *Kant's Ethical Thought* (Cambridge University Press, 1999), 325.

<sup>11</sup> As a side note, this represents another distinction in imperfect duties: imperfect duties toward others are not owed to anyone in particular, while imperfect duties toward yourself *are* owed to someone in particular, namely, you. Thus, their being imperfect cannot be based on their not being owed to anyone in particular.

have duties to *others* that also concern *them*, namely, their happiness. I have duties to *myself* that concern *others*, for instance not to lie to them. And I have duties to *others* that concern *myself* – for instance, the duty to try, as much as may be possible, to cultivate my sympathetic feelings. I can also be the one who puts others under obligations in the same ways, so others can also have these duties to me, concerning either myself or other people. And, finally, although only persons can be under obligation or put others under obligation, for Kant a duty can be toward *persons* even if what it requires does not concern *persons* at all, as is the duty to *myself* not to despoil *nature*. In this case, I am not required to do or avoid doing anything concerning myself or any other person, even if the duty is to *myself*.<sup>12</sup>

To summarize, Kant's account of obligation has three positions: the person who is *under obligation*; the person or things *with respect to* which he has the obligation; and the person(s) *to whom* he has that obligation, whose will(s) puts him under obligation, or to whom the duty is *owed*.<sup>13</sup> The circumstance under which you come to have a duty toward a person is that you treat (in a broad sense) her will as having instrumental value. Since you may not treat a rational will in a way that does not at the same time acknowledge the value of its nature as a capacity to set and pursue ends, you must treat wills in ways that they could rationally agree to.

The above account I believe begins to add detail to what others have said about the Kantian idea of having a duty *to* a person. Before I complete the account, it would be helpful to review why a couple of Kantian views of this relation on offer do not give enough detail. Allen Wood, for instance, argues that a duty is to a person, S, just in case it is “grounded on the requirement to respect humanity in the person of S.”<sup>14</sup> So far as it goes, this account amounts to understanding “toward” as giving the location of the thing to respect, its being “in the person of” the person *to* whom one has a duty. The “to” relation in effect is what gives your obligation a directionality by locating the person whose humanity grounds your obligation. This explanation is fitting, given Wood's overall construal of Kant's ethics as a value-based ethical theory in which the value of humanity is what grounds our obligations. The source of obligation is

<sup>12</sup> This is not to say that all duties have both someone to whom the duty is owed, as well as something *owed to someone*, since that leaves out meritorious duties – benevolence, for instance – in which one is required to do something that is not owed to a person.

<sup>13</sup> This is how Kant puts it, in his gloss on the duty of self-development: he states that a human being “owes it to himself (as a rational being) not to leave idle” capacities he might one day use (MS 6:444).

<sup>14</sup> Wood, *Kant's Ethical Thought*, 325.

the value of the humanity of a person, and hence it is the humanity of whoever is generating the obligation through its value that establishes the relation of being *to* that person.<sup>15</sup> It's as though humanity's value magnetizes it and obligation is an iron needle drawn toward it.

I think more needs to be said, however. As I explain below, imperfect duties, for instance *to others*, give others a claim to our performance of these duties, but this does not necessarily generate a claim for any individual person over our behavior. That we have a duty to others to promote their permissible ends gives the group "others" a collective claim over this. But no given individual in that group has a claim to our adopting their happiness as our end. Nevertheless, it is the humanity in the given individual person that is the ground of our obligation to adopt as our end the promotion of their permissible ends. So even if the humanity in a given person grounds our obligation, it is the group "others" to which the duty is owed.

Another elaboration of the Kantian understanding of the "to" relation is given by Andrews Reath. His view is that:

the directionality and sense of the "to" is given by fact that the fundamental moral requirement (expressed in one way) is to adopt certain attitudes towards those with whom we interact. We have a general duty to show proper regard which is specified by substantive principles that pick out certain facts about a person's condition, needs, interests, circumstances and so on, as the source of reasons for one to treat or view those individuals in certain ways.<sup>16</sup>

Reath's view aims to avoid an overly voluntarist account of duties *to* a person, one in which you have a duty *to* that person whose will constrains you through, for instance, a tacit voluntary agreement. He avoids this by adopting a common Kantian constructivist model of the origin of moral principles, joint legislation. The idea is that obligations originate in moral principles, and moral principles are just those principles that are arrived at by way of a specific deliberative procedure. In particular, moral principles are those principles we would, were we free and rational, agree to as the principles we, as we actually are, are to live by. Those principles would designate who has what duty, to whom or what, and regarding whom or what. Who has the duty is the "subject" in such principles, and to whom it is owed is the "object." When and under what circumstances, if any, the object can release the subject of an obligation is also determined through the joint legislation procedure. Thus, there is nothing mysterious about

<sup>15</sup> I argue against this value-based understanding of Kantian ethics in "Value and Autonomy."

<sup>16</sup> Reath, "Self-Legislation," 362.

the “to” relation in duties for Reath. A person has a duty to himself just in case some principle is or can be validated through the joint legislation of rational agents in which she is in both the subject and object.

The thought that the “to” position is the source of reasons for doing what I do, however, is not sufficient to distinguish the “to” relation from the “regarding” relation. Your child’s welfare is a source of reasons for me, even if my duty to care for her this afternoon is to you, because I promised you I would. Of course, it is not a source of reasons in the same way that my promise to you is. But what is this difference? The “to” as opposed to the “regarding” position in the principle is supposed to be in a certain way authoritative in the circumstance. That difference, according to the model of Kantian ethics Reath is adhering to, depends on how co-legislation establishes these positions in moral principles. But the question is, when co-legislators establish who is in the “to” position, what exactly is it they are establishing? It is surely something more than being a source of reasons. To anticipate a bit, the person in the “to” relation has certain claims and normally liberties. So in my view we need a principle of the sort that would justifiably establish the “to” relation in such a way that it establishes who has a right to the performance, who has a right to complain if there is a failure of performance, who has a liberty to waive the duty, and so on.

I think that a sharper explanation of the “to” relation is needed at least as an addition to views such as Wood’s and Reath’s. These explanations do not sufficiently distinguish the person to whom we have a duty and persons our duty is regarding. There is more that needs saying about this distinction.

### 3 OBLIGATIONS AND CLAIMS

If I promise you that I will help to take care of your child this afternoon, I have acquired a duty *regarding* your child, but no duty *to* your child. My duty is *to* you, the one to whom I have made the promise, albeit *regarding* her. My duty is not *regarding* you, at least insofar as the substance of my duty is to chat with your child and not you. A standard way of looking at this difference between duties “to” and “regarding” a person relies on the fact that the person *to* whom one has a duty has, indeed must have in virtue of one’s having a duty to her, a right or claim to one’s doing what one has a duty to do.<sup>17</sup>

<sup>17</sup> The idea that in some way duties and claims are reciprocal originates with Wesley Hohfeld in his *Fundamental Legal Conceptions*, W. Cook (ed.) (New Haven: Yale University Press, 1919); see, for example, 39. See also Joel Feinberg, “The Nature and Value of Rights,” *Journal of Value Inquiry* 4 (1970), 243–57; and John Mackie, “Can There Be a Rights-Based Moral Theory?” *Midwest Studies in Philosophy* 3 (1978), 350–59.



There is strong evidence that Kant accepted, or at least would accept, something at least very close to this standard view. In the introductory sections to the *Metaphysics of Morals*, he states that “rights have reference to duties” and that it is only

through the *moral imperative*, which is a proposition commanding duty, from which the capacity for putting others under obligation, that is, the concept of a right, can afterwards be explicated. (MS 6:239)

Passages of this sort imply that there is some sort of reciprocity between rights and duties. One might doubt the attribution of this reciprocity to Kant on the grounds that for him rights have to do with what is legally enforceable, and so with what is not in the domain of ethics. But, aside from the fact that these passages come prior to the division between the *Doctrine of Right* and the *Doctrine of Virtue*, there is plenty of textual evidence that this reciprocity is assumed in the latter as well. In speaking of the duties of virtue owed others arising from the respect due them, he states that “every human being has a legitimate claim to respect,” that the respect is “owed them” and is “a right to which he cannot renounce his claim” (MS 6:462–64). I take it, then, that Kant recognizes something like *moral* as well as *legally enforceable* rights, and recognizes their reciprocity. It is the categorical imperative that is the ground of our capacity to put others under obligation and of their capacity to put us under obligation, and is by way of this the source of our rights over them and theirs over us. It is this very reciprocity, as I see it, that explains the “to” relation.

On this standard view, then, to say one has a duty *to* someone is to say that that someone has a right or a claim to your doing whatever it is that you have a duty to do. This is a claim that a person *with regard* to whom one has a duty need not possess. The sort of “right” in question here is a claim grounded in the right-holder’s control over the performance of the person against whom she holds the right. The distinction between the person *to* whom one has a duty and the person or thing *regarding* whom one has it is thus based on the correlativity of claims *against* a person and that person’s duties *to* the rights bearer, or, at a first approximation:

C: S has a duty to P to  $\phi$  if and only if P has a claim against S that S  $\phi$  (and P under certain circumstances possesses a liberty to waive that claim).

On doctrine C, the person *to* whom one has a duty is necessarily the person with a claim to your performance, while the person (or thing) *with regard to* whom you have a duty is the person or thing you are to

do something to, with, about, or whatever, but not necessarily the person who has any claim on your doing it. It *might* be the same person, of course, as when I promise you, say, to return your lawnmower. In this case, I have a duty *to* you – that is, you are in possession of a claim against me that I return it – that is also *regarding* you. But my having the duty *regarding* you, to return your lawnmower, does not give you any control over my returning it. For instance, you might give me your lawnmower, but then your mother makes me feel guilty for having taken advantage of your shortsighted generosity and extracts a promise from me to her that I would return it to you. While you, having given the lawnmower to me, would have no direct right to its return, your mother would have such a right, since, in virtue of promising her, I came to have an obligation *to* her to return it to you.

In the same way, you could acquire a duty to your mother regarding yourself, if, say, you promised her that you would not engage in any more acts of rash generosity of the sort that led you to give me the lawnmower. In that case, your mother comes to have a right against you that is regarding you, in particular, that you not rashly give away your belongings.

The correlativity of rights *against* and duties *to* primarily concerns the *core* of rights.<sup>18</sup> If I have a duty to you to care for your child for the afternoon, you have a right over my behavior that afternoon, and the core of this right consists in your moral claim against me that I look after her, together with in this case a liberty to waive that claim against me. But the right may also consist in your having other claims and liberties, and my having duties, that are associated with this core claim, such as your having the right to complain if I fail and my having a duty to clear my calendar for the afternoon to make room for your child.

Kant held that we have an imperfect duty to help to promote the (morally permissible) ends of others. However, being under this obligation to others does not imply that any particular person has a right to our help. Thus, while if I, for instance, lie to some random person X, I have thereby violated X's right to the truth, if I fail to help some random person Y achieve some morally permissible end of his, I normally have not thereby violated any right that Y had to my help. If there are such duties, does it follow that C, the doctrine that duties *to* persons are reciprocal with rights held by those persons, is false?

<sup>18</sup> I understand the notion of the "core" of a right as it is explained in Carl Wellman, *Real Rights* (Oxford University Press, 1995).

## 4 DUTIES TO OTHERS AND COLLECTIVE RIGHTS

It does not if there is something answering to concept of a group or collective (moral) right. For it follows from C together with the claim that each person has a duty to others to promote their permissible ends that others thereby have a right to our help to promote their ends. If there can be rights held collectively, the right to our help would be held collectively by the group “others,” not by any individual in that group. There are of course worries about collective rights, at least insofar as these rights are not reducible to the rights of individuals who make up collections of persons. For instance, Will Kymlicka states that “it is individual, sentient beings whose lives go better or worse, who suffer or flourish, and so it is their welfare that is the subject-matter of morality.”<sup>19</sup> Many think such considerations rule out any sort of collective rights. However, we can understand collective rights in a relatively unproblematic way. Kymlicka’s view, for instance, of collective rights is as a right that individuals possess *in virtue of* being members of a group. Thus, a moral system may grant a right to certain individuals in virtue of their belonging to some group, for instance, *because she is a citizen, member of a religious group*, and so on. All that this would take is a normative structure that assigns such rights to individuals. This notion of a collective right does not require attributing a normative property to a collective, but only to individuals in virtue of their membership in the collective. And as long as one accepts the idea of some sort of normative structure that assigns rights to individuals, all that is required is an argument that there are some rights morality assigns to individuals in virtue of their membership in a collective.

This alone does not explain the puzzle concerning the Kantian obligation to others to promote their permissible ends. For that puzzle is that no individual within the collective “others” has a right to my help, and so no individual within that collective has such a right by being a member of that collective entity. So what I am proposing instead is that the collective right to your help that “others” possess against you is the right of some one or other of those persons to hold you accountable for not helping some one or other, where the person holding you accountable need not be the person whom you help, and the person whom you help need not be, in virtue of that, the person who has a right to hold you to account. Let me explain.

<sup>19</sup> Will Kymlicka, *Liberalism, Community, and Culture* (Oxford University Press, 1991), 242.

Our duty is to make the good of others our own end. That will ordinarily translate into particular helping actions at times when there is opportunity to help someone who needs it, there is no dramatic sacrifice required, no more important duties conflicting, and so on. Now suppose such an opportunity to help presents itself to you, and yet you do nothing to help. Suppose it is obvious that there is no reason for you not to help, and yet, with full awareness, you do nothing. Because you have a duty to make the good of others your end, it seems to me that others now have a right to ask for some account of yourself. Should it turn out that you never lift a finger for others, and some know this, it seems that those who know now have standing to blame you for your failure. However, although those who know may have a right to blame, perhaps even to chastise, none, no one in particular, has the right to blame or chastise you for your not helping *her*, not even the particular person you had an opportunity to help. They have standing to blame you only for rejecting the happiness of others as your end. This would also be the case for the person who needed your help. She would have a right to complain about your behavior – should she be in a position to know that you had opportunity, and so on – but not on her own behalf, rather as a member of the group whose happiness you have not made your end. That is, in general, it is normally *any* member of the group “others” who has standing to blame and resent your behavior, should they have sufficient evidence and be in a position to know these things about you.

Suppose, as C supposes, when we have a duty to others, then (in the case of Kantian imperfect duties) others acquire a right that we have a certain end, namely, their happiness. This right is composed of some set of standing, depending on the circumstance, to complain if we do not have this as our end, to tell us to adopt it as our end, and so on. *Who* is in the right normative position to execute these claims depends on other factors. If you do not make the happiness of others your end, the person who can complain cannot complain simply that you haven’t helped her. And although others possess a liberty to waive your obligation to make their happiness your end, no individual person, according to C, is in the right normative position to exercise that liberty. Such a person would have to be in an impossible fiduciary relationship with all others to be in such a position.

That Kant’s view assigns rights and liberties to persons in virtue of being in the class of “others” seems plausible at least in one respect: it is in virtue of our membership in the class of rational human agents that we have rights to certain kinds of treatment. It may seem as if this is only

an ornate way of saying that because we're rational agents we have certain rights. Is it really a collective right, in the sense that the above legal rights are collective rights? Perhaps it's not because we are members of a group, the group of rational human agents, that we have certain rights; it is simply because we *are* rational agents that we have them. However, this could be said *mutatis mutandis* of the collective legal rights above: to say that it is because I'm a member of the collective "citizens of the USA" that I have the rights delineated in the amendments to the Constitution is just an ornate way of saying that it is because *I am* a citizen that I have those rights. The only difference is that my legal rights are born of a voluntarily constructed system, while my humanity is a metaphysical fact about me. But this is insufficient to give the "in virtue of being a member of" any different weight in the legal case than it would have in the moral case. If I have rights in virtue of membership of a voluntary association, then my rights as a human rational agent I also have in virtue of my membership in the latter group.

Suppose there is something special about each person's humanity that makes a difference to whether we say that our human rights are rights we have in virtue of being members of the group, "humanity," and instead say that it is in virtue of our humanity *tout court* that we have such rights. The only important question is whether the group "others" constitutes a group, membership in which grants collective moral rights to help. It seems that if there is a justified moral principle assigning such rights, it does. The question is whether there is something fishy enough about "others" as a group to challenge the justification for any such principle. It doesn't seem to me that there is.

If what I have been saying about collective rights so far is right, then we must distinguish two groups within the collective "others" as it occurs in the duty: I have a duty *to others* to adopt as my own end the happiness *of others*. There are those others whose happiness I must further by pursuing such an end – the others *with regard to which* I have this duty – and there are those others who have a right to complain and blame me if I don't – the group *to whom I owe* this duty. These two groups need not be the same. The group "others," whose happiness I must adopt as my own end, is not, for instance, constituted by every other human being in existence. Kant does not explicitly rule this out, but his exposition of the duty appears to limit "others" more to "those who have to contend with great hardships" (G 4:423). That I count the projects of Bill Gates or Warren Buffett as my own somehow does not capture the spirit of the moral duty to make the happiness of others my end. But perhaps "great hardships"

is too limiting. Suppose the world contained only people with medium hardships. It seems we'd still have the obligation. I will assume, then, that when Kant speaks of the imperfect duty to others of promoting the happiness of others, he at least means to restrict the latter group "others" to "others *who need our help*," taking "need" to be a relatively easy bar to meet. But it remains true that, while Bill Gates and Warren Buffett do not normally belong in the category of "others" whose happiness we ought to promote, they do *ceteris paribus* belong in the category of others who are entitled to blame us if we fail to make this our end.

If we accept C, then if we have an obligation to others who need our help, the need of those others thereby provides the ground of a right for others – not necessarily the same others – to complain or blame us for not having the happiness of others as our end, should they be in a position to judge this to be the case. This, however, is at best a marginal improvement over saying that some particular person has a right to our help. A more plausible view would be that the right acquired by others by our having a duty to help others pursue their ends is a right some one or more of them have against us that we help some one or more of them to pursue their ends. But the person whom we help has no right to our help, in virtue of the fact that we are discharging our obligation to others by helping her. So the person whom we help does indeed have a right against us that we help. But he doesn't have a right that we help him in particular, only that we help some one or other of the people constituting "others who need our help." It just so happens that in helping him, we discharge this right he had against us that we help some one or other. And in having a right to our help, others have a claim that no one individually has a liberty to waive, because no one is in the right normative position to do so. No one can release me from my duty to others by exercising a liberty to forgo their claim to my help because no one has such a liberty.

Admittedly this has the sound of a detached morality. Perhaps we would find it somewhat alienating to receive someone's help because he says he "must help some one or other of you others, and I've decided it will be you." But my story is not inconsistent with the existence of special relationships to our friends and family. After all, you might decide to help this or that person because they are friends or family. It is certainly not inconsistent with the thought that we should care about them, or have feelings even for the "others" whom we help. That the duty to make the happiness of others our end has this feature does not really add any new layer to the question of whether Kantian ethics is too detached.

I think that we can, then, retain the doctrine in C while holding that we have duties to others by understanding the nature and content of the right that others acquire against us in a certain way. In particular, while the duty to make the happiness of others our end results in the rights of others to do so, it does not result in rights to particular actions, or rights owed to particular persons.

## CHAPTER II

# *Duties regarding animals*

*Patrick Kain*

### I INTRODUCTION

In some of the most widely cited, and certainly the most *criticized*, passages from the *Metaphysics of Morals*, Kant infamously insists that we human beings have duties to all human beings: “a human being is under obligation to regard himself, as well as every other human being, as his end” (MS 6:410); but “a human being has duties only to human beings (himself and others)” (MS 6:442).<sup>1</sup> While Kant recognizes many moral constraints upon our behavior toward non-human animals, he insists that these are only duties “*with regard to* these animals,” rather than duties “*to* those beings” (MS 6:442–43).<sup>2</sup> “Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other” (MS 6:462), but animals are “things,” not persons, and “*respect* is always directed only to persons, never to things” (KpV 5:76). If anyone thinks otherwise, that is due to an “amphiboly in his concepts of reflection.” It seems as if Kant thinks an animal is no more worthy of our concern than is a turnspit on which we might choose to roast it.

This position on the nature and scope of fundamental “moral status” and its practical implications, both for the treatment of so-called “marginal cases” of seriously immature or radically disabled human beings and for the treatment of non-human animals, has been a source of much consternation.<sup>3</sup> Prominent philosophers have suggested that one of Kant’s greatest mistakes *überhaupt* was his failure to appreciate the nature of non-

<sup>1</sup> Translations in this essay are those of the Cambridge Edition, except where noted otherwise, or where a quoted passage is not included in an already published Cambridge Edition work.

<sup>2</sup> In what follows, I will often use “animals” as shorthand for “non-human animals.”

<sup>3</sup> “To have moral status is to be morally considerable, or to have moral standing. It is to be an entity towards which moral agents have, or can have, moral obligations.” Mary Anne Warren, *Moral Status: Obligation to Persons and Other Living Things* (Oxford University Press, 1997), 3.



human animals and their moral significance.<sup>4</sup> Kant is regularly accused of (i) drawing an arbitrary distinction between the moral status of all human beings and that of non-humans which cannot be reconciled with the actual condition of human infants and severely disabled adults, (ii) a fundamental failure to consider the nature of non-human animals and acknowledge their similarity to humans, (iii) a failure to recognize that the moral constraints on human behavior toward non-humans should be based on the nature of those animals, rather than in incidental effects of our behavior upon humans which turn upon highly contingent features of human psychology, and (iv) a failure to regard animals as the proper objects of human concern in their own right.<sup>5</sup> These charges appear to cut to the heart of Kant's ethics, and addressing them has seemed to demand either the outright rejection of Kantian ethics or significant alteration of its trademark focus on human dignity.<sup>6</sup>

A better appreciation of Kant's commitments in a variety of disciplines reveals that Kant had a deeper understanding of human and non-human animals than is generally recognized and will help address, at least from Kant's perspective, many of the familiar objections to his account of our "duties regarding animals." In section 2 ("The basis of moral status," below), I will review some of Kant's core principles about the nature of moral obligation, which structure his thoughts about the moral status of human beings and non-human animals. In the next section ("The nature of animals"), I will consider in some detail Kant's account of the nature of and distinction

<sup>4</sup> Responses by Christine Korsgaard and Peter Singer in Vadim Vasilyev's "International Kant Interview 2004–2005," [www.philos.msu.ru/community/staff/vasiliev/Kant\\_Interview/Kant\\_Interview.htm](http://www.philos.msu.ru/community/staff/vasiliev/Kant_Interview/Kant_Interview.htm).

<sup>5</sup> For example, Alexander Broadie, and Elizabeth M. Pybus, "Kant's Treatment of Animals," *Philosophy* 49 (1974), 375–83; Peter Singer, *Animal Liberation: A New Ethics for Our Treatment of Animals* (New York: Random House, 1975); Tom Regan, *The Case for Animal Rights* (London: Routledge & Kegan Paul, 1984); Martha Nussbaum, *Frontiers of Justice: Disabilities, Nationality, Species Membership* (Cambridge, Mass.: Harvard University Press, 2006); and Warren, *Moral Status*. Elements of these criticisms can be found in the work of Arthur Schopenhauer and Albert Schweitzer; see Heike Baranzke, "Tierethik, Tiernatur und Moralanthropologie im Kontext von §17 Tugendlehre," *Kant-Studien* 96 (2005), 336–63.

<sup>6</sup> For significant concessions by Kantians on some of these points, see Christine M. Korsgaard, *The Sources of Normativity*, Onora O'Neill (ed.) (Cambridge University Press, 1996); "Fellow Creatures: Kantian Ethics and Our Duties to Animals," in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values*, vol. xxv (Salt Lake City: Utah University Press, 2005), 77–110; and "Interacting with Animals: A Kantian Account," in Thomas Beauchamp and R.G. Frey (eds.), *The Oxford Handbook on Ethics and Animals* (Oxford University Press, in press); Jens Timmermann, "When the Tail Wags the Dog: Animal Welfare and Indirect Duty in Kantian Ethics," *Kantian Review* 10 (2005), 128–49; Allen W. Wood, "Kant on Duties Regarding Nonrational Nature 1," *Proceedings of the Aristotelian Society, Supplementary Volumes* 72 (1998), 189–210, *Kant's Ethical Thought* (Cambridge University Press, 1999), and *Kantian Ethics* (Cambridge University Press, 2008).

between humans and non-human animals. With this account in hand, I will turn, in the section that follows (“Moral implications”), to Kant’s case for claiming that we have duties to every human being and significant duties regarding non-human animals that are grounded in their nature. Finally, in “Two Kantian alternatives,” I will consider Kant’s account in relation to some recently proposed Kantian alternatives.

## 2 THE BASIS OF MORAL STATUS

Kant insists upon a sharp distinction between beings with *dignity* (*Würde*) and those with mere *price* (*Preis*). Price is a kind of relative value, a value something has if it is related in the correct way to something else, in particular to the needs or desires of human beings. By contrast, dignity is a kind of absolute and intrinsic value; something with dignity “is raised above all price and therefore admits of no equivalent,” it cannot “be replaced by something else” (G 4:434).<sup>7</sup> Kant claims that what gives a being dignity and marks it out as an “end in itself” is its innate rational capacity (*Fähigkeit*) for autonomy, a predisposition (*Anlage*) to “personality,” the capacity to “legislate” the moral law and to act out of respect for the moral law, “freedom ... under moral laws” (MS 6:223, 418, 434–35; G 4:428, 435–36; RGV 6:27).<sup>8</sup>

In Kant’s theory there is a deep connection between dignity and moral obligation. In Kant’s terms, only beings with dignity are capable of “passive” and “active obligation”: only beings with dignity can be obligated or obligate others. “Duty to any subject is moral constraint by that subject’s will” (MS 6:442). Moral obligations can be articulated as the demand to respect the dignity and autonomy of every rational being (G 4:428–36). Thus, the second formula of the categorical imperative demands: “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” (G 4:429). Lest it appear that Kant is simply begging all of the relevant questions about the scope of moral status, we must note that Kant employs the terms “humanity” and “personality” in a technical sense to refer to certain capacities or predispositions of the will, which may or may not turn out to be ascribable to all and only human beings.

<sup>7</sup> This is not to deny that Kant draws some distinctions within “price.” See, for example, G 4:428, 434; MS 6:434.

<sup>8</sup> Kant rejects the possibility that organisms, in general, could be “final ends” or ends in themselves (KU 5:425–35), contra G.F. Meier, *Philosophische Sittenlehre* (Halle: Hemmerde, 1753–61), §975.

Since a “duty to any subject is moral constraint by that subject’s will,” an obligator (a being *to whom* one can have a duty, a being capable of “active obligation”) must have a will that can impose a moral constraint upon the obligated, and the obligated (one capable of “passive obligation”) must have a will that can be constrained by the obligator. Thus, Kant isolated two necessary conditions for genuine moral status: we can be obligated only to a being that is both (i) a “person,” a being with a free will “standing under the moral law,” and (ii) is “given as an object of experience,” so that we can recognize that it can obligate us and so that we can, through our actions, have some bearing upon it and/or its ends (MS 6:442).

Regarding the first condition, Kant famously argued that neither “theoretical” philosophy nor empirical investigation can establish that there is *any* such absolute freedom, any “freedom under moral laws.” “Experience lets us cognize only the law of appearances and hence the mechanism of nature, the direct opposite of freedom” (KpV 5:29). Kant came to insist that the reality of absolute freedom, or freedom under moral laws, can only be established in practical philosophy, by the “fact of reason.” We are each “immediately conscious (as soon as we draw up maxims of the will for ourselves)” of the moral law; the moral law is given to us as “the sole fact of pure reason” and this fact leads us to the concept of freedom and the postulation of its reality in us (KpV 5:29–31). “The categorical imperative proves for morally practical purposes” that at least some of us “human beings” are free” (MS 6:280n).

Yet we must not neglect the second condition and its implications for the determination of moral status. Kant’s insistence that we can only have obligations to persons who are “given as an object of experience” suggests that experience and the biological, psychological, and anthropological theories, concepts, and judgments through which we make systematic sense of the objects we are given in experience must play a significant role in helping to determine in a naturalistically respectable way which objects of experience should be considered to be the presentation of the relevant kinds of predispositions; a suggestion confirmed by Kant’s appeal to “experience” and his employment of biological and psychological terminology in the discussion of our duties regarding non-human animals and of the moral relationship between human parents and the children they conceive (MS 6:280, 442). We must investigate salient aspects of Kant’s investigations in these disciplines.

## 3 THE NATURE OF ANIMALS

Kant articulated a naturalistic framework for systematic biological and psychological investigations.<sup>9</sup> Kant insists that, in natural science, we must seek to identify a system of efficient or “mechanical” causal laws responsible for observable regularities, but there are phenomena that resist such an understanding (KU 5:387–88, 372–76, 401–04). To bring such regularities “under laws,” a set of teleological concepts is needed, including the concept of an *organism*, a “natural end” which is a teleologically organized and self-organizing whole, organized for life and reproduction (VR 2:429; KU 5:376). When using such concepts, we must still observe the maxim that “in a natural science everything must be explained *naturally*” (GtP 8:178; cf. KrV A544/B572, A773/B801). One should seek a systematic and parsimonious account which relies upon analogies to observed powers and eschews both unnecessary and unhelpful complexity and direct appeals to divine intervention. A “philosophically appropriate,” “naturalistic” explanation of the regularities observed among organisms favors an “epigenetic” theory of the reproduction of organisms combined with a commitment to real biological species and a doctrine of original “predispositions” (GtP 8:168–69; BBM 8:102; GtP 8:178). In reproduction, adult organisms of a species produce a new organism of their species, endowed at conception with the species’ specific organization, a set of “predispositions” (*Anlagen*) and “germs” or “seeds” (*Keime*) that were originally implanted in the species’ first members (KU 5:423).

In psychology we find an account of animals or “living” organisms, those endowed with “sensation and choice” (MS 6:442), which extends this biological framework. Since animals can perceive and respond to changes in their immediate environment in ways that (most) plants cannot, Kant insists upon judging animal behavior as a product of inner principles (even if less than fully conscious or self-conscious ones): living beings have the capacity to move themselves according to the power of choice, that is, in virtue of their representations. Kant argued that although mental representations can, in general, be cognized and explained naturalistically, neither can be fully explained “materialistically.” The mental representations that are essential constituents of the genuine psychological regularities we observe, Kant argued, must be regarded as states of an immaterial

<sup>9</sup> The argument of the next several paragraphs is developed in more detail in Patrick Kain, “Kant’s Defense of Human Moral Status,” *Journal of the History of Philosophy* 47 (2009), 59–102.

soul (though not necessarily a simple, substantial, or immortal soul) (KrV B419–20; KU 5:460). Kant insisted that animals are not “mere machines,” but have souls with a *vis locomotiva*, because the mental representations that guide their behavior cannot be realized in matter (KU 5:457, 464n). In animals, the “faculty of desire” is linked with a “faculty of cognition” or “intuition” which gives rise to representations (via the senses, but also via reproductive and anticipatory imagination) and a “faculty of feeling pleasure or displeasure” in conjunction with a representation (MS 6:211).<sup>10</sup> For systematic reasons, Kant favored an account of animal reproduction and original ensoulment according to which each animal is endowed from its conception with the biological and psychological predispositions of its species.<sup>11</sup> The predispositions and propensities of an animal species, which may underlie or manifest themselves in a variety of instincts, acquired inclinations, and habits, serve as causal grounds for the occurrence of certain thoughts, feelings, desires, and behaviors.<sup>12</sup>

We humans can be “immediately aware” of our own representations, especially those representations upon which we act; based on observable similarities between our actions and the behavior of non-human animals (such as dam-building beavers), we infer that they have some capacities, analogous to, if yet specifically different from, our capacity to reason and our capacity to act from reason (KU 5:464n). Animals can represent, perceive, and be acquainted with objects through their representations and are capable of subtle differentiations amongst objects.<sup>13</sup> Some animals have more refined external senses than we humans (VM 28:277). In some cases, it seems “the acts of animals arise out of the same *principium* from which human actions spring, and the animal actions are analogues of this” (C 27:459). We have no access through our own introspection, however, to evidence that animals have inner sense, concepts, or cognition, which we encounter in our own case. Kant thought animals do not possess a capacity for language use, which would indicate concepts and higher cognition, much less a first-person pronoun. As for the aforementioned “artistry”

<sup>10</sup> See also VM 28:115–17, 274–77, 448–49, 594, 690; 29:906, 1026.

<sup>11</sup> Kain, “Defense,” 82–87.

<sup>12</sup> Patrick Frierson, “Kant’s Empirical Account of Human Action,” *Philosophers’ Imprint* 5 (7) (2005), 1–34.

<sup>13</sup> LJ 9:64–65; FS 2:59–60; PS 2:285; HN 15:161–62, 713; VM 28:66–67, 78–79, 98–99, 857. For a careful analysis of Kant’s account of the nature and limits of animal psychology, upon which I rely in this paragraph, see Steve Naragon, “Reason and Animals: Descartes, Kant, and Mead on the Place of Humans in Nature,” unpublished doctoral dissertation, University of Notre Dame, 1987, and “Kant on Descartes and the Brutes,” *Kant-Studien* 81 (1990), 1–23. See also Karl Ameriks, *Kant’s Theory of Mind* (Oxford University Press, 1982, 2000), 242.

of beavers, Kant endorsed Bonnet's contention that beavers always build dams according to a single model or plan, an indication that whatever their artistry and the complex form of social cooperation they employ, they lack the ability to reflect upon, modify, and improve their craft or inhibit their instincts.<sup>14</sup> More generally, animals' behavior appears to be guided by rather determinate and pervasive instincts; they are incapable of impulse control and many are easily duped; their behavior does not progress cumulatively over the course of generations. Parsimony counsels not ascribing more sophisticated mental capacities than necessary to explain the phenomena, so Kant concluded that nothing in their behavior required positing full-blown "consciousness," a capacity for "inner sense," for second-order representations, including representations of oneself or one's entire condition: animals lack concepts, judgment, apperception, and self-consciousness, and thus genuine cognition of objects. Unable "to represent to themselves the ground of their movement [*Beweggründe*]," they cannot reflect upon their desires or have "a desire within a desire" (VM 28:99). Unable to conceive of "what is useful or injurious" or "desirable in regard to [their] condition as a whole," they are unable to pursue or experience happiness as such. Perhaps most important for present purposes, absent the capacity to represent what is "unconditionally good," animals must lack the capacity to act upon (or against) the representation of such an unconditional law (KrV A802/B830).

In contrast, we human beings have language, "inner sense," and second-order representations, concepts, apperception, self-consciousness, cognition, and capacities for reflection and inhibition in light of general representations. In his *Anthropology* text, Kant claimed that each of humans' three practical predispositions, the "technical, pragmatic and moral" predispositions, distinguish human beings from all other terrestrial animals. The profound indeterminacy of our instincts and skills, and the connection between our "consciousness" and our technical skill at manipulating things (especially with our hands) itself distinguishes us from all other animals with which we are familiar; our capacity to use other humans in pursuit of happiness and culture and to govern ourselves according to rational principles distinguishes human beings yet further

<sup>14</sup> VPG-Hesse, 122–23; see also VM 28:117. My thanks to Werner Stark for sharing his transcriptions of the notes from Kant's "Lectures on Physical Geography" (VPG = *Vorlesungen der physischen Geographie*), some of which will appear in vol. xxvi of the Akademie Ausgabe (forthcoming), and his invaluable assistance in working with them. I cite passages from these lectures by name (e.g. VPG-Hesse) and the pagination in the original manuscripts. Translations are my own.

(ApH 7:321ff.).<sup>15</sup> This creates an opening, in the human case, for Kant to contend that we have also a capacity for a rational will: to maintain that “the categorical imperative proves for morally practical purposes” that at least some of us “human beings” are free “persons” with the predisposition for freedom under moral laws (MS 6:280n).

It is seldom recognized that, in addition to his interest in distinguishing human beings, and human behavior and mental capacities, from those of non-human animals in general, Kant had a significant interest in animal ethology, comparative morphology, and natural history, as part of a proper “pragmatic” knowledge of the world. Freshly transcribed and edited notes from his lecture course on “Physical Geography” show Kant synthesizing the observations of leading biologists and travelers into characterizations of non-human animals that go beyond the occasional comments in his published works (including the *Physical Geography* text he allowed to be published in 1802).<sup>16</sup>

On the basis of Kant’s comments in the *Anthropology* and the morphological similarities between humans and monkeys (particularly the hand, so emphasized by Linnaeus and Buffon) we might expect Kant to have had particular interest in monkeys.<sup>17</sup> While impressed by their manual dexterity and its deployment for catching mussels, making beds, putting on clothes, and other things, Kant was less than fully impressed, given reports that they steal produce from field and garden and band together to slay lions, tigers, or even humans (PG 9:336–37).

Although the monkeys have an *analogon rationis*, no *analogon moralitatis* will be found in them, as they are always wicked, spiteful and obstinate, and everywhere they go, they wreak havoc.<sup>18</sup>

Wickedness is [the monkey’s] primary attribute; it is never capable of complete trust; with respect to its mental powers, so to speak, the dog and elephant are much to be preferred.<sup>19</sup>

Indeed,

[Dogs] seem to be the most perfect animal, and to manifest most strongly the *analogon rationis* ... they carefully look after their responsibilities, remain with

<sup>15</sup> There is an ambiguous relationship between this description of the practical predispositions and the description found in the *Religion* (animality, humanity, and personality) (RGV 6:26–28).

<sup>16</sup> The course originally included some anthropological topics (as did the metaphysics course); by the mid-1770s, Kant conceived of “anthropology” and “physical geography” as complementary “pragmatic” disciplines which he then taught in alternating semesters (VR 2:443; Br. 10:146). For a brief overview in English, see Steve Naragon’s “Kant in the Classroom” internet resource, available online at [www.manchester.edu/kant/](http://www.manchester.edu/kant/) (accessed April 19, 2010).

<sup>17</sup> ApH 7:322; VPG-Pillau 252, 266. <sup>18</sup> VPG-Kaehler 405; VPG-Messina 248.

<sup>19</sup> VPG-Pillau 266. The comparison of beavers, monkeys, dogs, and elephants seems to have been a common trope; see, for example, Buffon’s discussion in his volume on elephants.

their master; if they've done something wicked they become disturbed; and if they see their master angry, try to win him over with a submissive posture.<sup>20</sup>

While dogs may be Kant's prime example of brutes' necessitation *per stimulus* and the lack of impulse control – “a dog must eat if he is hungry and has something in front of him” (C 27:267) – Kant notes how dogs learn to howl or open a gate-latch, and how with practice they can learn a rabbit's tricks and outwit a rabbit. Their instinct, by repetition of similar cases, “forms an experience which serves the dog as a guiding thread,” despite its lack of concepts.<sup>21</sup>

Kant's greatest sense of wonder, though, is reserved for elephants. “When one observes their strength and their similarity to man, [an elephant] is an animal worthy of admiration [*ein bewunderungswürdiges Thier*].”<sup>22</sup> The elephant's trunk is “the most noble tool,” comparable to a hand in its dexterity and sensitivity, and with a wider range of uses as well; an elephant can use its powers more generally than any other animal.<sup>23</sup> Elephants are very useful, because of their strength and speed on land and in water, and because they are teachable (*gelehrig*) and prudent (*klug*). “Unprovoked, an elephant does no one harm” (cf. PG 9:329). “It is often so gentle that one can break coconuts open on its head, although it must be given some or it will avenge itself with its trunk.” They can be not only tamed, but also “disciplined” (perhaps the only animal that is capable of discipline). Kant notes that people in Surinam use an elephant in place of a servant, a role which it carries out well and patiently.<sup>24</sup> In one set of notes, Kant is reported to have concluded his comments on elephants thus:

An elephant is a gentle animal, and seems to be an *Analogon* of Morality. It understands jokes, but cannot be duped.<sup>25</sup>

Unfortunately, neither the precise basis of such remarks, nor their implications, is further elaborated. Clearly reports about elephant behavior (or at least the parts that he found credible or worth collecting and remarking upon) made an impression upon Kant. Rather than emphasize differences between or the distance between elephants and humans, Kant attributes

<sup>20</sup> VPG-Kaehler 401–02. On faithfulness to their master, see also VPG-Hesse 117; MS 6:443; C 27:459.

<sup>21</sup> VM 29:949; 28:116; VA 25:1196.

<sup>22</sup> VPG-Pillau 252. The most detailed discussion is found in the parallel Pillau and Barth notes. These are the primary source for the rest of this paragraph, unless otherwise noted.

<sup>23</sup> VPG-Kaehler 397; cf. PG 9:328. <sup>24</sup> VPG-Kaehler 396; cf. VPG-Messina 238.

<sup>25</sup> VPG-Pillau 253. This is an important contrast with most other animals, which Kant thinks are easily deceived (VM 28:116).



significant mental sophistication to elephants and uses words with significant positive ethical overtones (prudence, good-natured, patience, discipline) without reservation.<sup>26</sup>

This survey of Kant's systematic, "naturalistic," and empirical biology, empirical psychology, and pragmatic anthropology and "physical geography" establishes that Kant had a serious account of the nature of animals. While many of the details and assumptions of Kant's account have been superseded by subsequent scientific and philosophical developments, it is not clear that his primary conclusions have been.<sup>27</sup> Kant concluded, as a contingent empirical matter, that human beings have *rational* souls, while no other animals with which we are familiar do.<sup>28</sup> This is the account to which his moral philosophy makes reference.

#### 4 MORAL IMPLICATIONS

##### 4.1 Kant's defense of human moral status

As we have seen, on Kant's account, moral status requires the possession of "freedom under moral laws" by something "given as an object of experience." Kant claims that, in our "immediate consciousness" of the moral law "the categorical imperative proves for morally practical purposes" that at least some of us "human beings" are "free" (KpV 5:29; MS 6:280n). Of course human infants and the severely disabled fail to manifest in their behavior much complex consciousness at all, much less an immediate consciousness of the moral law.<sup>29</sup> I have argued elsewhere that Kant has a principled basis for his ascription of moral status, even to humans in so-called "marginal cases."<sup>30</sup> First, Kant's analysis of freedom contends that freedom must be an original and essential predisposition of any being that can possess it. Kant insisted that it must be possible for finite beings endowed with freedom to come into being, since "the

<sup>26</sup> Apart from distinguishing discipline from mere learning, Kant does not elaborate. He appears to accept the myth that elephants do not mate in captivity, but does not mention Buffon's interpretation of this as a form of modesty or self-control (VPG-Pillau 235). Nor does he elaborate an interpretation of an elephant's desire to avenge itself or resist being duped.

<sup>27</sup> Interestingly, Korsgaard seems independently to arrive at some similar conclusions in "Interacting."

<sup>28</sup> See note 10, above, and Kain, "Defense," 82n70. For additional historical context, see Hans Werner Ingensiep, "Tierseele und tierethische Argumentationen in der Deutschen philosophischen Literatur des 18. Jahrhunderts," *NTM: Internationale Zeitschrift für Geschichte und Ethik der Naturwissenschaften, Technik und Medizin* (N.S.) 4 (1996), 103–18.

<sup>29</sup> It is sometimes supposed that Kant's claims about "personality" suggest a "Lockean" approach to personal identity and moral status, but this is dubious. Kain, "Defense," 65n16.

<sup>30</sup> *Ibid.*, 90–100.

categorical imperative proves for morally practical purposes” that at least some of us “human beings” are free, but he argues that it is logically impossible for free beings to be the product of any physical operation. Kant suggested that the most appropriate way to think about the origins of a human being is that rational souls are created endowed with freedom and that these souls are embodied or “brought over into the world” by human parents when they generate and ensoul a human organism (MS 6:280).

Second, Kant’s thesis about freedom as a predisposition, taken in conjunction with Kant’s biological, psychological, and anthropological commitments, provides support for his judgment that every human being possesses it. Kant’s biological theory maintains that each organism can be considered the presentation of a being with predispositions, and we must consider them to be such presentations “from procreation” or conception. Kant’s psychological theory maintains that each animal gets its soul at the point of its reproductive origin. The practical doctrine of original freedom entails that free rational souls must be *essentially* free rational souls, which implies that moral status attaches as soon as an organism endowed with such a soul is generated or conceived. The patterns of pragmatic and moral development across human populations strongly suggested to Kant that the predisposition to personality should be considered a predisposition of the human species, as opposed to a predisposition of only *some* of its members. “The human procreative faculty is the faculty of a human being, with a human of the other sex, to put a person in the world” (HN 23:357). Kant’s commitments provide a principled, if debatable, basis for his judgment that all human beings, even the apparently “marginal cases,” are intrinsically worthy of respect and each is capable of directly obligating us. Kant’s substantive judgments about human marginal cases may not require the rejection or radical revision of his account of moral status.

#### 4.2 *Kant’s rejection of duties “to” animals*

In this context, the question becomes whether careful attention to the nature and behavior of any non-human animals provides evidence that it, and by extension the other members of its species, possess the predisposition to personality. Kant’s conclusion was that it does not. Indeed, his judgment was that there was insufficient evidence to even ascribe to non-human animals many of the predispositions and capacities which are necessary components of the predisposition to personality: they lack the

capacity for concepts, self-consciousness, judgment, and so forth. While it is not clear precisely why he interpreted the behavior of monkeys, dogs, and elephants as he did, absent the manifestation by some of those animals of rather full-blown “Kantian” moral consciousness, or at least the manifestation that such consciousness was developing, this judgment is hardly arbitrary. Animals are “endowed with sensation and choice” yet are “non-rational,” they are incapable of rational cognition and, most importantly, they lack a free rational will (MS 6:442–43). Love, fear, admiration, and amazement are proper for a variety of objects, especially for animals, but the “proper object of respect” is the moral law and those beings with dignity, ourselves and other human beings, with the capacity to “legislate” the law and to hold it before us (KpV 5:76–78; MS 6:443; G 4:435–36, 440).<sup>31</sup> This is why we cannot have any duties *to* animals.

Perhaps what strikes many readers as fundamentally objectionable about Kant’s denial of duties to animals is the apparent implication that they are completely devoid of moral significance, mere “things” at best only accidentally distinguishable from any arbitrary hunk of matter. But before jumping to such a conclusion, careful attention must be paid to the details of Kant’s positive account of the place of animals in the moral life.

#### 4.3 Kant’s account of duties “regarding” animals

In the *Metaphysics of Morals* and in notes from his “Lectures on Ethics,” Kant identifies a general duty to oneself to refrain from unjustified “violent and cruel treatment of animals,” as well as a number of more particular moral requirements regarding our behavior towards certain animals. After laying out Kant’s core argument for this general duty and considering the basis for some of the particular duties he mentions, we will examine some important and illuminating objections to it.

Kant’s contention in the *Metaphysics of Morals* is that the fundamental moral problem with “violent and cruel treatment of animals” is its rather “intimate opposition” to “a human being’s duty to himself” (MS 6:443). As Baranzke has recently emphasized, Kant’s discussion of duties regarding animals comes at the conclusion of his discussion of *perfect duties to oneself*, before he proceeds to his detailed examination of imperfect duties to oneself or any duties to others.<sup>32</sup> In the *Metaphysics of Morals*, duties to

<sup>31</sup> See also, KU 5:372, 482n; C 27:459; V 27:709–10.

<sup>32</sup> Baranzke, “Tierethik.”

oneself are tied to the ethical requirement to have “one’s own perfection” as an end (MS 6:385–87). The perfect or limiting or “negative duties [to oneself] *forbid* a human being to act contrary to the **end** of his nature and so have to do merely with his moral self-preservation.” In contrast, positive, widening, imperfect duties to oneself “command him to make a certain object of choice his end, concern his *perfecting* of himself ... they belong to his *cultivation* (active perfecting) of himself” (MS 6:419). Suicide, for example, is contrary to one’s perfect duty to oneself because of the way it conflicts with the agent’s natural inclination to self-preservation; it involves “renouncing his personality” and “debasement of humanity in [his] person” (MS 6:420, 422–23). Because of what the agent expresses about his nature when he violates a perfect duty to himself, such actions are particularly dishonorable.

Kant contends, most fundamentally, that the “violent and cruel treatment of animals” violates a perfect duty to oneself. As Denis has explained, Kant insists that “the ways that we treat animals reflect and affect morally important attitudes and feelings.”<sup>33</sup> This approach emphasizes two points: one about the moral significance of certain of our feelings, the other about the nature of animals and how, given that nature and our own, animals properly engage these feelings. First, “certain emotional predispositions are extremely useful natural tools for us as moral beings,” useful both motivationally and epistemically, and they “may also reflect certain moral commitments” insofar as they “can be shaped” by our choices.<sup>34</sup> In particular, Kant singles out the “disposition of sensibility ... to love something ... even apart from any intention to use it” and, especially, the “natural predisposition” to the “shared feeling of [others]’ suffering,” as feelings that may “promote morality or at least prepare the way for it” and are “very serviceable to morality in one’s relations with other people” (MS 6:443).

We might go even further, once we note that Kant recognizes some “feelings,” namely “moral feeling, conscience, love of one’s neighbor [*die Liebe des Nächsten, Menschenliebe*], and respect for oneself (self-esteem),” as “moral endowments” that “lie at the basis of morality, as subjective conditions of receptiveness to the concept of duty” (MS 6:399). Although even these are not feelings one could have a duty to have (both because their presence is necessary for beings like us and a precondition of duty

<sup>33</sup> Lara Denis, “Kant’s Conception of Duties Regarding Animals: Reconstruction and Reconsideration,” *History of Philosophy Quarterly* 17 (2000), 405–23, 417.

<sup>34</sup> *Ibid.*, 406–7.

itself, and because they may not be produced or increased, either *ex nihilo* or simply at will); nonetheless, these are feelings that ought to be cultivated, and more importantly in the present context, they ought not be degraded, demeaned, or devalued. The two feelings to which Kant directly appeals in his discussion of duties regarding animals and inanimate nature, love and sympathy, are “intimately” connected with the feelings on this list. The general capacity for love as “delight,” (*Liebe des Wohlgefallens, amor complacentiae*) “pleasure joined immediately to the representation of an object’s existence,” is discussed as part of Kant’s treatment of the “moral endowment” of *Menschenliebe*, the latter being either a special instance or a particular development of the former (MS 6:402, 449, 450). Although sympathy does not itself appear explicitly on the list of subjective preconditions of duty, it seems to have a similar status. “Sympathetic joy and sadness (*sympathia moralis*) are sensible feelings of pleasure or displeasure . . . at another’s state of joy or pain” (MS 6:456). Humans, Kant claims, have a natural receptivity to such shared feeling, often called “humanity” or “humaneness” (*Menschlichkeit, humanitas aesthetica*), which is a precondition for the willingness to share in others’ feelings. “While it is not in itself a duty to share the sufferings (as well as the joys) of others, it is a duty to sympathize actively in their fate; and to this end it is therefore an indirect duty to cultivate the compassionate natural (aesthetic) feelings in us” (MS 6:457). At least for “animals endowed with reason,” such as ourselves, sympathetic feeling is a necessary precondition for moral obligation. In other words, these feelings of love and sympathy are not simply morally useful, as merely one means among others, or merely useful because of some highly contingent facts about human psychology; they may be “an essential part of the fulfillment of duty itself,” at least for beings anything much like us.<sup>35</sup> At least, they are intimately connected to such essential feelings. They are feelings that we have a perfect duty to ourselves to preserve and neither denigrate nor demean, in addition to being feelings that we have an imperfect duty to ourselves to cultivate.

The second crucial point in Kant’s case for this perfect duty to ourselves regarding animal cruelty is that, on Kant’s account of the nature of animals, animals by their nature properly engage our morally significant feelings. An animal is not only a beautiful and teleologically organized

<sup>35</sup> Paul Guyer, *Kant and the Experience of Freedom* (Cambridge University Press, 1993), 390. (Guyer does not endorse this specific analysis, or the point to which I am putting it.)

creature, but also a creature that can feel pleasure and pain, represent the world and have desires (including desires conducive to its self-preservation, reproduction, and enjoyment), and act upon those desires and “principles” analogous to ours. Such a creature is a proper object of our love and sympathy in ways that plants, machines, and crystal formations are not. It is “because of these analogies” between human and animal nature that “the ways that we treat animals reflect and affect morally important attitudes and feelings.”<sup>36</sup> “Many of our morally important sentiments do not discriminate between animals and humans,” and this is no accident or psychological quirk.<sup>37</sup> It is love and/or sympathy which we feel, or at least have a predisposition to feel, towards animals as well as human beings, and in many cases, such feelings may be based upon the presence of some of the same, or closely analogous, features present in animals and humans. Choices to deny, avoid, trivialize, or cavalierly violate such bonds of love or sympathy (or predispositions to them) express disrespect for ourselves. In general, the violent or cruel treatment of animals (at least when unjustified), is incompatible with respect for ourselves because it essentially involves the disregard, denial, or demeaning of these predispositions, feelings, and bonds which are integral to our own nature as moral animals. Animals ought not to be harmed or destroyed “without reason” (C 27:459).

In the case of certain kinds of animals and particular individual animals, Kant suggests a few additional conclusions. An animal’s specific capacities, not just for experiencing pain but for excessive strain, or for loyalty, may come into play, as may its individual history. The kinds of work to which an animal or kind of animal may permissibly be put should accord with their capacities; they “should not be strained beyond their capacities” (MS 6:443). Horses and dogs may provide service over many years, and dogs in particular may do so with particular loyalty and attachment to their master, as we have seen. Having done so, they must be rewarded with gratitude, “just as if they were members of the household”; “once the dog can serve no longer, [we] must look after him to the end” rather than “turn him out,” starve him, or have him shot. Failure to do so reveals “a very small mind,” and is contrary to one’s humane or sympathetic feelings (MS 6:443; C 27:459; V 27:710). A dog’s capacities for particular kinds of feelings, desires, and attachments make it the proper object of greater love and sympathy than is appropriate to feel for a grub, and one’s own dog’s particular devotion makes it especially apt

<sup>36</sup> Denis, “Kant’s Conception,” 417.    <sup>37</sup> *Ibid.*, 407.

for a significant measure of one's love, sympathy, and gratitude.<sup>38</sup> One can see how Kant's analysis would entail similar, indeed stricter, requirements for the treatment of elephants, given his understanding of their nature, especially their "analogy of morality."

Of course, it is not that feelings of love or sympathy for animals, all by themselves, provide a rule for action. No feeling, not even "moral feeling" itself, plays such a role in Kant's theory, and feelings of sympathetic love, even when directed at other humans, are neither an infallible guide to other's needs nor by themselves a rule for action (MS 6:400; G 4:398). Moreover, Kant explicitly allows the killing of some animals "quickly (without pain)" and even some "agonizing physical experiments" for important ends, though not for sport or pure speculation (MS 6:443; C 27:460). As with other perfect duties, what needs to be determined in each domain is which courses of action, or, better, which maxims of action, are incompatible with respect for one's rational nature, in this case, incompatible with one's moral self-preservation. Just as the permissible assumption of some risks to life and bodily integrity is compatible with the prohibition on suicide (and with proper regard for the inclination to self-preservation), so may some use, some killing, even some cruel treatment of animals for important human ends, be permissible or even required.<sup>39</sup> In the *Metaphysics of Morals* Kant intends to outline some first principles that provide a basic framework for such deliberations and determinations, rather than to provide an algorithm or exhaustive treatment of examples. In this case, Kant's principles may raise significant questions about a wide range of human conduct, from animal research, to our eating and farming practices, to some of our leisure activities; not just any human interest may justify the killing of or cruelty to an animal.<sup>40</sup> There are both general protections for all sentient creatures and various particular requirements regarding specific kinds of animals and specific kinds of human-animal relationships, requirements which depend significantly upon the nature of the animals in question.

It is important to note how this core argument differs from the argument often attributed to Kant. It is often thought that Kant's only

<sup>38</sup> While this is a duty to oneself, the duty requires *gratitude to the dog*, contra Timmermann, "Tail," 132; or, if gratitude proper entails respect, some analogue of gratitude to the dog (MS 6:454). Kant does condone killing dogs if they become rabid, however (VPG-Hesse 117).

<sup>39</sup> Denis, "Kant's Conception," esp. 413–14. See MS 6:422–24, 434–37.

<sup>40</sup> For a sketch of some such arguments, see Denis, "Kant's Conception." Without endorsing all of her conclusions, one can see how this approach might address a remarkably wide range of ethical questions. See also Lara Denis, "Animality and Agency: A Kantian Approach to Abortion," *Philosophy and Phenomenological Research* 76 (2008), 117–37.

objection to animal cruelty focuses on the putative psychological *effects* of violence and cruelty toward animals on the human agent that perpetrates it and, especially, the effects on other humans that the agent may subsequently encounter and be more likely to mistreat. While Kant cannot resist endorsing plausible empirical hypotheses about the long-term effects of animal cruelty, this should not be confused with the particularly “intimate opposition” of such cruelty to one’s duties to self that Kant intends to highlight. The more familiar “brutalization argument” is vulnerable to the familiar objection that a single act of gratuitous cruelty may fail to have a discernable long-term impact, and to the objection that the contingencies of human psychology upon which brutalization claims depend are not deep enough to properly secure significant prohibitions on animal cruelty. But Kant’s core argument evades both of these complaints: his focus is upon the immediate disregard for one’s morally significant feelings that is integral to the mistreatment of animals, even in isolated instances, and this is independent of many psychological contingencies.

One objection to Kant’s account focuses on Kant’s characterizations of our duties regarding animals as “indirect” duties (MS 6:443). Timmermann has recently argued that, within Kant’s theory, the identification of something as an indirect duty reveals that it is “really no duty, nor part of a duty, but a mere accidental means to fulfilling a duty.” Thus, in the case of indirect duties regarding animals, “there is nothing about the animal that makes treating them decently morally good. Treating animals decently is a mere means to taking care of your own moral well-being.” “There would be no duty to do it if neglect did not lead to adverse *effects* on our moral capacities.”<sup>41</sup> Surely, it is alleged, this fails to do justice to ordinary moral intuitions about the mistreatment of animals. This objection falters on several counts. First, in the present case, it misunderstands Kant’s position: Kant’s emphasis is upon what mistreatment of animals *expresses* about one’s feelings and moral perfection, rather than on the effects of mistreatment, for oneself or another, or on the ineffectiveness of mistreatment as a means to one’s obligatory ends. On Kant’s account, the relevant feelings are also much more than accidental or incidental means to moral compliance. Second, there is indeed something about the animals in question that grounds Kant’s demands to treat them decently: because of their nature or behavior, animals are

<sup>41</sup> Timmermann, “Tail,” 140, 143n10, 144n11. See also Jens Timmermann, “Kant on Conscience, ‘Indirect’ Duty, and Moral Error,” *International Philosophical Quarterly* 46 (2006), 293–308.



the proper object of one's sympathy and love. Again, proper treatment of animals is a necessary condition for and perhaps a constitutive part of one's moral well-being, rather than a mere "instrumental" means to it. Thus, regardless of whether other cases may fit Timmermann's general characterization of "indirect" duties, Kant's account of our duties regarding animals does not manifest its objectionable aspects.

A second objection points to another kind of apparent "indirectness" in Kant's account. Focusing, as Kant's account does, on the human agent and her own self-regarding psychological states allegedly marginalizes, distorts, or attenuates the proper consideration of the animals' nature or proper concern for the animals and their well-being. By focusing on the agent's self-respect, the Kantian account seems to foreground the agent's self-concern (if only for her own integrity or "self-righteousness") and background her concern for the animals.<sup>42</sup> Such an orientation, it is objected, is both psychologically peculiar and ethically deficient. However, this objection may involve confusion, at least as it is applied to the account outlined above. Indeed, part of what Kant insists upon is the fact that a self-respecting person is directly concerned with the fate of animals: he regards animals as proper direct objects of love and sympathy and he acts in ways that preserve his own disposition to such love and sympathy. To be sure, Kant will insist that one's love and sympathy for animals (as with such feelings for other humans) should, in action, be regulated by reason. But, it is not clear that self-respect plays a larger psychological role in the case of duties regarding animals than it does in the case of duties to other humans; rather, it is simply that there is no need to appeal, in the present case, to the agent's *respect* for anyone other than the agent. Put another way, if the Kantian account of an agent's self-respect leaves sufficient psychological room for genuine respect, love, and sympathy for other people when we discharge our duties to them (and manifest love or sympathy for them), then there may be no special problem about having direct love or sympathy for animals when we discharge our duties regarding them.

Understood in its proper context, Kant's insistence upon duties to all human beings and duties regarding animals is reasonably well grounded and responsive to many familiar objections. Of course, some, including some Kantians, may still insist that animals are due greater regard than Kant allows.

<sup>42</sup> Wood, "Duties," 194. Part of Wood's endorsement of this objection may depend upon his acceptance of Guyer's claims that that the duties, on Kant's account, must be only imperfect, rather than perfect, duties to oneself (210n18).

## 5 TWO KANTIAN ALTERNATIVES

Two distinguished Kantian ethicists, Allen Wood and Christine Korsgaard, have recently proposed modifications of Kant's account, designed to accord animals greater significance in Kantian ethics. It may be instructive to consider these alternatives and a few of the ways they compare with Kant's position as described above.

Wood rejects Kant's claim that all duties must be duties *to some person* or duties to respect rational nature "in the person of some being who has it" (and thus to respect "persons themselves"); Wood contends, "we should *also* respect rational nature *in the abstract*, which entails respecting fragments of it or necessary conditions of it, even where these are not found in fully rational beings or persons."<sup>43</sup> Some of the features of animals (e.g. their capacity for suffering, or for desire, or for caring) constitute "substructures, fragments, and analogues of rational nature"; they are of the sort to be large and rather immediate components of rational nature, at least when possessed by beings with a rational nature. Because of this special relationship these features bear to rational nature, each instance of such a feature deserves respect in its own right.<sup>44</sup> On this account, Kantian duties regarding animals are established without any need for special reference to the agent's own self-respect (or for his respect for other human beings) and respect is not limited to persons, even while all value is still determined in relation to rational nature.

One point of concern about Wood's account is that it remains unclear precisely what "respect for rational nature *in the abstract*" is supposed to denote.<sup>45</sup> More importantly, it is unclear why the relationship that substructures, fragments, and analogues of rational nature allegedly bear to "rational nature in the abstract" entails that they are worthy of the genuine respect rational nature is. Indeed, because those features are at best only analogues of a rational nature, or only tokens of a type of feature that may be a component of an individual with a rational nature, one might suppose that what those features are worthy of is an analogue of respect, or merely a token attitude of a type that may be a component of respect, rather than that those

<sup>43</sup> Wood, *Kant's Ethical Thought*, 143; "Duties," 193, 198.

<sup>44</sup> Wood, *Kantian Ethics*, 100–03; "Duties," 197. Occasionally Wood seems to suggest that what deserves respect in these cases are *the animals themselves*, rather than the features. Perhaps this is a further inference – they deserve respect because they are bearers of features that deserve respect.

<sup>45</sup> If it simply denotes respect for the moral law, considered as an abstract object or principle, that is fine, but, until it is independently determined what the moral law demands regarding animals

features deserve full-blown respect. If this is correct, one might note that the resulting view is not far from Kant's own view, as long as the love and sympathy of which he claims they are proper objects are sufficiently plausible "analogues of respect." After all, such love and sympathy are direct forms of concern for the animals in virtue of their analogous characteristics; and they do generate significant constraints on our behavior toward the animals.<sup>46</sup>

Korsgaard argues that Kantians should recognize all animals (and perhaps all functionally organized objects) as "the source of legitimate normative claims ... that must be recognized by all rational agents"; animals and their interests "have a direct normative claim" on us, and it is their protection that the moral law demands in a fundamental or ultimate way.<sup>47</sup> On Korsgaard's "constructivist" interpretation of Kant, all norms are constructed by and all value is conferred by our acts of legislative volition.<sup>48</sup> In pursuing my interests, I claim that my interests and my "natural good" are worthy of pursuit by any rational being and I claim that I possess absolute worth, worthy of respect by any rational being; my "legislative volition" confers value upon myself and upon my interests and constructs universal norms for my protection and the promotion of my interests. Of course beings such as animals or human infants that are incapable of or simply fail to exercise legislative volition ipso facto do not construct any norms or confer value on anything. Yet, Korsgaard explains, this need not preclude *someone else* from constructing norms for their protection or conferring value, even fundamental value, upon them. Indeed, if it is my "animal nature, not just [my] autonomous nature, that [I] take to be an end-in-itself" or what is of fundamental value, and if it is on my "natural good" as an animal that I "confer normative value" when I value myself as an end-in-itself, then my acts of "legislative volition" (which must be universal in scope) commit me to endorsing the fundamental normative significance of all other humans and animals and of their interests.<sup>49</sup> So the moral law demands respect

and their features, it cannot carry much weight in Wood's argument. If it denotes some kind of respect for the human species and its historical vocation, it may again be unobjectionable, but would, in any event, require a detour much like the one rejected in eschewing appeal to the agent's own self-respect.

<sup>46</sup> This may be all that Wood's position is intended to capture, since he resists ascribing to animals any moral status equivalent to that of human beings, even the most marginal human beings or what he calls "persons in the extended sense." *Kantian Ethics*, 97, 101.

<sup>47</sup> Korsgaard, "Fellow Creatures," 95. (See also Korsgaard, "Interacting" and *Sources*.)

<sup>48</sup> Korsgaard, "Fellow Creatures," 95, 101.

<sup>49</sup> *Ibid.*, 104. In "Interacting" Korsgaard distinguishes "weaker" and "stronger" versions of this argument. As far as I can tell, to reach the conclusion that morality requires *respect for animals* themselves, the stronger version is necessary.

for animals and protection for them and their interests, and it does so for their sake.

For present purposes, we should focus on Korsgaard's case for identifying *animal selves* and an animal's natural good as the objects of fundamental normative significance (and proper direct objects of respect).<sup>50</sup> First, it may be important to distinguish between the normative significance of a particular being or self, on the one hand, and the normative significance of that being's interests, on the other. Even in the straightforwardly human case, it seems important for Kantians to distinguish between respect for a person and the concern for her interests, or even her happiness or well-being as a whole, that is rooted in that respect.<sup>51</sup> Second, it is important to recall that Kant's own account already requires serious concern for the interests of animals, tied to our love and sympathy; what Kant does not allow is concern in the form of *respect* for the animal itself.<sup>52</sup> In support of the claim that it is animal nature per se upon which we confer absolute value, Korsgaard adopts a thought experiment: "imagine that [you are] about to be deprived of [your] rational nature, but may now settle the question whether [you] will afterward be tortured or not. Can [you] really say: 'In that case it won't matter?'"<sup>53</sup> However, even if there is agreement that it would "matter," this thought experiment does not isolate the precise reason for this concern, whether the reason is the same as in the ordinary case; but this is what is needed to distinguish the accounts of Korsgaard and Kant here. Does torturing "me-sans-my-rational-nature" matter because it is disrespectful to the victim, or because it is painful, or because it is destructive? If it is simply my love or sympathy that is, or should be, engaged in such a case, then I may not conclude with Korsgaard that it must be "my animal self" upon which I confer absolute value, or that, by extension, all animals must be *respected*.<sup>54</sup>

<sup>50</sup> For a discussion of Kant's alleged moral constructivism, see Patrick Kain, "Self-Legislation in Kant's Moral Philosophy," *Archiv für Geschichte der Philosophie* 86 (2004), 257–306.

<sup>51</sup> As Korsgaard herself might say, the former has intrinsic or unconditional value while the latter has conditional yet objective value. *Creating the Kingdom of Ends* (Cambridge University Press, 1996), 260–62.

<sup>52</sup> This is a further reason why Korsgaard's "stronger" argument may be required.

<sup>53</sup> For a similar argument, see Timmermann, "Tail," 135. I doubt it is metaphysically possible to be deprived of one's nature, or a part of one's nature, while continuing to exist; it remains unclear whether there is a coherent reformulation of the point appropriate for the task, but I won't press this issue here.

<sup>54</sup> Other significant questions about Korsgaard's account concern the precise nature of this respect for animals – whether it is substantially the same as that for humans – and whether respect, perhaps equal respect, is also required for plants and machines.

Consideration of these two alternatives puts us in a position to consider two final points, one critical of Kant, the other complimentary. First, the criticism. Each of these alternatives recognizes the need for a Kantian theory of value that goes, in some respects, beyond what has been found in Kant. Wood argues for the value of substructures, fragments, and analogues of rational nature found in animals. Korsgaard appeals to an Aristotelian account of the final ends (or natural good) of animals and insists upon the centrality of our animal nature to our practical identity.<sup>55</sup> I have suggested that Kant points in a slightly different direction, in the case at hand, namely to claims about animals as the proper objects of human beings' love and sympathy. But here, too, Kant hardly provided an exhaustive account of the distinctive ways in which animals properly engage these feelings, and what he did suggest needs much more philosophical attention than it has received. There are questions about the justification of this account within Kant's system and its philosophical adequacy for the tasks at hand. What a comparison of these alternative highlights is that Kant left many unresolved questions about the nature of non-moral value, its various species, and their precise relations to dignity, and that this presents a challenge for Kantian accounts of duties regarding animals.

On the positive side, Kant's warning about an "amphiboly in moral concepts of reflection" may contain more insight than is generally appreciated. In the *Critique of Pure Reason*, Kant identified an amphiboly, alleging that Leibnizian metaphysical principles mistakenly result from a failure to distinguish properly between two different sources of representations (namely sensibility and the understanding) (KrV A260–92/B316–49). In moral philosophy, Kant suggests that a similar confusion amongst sources of cognition is involved when we mistake a "duty *with regard to*" animals "for a duty *to* those beings" (MS 6:442). Carelessness with the rational concept of obligation (which only allows the thought of obligation to persons), combined with a failure to distinguish properly amongst our feelings, generates confusion. Our feelings of love and sympathy do help us to "recognize ... something improper" in the mistreatment of animals, but when these feelings are not carefully distinguished from that of respect, we mistakenly "represent to ourselves" (*sich vorstellen*) that we have duties to animals, even though this is contrary to what "reason alone can judge": such a relation cannot even be "thought"

<sup>55</sup> If I am correct about Kant's biology and psychology, the former point may be less foreign to Kant than Korsgaard may realize.

(*gedenken*) with such beings, since they lack the relevant predisposition or capacity (V 27:710; MS 6:442).<sup>56</sup> Nonetheless, the strictness of duty and the immediacy of loving and sympathetic concern for an animal make it feel as if we have a direct duty to the animal. Regardless of Kant's particular judgment about animals, given Kant's moral philosophy and moral psychology, some amphiboly should be expected in moral philosophy. Moral obligations will have some implications for the treatment of beings lacking moral status, and it is only natural that we might misinterpret our feelings in such cases as indicative of duties to such beings. This is why Kantians, such as Kant, Wood, and Korsgaard, might be seen as arguing amongst themselves, at least in part, about *where* the amphiboly occurs. If Kant is right, those not privy to his theory of obligation and his moral psychology may be especially vulnerable to the amphiboly, since it is hard to identify it at all without these philosophical resources. As it turns out, many complaints about Kant's account of duties regarding animals, especially those coming from theorists who reject his distinction between respect and love or sympathy, may miss the mark because they fail to grasp the amphiboly. While non-Kantians complain on the basis of the amphiboly that Kant trivializes or distorts our duties regarding animals, it is not clear that Kant's theory demands, at a fundamental level, much less regard for non-human animals than many of its rivals do. Indeed, his account of duties regarding animals endorses direct appreciation of and concern for animals and recognizes significant moral requirements on us that are grounded in and can vary with the nature, behavior, and history of the animals. If this were all that is involved in "moral status" or a duty to something, then there might be little difference between Kant and many of his rivals. What Kant does argue for is something more, namely respect, for human beings, and this may be something that many of his rivals cannot accommodate. Whatever the outcome of this dispute, careful attention to the charge of amphiboly enables us to distinguish these issues.

## 6 CONCLUSION

Examined carefully in the light of Kant's corpus, Kant's account of our duties regarding non-human animals is less vulnerable to many familiar criticisms than ordinarily thought. Perhaps Kantian theorists, if they are

<sup>56</sup> Here I slightly modify Gregor's and Heath's translations to better capture technical epistemological features of the amphiboly.

willing to follow Kant's lead and able to integrate contemporary scientific accounts of human and non-human animals into Kantian theory, can defend the foundations of Kantian moral philosophy while both affirming the importance of genuine concern for animals and distinguishing such concern from the respect due to human beings.

*Kant's Tugendlehre as normative ethics*

Thomas E. Hill, Jr.

## I INTRODUCTION

In *The Metaphysics of Morals*, especially the *Tugendlehre* or *Doctrine of Virtue*, Kant clarifies, develops, and extends ideas that he presented in the *Groundwork for the Metaphysics of Morals* and the *Critique of Practical Reason*.<sup>1</sup> These earlier works attempt to articulate and defend the fundamental moral law, and they argue that all previous moral theories fail to appreciate the autonomy of will that we must attribute to ourselves as rational moral agents. They provide only a few brief examples of how this supreme principle can guide moral deliberation, a task that the *Groundwork* explicitly postpones for a later “metaphysics of morals” (G 4:391). Kant published a two-part work of this title late in his life (1797–98), and it is here that we find his fullest official presentation of his normative ethical theory (or here “normative ethics”).

My plan is to review and highlight certain features of Kant's normative ethics as I understand it. My focus will be primarily on general features, especially its aim and structure, rather than on specific first-order duties. The discussion will be wide-ranging, though not comprehensive. The interpretations that I propose may be controversial at points, but I shall not defend them here.<sup>2</sup> My hope is that together they present the *Tugendlehre* as a normative ethics that is coherent and contrasts with other normative ethical theories in interesting ways.

More specifically, the plan is this: first, I review some features of normative ethics that distinguish it from science, metaphysics, metaethics, and theories of law. Second, I discuss the role of the basic moral principles

<sup>1</sup> All references to Kant's works are to the Cambridge Edition, with the exception of those to the *Groundwork for the Metaphysics of Morals*, which are to Immanuel Kant, *Groundwork for the Metaphysics of Morals*, T.E. Hill, Jr. (ed.), A. Zweig (trans.) (Oxford University Press, 2002).

<sup>2</sup> For other perspectives, see Allen Wood's essay, “The Final Form of Kant's Practical Philosophy,” and other essays in Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 1–21 and following.



in Kant's theory and how they relate to more specific principles. Third, I consider Kant's idea of duties to oneself and their relevance to certain contemporary discussions. Fourth, I discuss second-order duties to oneself that anticipate our liability to errors in moral judgment, ulterior motives, and weakness of will. Fifth, I comment briefly on Kant's idea of what should motivate us to fulfill our ethical duties. Finally, I note some ways in which the *Tugendlehre* is incomplete.

## 2 NORMATIVE ETHICS

Normative ethics may be distinguished from science, metaphysics, meta-ethics, and theories of law. Kant has much to say about all of these, but they are not the same. The boundaries are sometimes unclear, but rough distinctions should serve well enough here. Normative ethics traditionally proposes general answers to such questions as "What ought I to do?", "What may I do?", "What ends are ultimately worth pursuing?", and "What are features of a morally good person?" These contrast with questions of sociology and psychology, such as "What do people in various groups *believe* about moral requirements, permissions, worthy ends, good character?" and "What are the causal influences on their moral beliefs and behavior?" In the *Tugendlehre* Kant addresses normative ethical questions as distinct from scientific ones. Scientific inquiry uses "theoretical reason" focused on the world as it is – the objects that exist, events that occur, and their causes. In Kant's view, ethics must use "practical reason," which is focused on the ends, conduct, and character that we *ought* to strive for.<sup>3</sup> Insofar as history and anthropology (in Kant's sense) attempt to describe and explain behavior, they too contrast with normative ethics.<sup>4</sup> Even if history and anthropology never rise to the level of science, they primarily use theoretical reason to investigate what is and has been, rather than what ought to be.

In contemporary philosophy *normative ethics* is also distinct from *metaphysics*. Kant works with a similar distinction, but his term "metaphysics of morals" may mislead readers accustomed to current usage. "Metaphysics" now commonly refers to philosophical inquiry into the most general

<sup>3</sup> Although he never gave a full explanation of "the unity of reason," Kant held that "theoretical reason" and "practical reason" refer to reason used in two different ways, a theoretical use and a practical use (KpV 5:119ff.; KrV A804f./B832f.).

<sup>4</sup> When history and anthropology are strictly empirical, they are uses of theoretical reason; but for certain purposes Kant allows moral concerns to shape the interpretation of historical events and to draw practical lessons from descriptive anthropology.

nature of things that exist (“Being”), including space, time, substance, causation, particularity and generality, and mind and body. Metaphysics is also distinguished from epistemology and empirical science, though the boundaries are sometimes unclear. For Kant “metaphysics” can mean different things, including *unwarranted speculation* about entities beyond our comprehension and an *appropriately modest attempt* to lay out systematically the most general features of things as we can know or justifiably view them. Kant divided the more modest aspirations into *metaphysics of nature* and *metaphysics of morals*. The *Tugendlehre* is Kant’s presentation of the part of his metaphysics of morals that is concerned with character and moral ends, as distinct from law and justice. This metaphysics of morals is not speculative metaphysics that tries to use theoretical reason beyond its limits, even though its use of practical reason borrows some “ideas” from speculative metaphysics for its choice-guiding (practical) purposes. A metaphysics of morals is also distinct from a *metaphysics of nature*, which attempts to use theoretical reason, appropriately limited, to develop a system of general principles about the natural world.<sup>5</sup>

By contrast with both speculative metaphysics and a metaphysics of nature, Kant’s metaphysics of morals relies primarily on practical reason, not theoretical reason. That is, it uses reason to develop a system of principles to explain not the natural world as it *is*, but what we *ought* to do, treat with respect, and set as ends. The system outlines what we can justifiably regard as moral ends and virtuous character but which we can legitimately claim only to be adequate for purposes of rational deliberation and choice.

A metaphysics of morals is also not *metaethics* in the current sense. Metaethics is conceived as investigation of the *meaning* of terms used in ethical judgments (e.g. “good” and “ought”) and the ontological *status* of alleged moral facts and properties (e.g. whether “real” or “unreal”). Sometimes it encompasses moral epistemology, which asks how can we know or justify our moral claims. Although it uses some definitions, metaphysical ideas, and justifying arguments, Kant’s metaphysics of morals is not primarily a work of conceptual analysis, ontology, or epistemology.

Finally, Kant’s normative *ethics*, as presented in the *Tugendlehre*, is distinct in various ways from his theory of Right.<sup>6</sup> The distinction has been

<sup>5</sup> A metaphysics of nature would be a system of rational principles distinct from particular empirical sciences and beyond the principles of causation, substance, etc., discussed in the first *Critique*. It is not surprising that Kant found this aspiration hard to achieve. See KrV Axix–xxii, A832/B860–A851/B879, and G 4:387–89.

<sup>6</sup> Kant’s *Rechtslehre* is about “rightful” relations in private, public, and international relations. The term “Right” is broader in English whereas *Recht* only concerns rights, justice, and enforceable laws.

variously interpreted, but some points are evident.<sup>7</sup> Both are included in *The Metaphysics of Morals*, but only the *Tugendlehre* concerns duties and ends that are *ethical*, in a sense evidently narrower than “morals” in the work’s title (MS 6:218–21, 379–84). The basic ethical duties are to adopt the maxims to respect humanity as an end in itself and, more specifically, to pursue self-perfection and the happiness of others. Unlike juridical duties, ethical duties cannot be externally legislated or coercively enforced. Their source is inner legislation, and full compliance with them requires being motivated by the moral law. They include wide *imperfect* duties that leave some room for choice about how and when to pursue moral ends. By contrast, juridical duties, as *perfect* duties without latitude, prescribe particular actions, and insofar as these are legal requirements they do not require moral motivation.

These differences are important, but we should not overlook the fact that principles of Right limit and supplement duties of virtue. The *Rechtslehre* not only comes before virtue in the order of exposition, but there is a structural priority as well. Principles of law and justice limit what we may do to promote morally good ends. Duties of virtue are to be understood as qualified, “Promote these ends, but *only* by means compatible with Right.” In addition, the *Rechtslehre* adds content to the *Tugendlehre* because, Kant says, it is an indirectly ethical duty to conform to juridical duties (MS 6:219–21). This implies that fully virtuous persons, for moral reasons, will not only avoid taking unjust and illegal means to their ends but will also obey the positive requirements of their legal system, within certain limits.<sup>8</sup> But the structural point remains that Right constrains and adds to the requirements of virtue.

### 3 THE ROLE OF THE FIRST PRINCIPLES

The principles in the *Metaphysics of Morals* have a hierarchical structure in which more basic comprehensive principles articulate reasons for more specific ones. Kant states the supreme moral principle as follows: “[A]ct on a

<sup>7</sup> Among the less evident and still controversial questions is how the first principles of the *Rechtslehre* and the *Tugendlehre* are related to the categorical imperative.

<sup>8</sup> Kant acknowledges that official orders to do something contrary to “inner morality” should not be obeyed (MS 6:371). Examples, presumably, would include orders to bear false witnesses or to commit rape. (See also RGVI 6:154n.) As Lara Denis notes in chapter 9 of this volume, perfect duties to oneself grounded in “the right of humanity” account for some limits to what law and justice can legitimately require. Also, there are arguably some limits to how unjust and contrary to the rule of law an alleged government can be before losing all claim to being a legitimate government as opposed to being a mere rogue regime without any legal or moral authority.

maxim which can also hold as a universal law. – Any maxim that does not so qualify is contrary to morals” (MS 6:226). Each Part, however, starts from a general principle more specific to its subject matter. In Part I, the *Rechtslehre*, this is “The Universal Principle of Right: Any action is *right* if it can coexist with everyone’s freedom in accordance with a universal law, or on its maxim the freedom of choice of each can coexist with everyone’s freedom in accordance with a universal law” (MS 6:230). In Part II, the *Tugendlehre*, the first principle is: “act in accordance with a maxim of *ends* that it can be a universal law for everyone to have” (MS 6:395).<sup>9</sup> Here my concern is just with how these relate to more particular principles and judgments, and my main aim is to relate Kant’s theory to contemporary discussions.

3.1 Details depend on interpretations of the first principles, but several general points regarding their use seem evident. For example, although Kant’s metaphysics of morals is supposed to start from a priori first principles, its development and application must take into account at least general empirical facts about the human condition. Also, arguments for more specific ethical conclusions cannot always be strictly deductive. Kant acknowledges that experience, judgment, and wisdom are required.<sup>10</sup> This is especially true when trying to fulfill the imperfect duties regarding the ends of self-perfection and the happiness of others. Even perfect ethical duties raise “casuistical questions” that have no simple deductive answers.<sup>11</sup> Moreover, arguments to and from its substantive principles may require *interpretation* of normative ideas. Applying basic ethical principles may be sometimes more a matter of finding a reasonable instantiation than of finding a unique solution. This fits with Kant’s understanding of ethics as employing ideas of reason that are neither reducible to empirical concepts nor as precisely determinate as mathematical concepts are.

3.2 In what sense are the first principles prior to more specific ethical principles? This is an especially pressing question if the first principles are not viewed as the axioms of a purely deductive system or as referring to metaphysical (ontological) bedrock.<sup>12</sup> An alternative that can be neutral

<sup>9</sup> Kant adds: “In accordance with this principle a human being is an end for himself as well as for others, and it is not enough that he is not authorized to use either himself or others merely as means (since he could then still be indifferent to them); it is in itself his duty to make man as such his end” (MS 6:395). Whether intended as a paraphrase, an explanation, or application, this articulates the standard that Kant most often appeals to throughout the *Tugendlehre*.

<sup>10</sup> See G 4:389–90; MS 6:217–18, 411–12, 478.

<sup>11</sup> See MS 6:423–24, 426, 428, 431, 433–34, 437.

<sup>12</sup> Some commentators suggest that Kant held, as a metaphysical/ontological thesis, that “humanity” is an ultimate “value” that exists and is the ground of our various ethical duties.

regarding the ontological status of values and duties is to think of the first principles as epistemologically prior to more specific duties. That is, the first principles may be known independently of more specific duties and they are the only basis for inferring more specific duties. Unlike contemporary theories that rely on the method of reflective equilibrium, this interpretation would give no weight to common "intuitions" about what we ought to do in various situations.

Alan Donagan interprets Kant's moral epistemology differently.<sup>13</sup> That is, although Kant presents the results of his rational reflections on morals in the form of a hierarchy of principles, this does not necessarily mean that the force of evidence flows from top to bottom. As in mathematics, Donagan suggests, we may sometimes reasonably have more confidence in some mid-level principles than in the abstract first principles that we later accept as expressions of the normative concerns that they have in common. The order of discovery and grounds for belief may not be all top-down even though the principles are finally presented in an ordered structure analogous to a system of mathematics. Even the neat structured arguments in that discipline, e.g. proofs in geometry, do not necessarily reflect the mathematicians' order of thought or even grounds for belief. This opens the possibility for a more holistic or reflective equilibrium epistemology in ethics. This would fit with Kant's brief remark in the Preface to the *Groundwork* that, although "apparent adequacy" is not "secure proof of correctness," his first principles would be "clarified" by working out the whole system and "strongly confirmed by the adequacy the system would manifest throughout" (G 4:392). It would fit too Kant's use of examples to invoke awareness of "the fact of reason" in the *Critique of Practical Reason* (KpV 5:31, 46–50).

Whatever we conclude about logical and epistemological priority, it is clear that Kant attributed a normative *moral* priority to the moral law. A persistent theme is that we have specific duties *because* the moral law requires them, and we should strive to fulfill them because we respect the moral law. This contrasts with one form of contemporary intuitionism, represented prominently by Jonathan Dancy, which holds that what we ought to do is determined by the complex cluster of particular facts that

<sup>13</sup> See Alan Donagan, "Moral Dilemmas, Genuine and Spurious: A Comparative Anatomy" and "The Relation of Moral Theory to Moral Judgment: A Kantian Review," in J.E. Malpas (ed.), *Philosophical Papers of Alan Donagan* (University of Chicago Press, 1994), vol. II, 153–68 and 194–216. Donagan presents a modified Kantian "metaphysics of morals" in *The Theory of Morality* (University of Chicago Press, 1977).

constitute “reasons” of various kinds in each particular case.<sup>14</sup> General moral principles, in Dancy’s view, are at best only approximate heuristic generalizations and often useless or harmful. Kant’s view of the prior moral authority of the law also contrasts with any holistic theory that counts particular moral intuitions as among the factors that determine the truth or constitute the authority of the moral law. Kant’s view, I think, is not that the ordinary person’s immediate grasp of the wrongness of false witness, and the like, is what makes the moral law true or authoritative.<sup>15</sup> Rather, the strength and immediacy of such judgments make one aware of the fundamental difference between mere self-interest and morality that is expressed by the moral law. If so, Kant’s theory treats the moral law as the fundamental moral authority even if it is not first in our moral awareness and not a decision procedure that entails determinate answers to all moral questions. A better analogy would be an ideal constitution for a moral commonwealth that authoritatively fixes basic values and procedures for legislating principles that represent the general will (or practical reason) of each citizen.

3.3 Kant’s first principles and most general principles regarding ends are categorical imperatives. This means that they are strict duties with which everyone must comply, even though they allow considerable latitude of choice in application. To respect and value the end of humanity is fundamental and not optional. To adopt and pursue the ends of self-perfection and others’ happiness is also rationally and morally required. These are obligatory ends, not merely ends prescribed by *prima facie* duties.<sup>16</sup> We must also make these our ends for the right (moral) reason – respect for the moral law. When, how, and to what extent we should further these ends, however, is somewhat indeterminate. The imperfect duties to develop our natural powers and to promote the happiness of others permit some “playroom” for free choice (MS 6:390). For example, each profession calls for different skills, but choice among responsible professions is normally optional; and, barring emergencies and special obligations, we have some choice about where to spend our charitable effort, time, and money. The end of moral perfection is “narrow and perfect in quality” but “wide and imperfect in degree” (MS 6:446): we must strive to fulfill *all* of our duties from a moral incentive, but *how* exactly to do this and the *extent* to which we can at any given moment is somewhat indeterminate.

<sup>14</sup> Jonathan Dancy, *Ethics without Principles* (Oxford University Press, 2004).

<sup>15</sup> See Kant’s example, KpV 5:30.

<sup>16</sup> W.D. Ross, *The Right and the Good*, Philip Stratton-Lake (ed.) (Oxford: Clarendon Press, 2001), 16–64.

Despite their latitude and indeterminacy in practice, Kant's duties to promote the obligatory ends still differ from Ross' prima facie duties of self-improvement and beneficence.<sup>17</sup> Most obviously Kant's duties imply that it is morally necessary to *adopt these ends* as one's own – merely promoting ends by chance or in ways that are incidental to another purpose does not suffice. Also, as noted, Kant's duties cannot be completely fulfilled without adopting and pursuing the end from a moral motive. Perhaps surprisingly, Ross' famous list of prima facie duties allows *less latitude* for choice than Kant's imperfect duties. This is because Ross' system implies that, unless there is another conflicting prima facie duty (e.g. fidelity, justice, reparation, gratitude, non-injury, or self-improvement) one must promote a small pleasure for someone else even at the sacrifice of one's own much greater pleasure.<sup>18</sup> Both Ross and Kant accept that it is not in general *a duty* to promote one's own pleasure or happiness, but (arguably) the latitude in Kant's imperfect duties allows *permission* to forgo doing minor favors for others, at least sometimes, so that we can pursue our own ends.<sup>19</sup>

3.4 Because the idea of natural teleology dominated much of ancient and medieval ethics, we should consider briefly its role in Kant's *Tugendlehre*. Is it a basic assumption expressed or presupposed by the first principles of his ethical theory? Modern philosophers such as Hobbes, Hume, and Spinoza rejected the traditional natural teleology, and so Kant's repeated appeals to "nature's purposes" in the *Tugendlehre* may seem an unfortunate regression. Nevertheless, the claim that an end is "natural" in Kant's theory cannot be a foundational moral claim. All of Kant's major ethical writings make clear that the first principles are necessary principles of practical reason that require (and admit) no empirical justification. Kant does argue that there are moral and epistemological reasons for the "regulative" principle that we should try to conceive of natural events as ordered towards an ultimate end, and in the *Tugendlehre* he writes as if specific human capacities (e.g. for speech and reproduction) have natural ends.

<sup>17</sup> Ibid., 21–22.

<sup>18</sup> Ross attempts to rectify this bizarre result by asserting, implausibly, that it is a prima facie duty to promote one's own pleasure if one sees it just as someone's pleasure rather than one's own. Ibid., 24–26.

<sup>19</sup> There has been some controversy about the kinds and extent of latitude that Kant's principles allow. See, for example, Marcia Baron and Melissa Seymour Fahmy, "Beneficence and Other Duties of Love in *The Metaphysics of Morals*," in Thomas E. Hill, Jr. (ed.), *The Blackwell Guide to Kant's Ethics*, (Oxford and Malden, Mass.: Wiley-Blackwell, 2009), 211–28; and Thomas E. Hill, Jr., "Meeting Needs and Doing Favors," in *Human Welfare and Moral Worth: Kantian Perspectives* (Oxford University Press, 2002), 201–44.

Underlying his familiar teleological language in presenting his specific duties, however, is the conviction that practical reason endorses the aims and dispositions that are said to be natural ends. Often this is only implicit, but there is ample precedent for this understanding. Even Thomas Aquinas identifies the basic natural ends relevant to the first precept of natural law as natural inclinations *approved by reason*.

In his gloss on the first principle of the *Tugendlehre* (as well as in the *Groundwork*), Kant presents the fundamental moral imperative in teleological language (“the duty to make man as such his end” [MS 6:395]) but it has moral force because it is a necessary requirement of reason, not because of morally neutral empirical facts about human nature. In the *Groundwork*, drawing from “common *rational* knowledge of morality,” Kant famously says that nature’s purpose for human reason is to produce a good will, not happiness (G 4:396). It soon becomes evident, however, that the special value of a good will lies in the fact that its basic principle is always to follow the unconditional requirements of reason (G 4:402). Practical reason unconditionally requires us to strive for a perfectly good will, and ultimately this is why we should think of it as “nature’s purpose” for giving us reason. In Kant’s teleological framework, the reason we *ought* to seek a certain end is not that it is *nature’s purpose* for us, in an independent, morally neutral sense. To the contrary, the deep rationale for thinking of an end as nature’s purpose for us is in large part the prior underlying belief that there are compelling reasons for us as human beings to pursue it.

#### 4 DUTIES TO ONESELF

A striking difference between Kant’s theory and most contemporary ethical theories is that Kant not only includes but gives special priority to duties to oneself. Many people today see morality as only concerned with how we treat others. Harming yourself, for example, is said to be foolish but not morally wrong unless it indirectly harms others. Advocates of “virtue ethics” tend to treat harming or degrading yourself as reflecting defects of character rather than violations of duty. Classical utilitarians, such as Bentham and Sidgwick, hold that one’s ultimate duty is to consider one’s own good as only a tiny part of the aggregate good that determines right and wrong. Special attention to oneself and one’s inner circle may be justified in practice but only from a moral point of view in which one’s own welfare is ultimately no more important than the comparable welfare of each other sentient creature. On this view, self-harm and self-



benefit have some weight on the grand scale that determines moral right and wrong, but not much. In addition, many philosophers hold that, independently of such substantive issues, it is conceptually confused to suppose that one could have a duty *to oneself* in the strict sense in which we have duties *to others*. They argue, for example, that one cannot literally *violate one's own rights* or create a binding obligation by making a promise to oneself.

In Kant's theory, by contrast, ethical duties to oneself are prominent.<sup>20</sup> Suicide, gluttony, drunkenness, lying, servility, and sexual "self-defilement," Kant argues, are contrary to perfect duties to oneself. Self-perfection is an end that one has an imperfect duty to oneself to adopt and pursue. Leaving the details aside, let us compare and contrast Kant's position with the contemporary views just mentioned.

*First*, like many today Kant distinguishes morality from mere prudence. A foolish use of one's money, for example, becomes (indirectly) a moral issue for Kant only when it threatens to throw one into poverty that leads to vice or makes one unable to fulfill one's responsibilities (MS 6:388). Kant condemns certain ways of harming and debasing oneself, however, as contrary to perfect duties to oneself. For example, severely damaging one's rational capacities by abuse of food, drink, and (presumably) other drugs is wrong, regardless of its harm to others; and the same holds for degrading sexual practices and servility (for example, "making oneself a worm" before others) (MS 6:422–28, 434–38). Also, Kant held that deliberate failure to develop one's natural and moral powers is culpable and to neglect them shows "lack of virtue" (MS 6:390).

*Second*, unlike the position associated above with "virtue ethics," Kant's theory does not relegate all acts that merely harm oneself to a realm of vice and virtue that is *distinct from duty*. In fact, for Kant virtue *is* the strength of moral will to fulfill all one's duties despite opposing inclinations.<sup>21</sup> Acts contrary to perfect duties to oneself or based on rejection of the obligatory end of self-improvement are morally wrong, not merely signs of imperfect virtue.

*Third*, Kant is obviously no utilitarian. More specifically, for him the question whether abusing and degrading oneself through drugs, sexual practices, servility, and lying is wrong is not determined by weighing the consequences for oneself and all others. Along with law and justice, these

<sup>20</sup> For an excellent detailed examination, see Lara Denis, *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (New York and London: Garland Publishing, 2001).

<sup>21</sup> See MS 6:380, 383–84, 390, 394, 404–10.

duties are firm constraints on what one may otherwise reasonably do to promote the happiness of anyone.

Finally, Kant insists that we have duties *to* ourselves even though he was aware of a distinction between duties *to* persons and duties merely *towards* persons or things. He affirms duties *to* oneself and *to* other human beings, but he denies that we have duties *to* God, nature, and non-human animals (MS 6:240–41). Instead, he holds that we have duties *towards* (or regarding) nature, animals, and at least the idea of God (MS 6:443–44). It is difficult to be sure that we understand the relevant terms in exactly the same way, but it seems clear that Kant meant to say that we have duties *to* ourselves and *to* others in the same sense. A duty to oneself in this sense, Kant implies, means that one is both the person under obligation and the person who imposes the obligation (or to whom it is owed). He anticipates the current objection that a duty to oneself is conceptually impossible because as the obligating person (to whom it owed) one could always release oneself (as the obligated person) from the duty (MS 6:417–18). If true, Kant says, this objection would undermine all duties, but he finds a “solution” in the different perspectives one takes when conceiving oneself as the one who is obligated and as the one who imposes the obligation.

In several ways Kant’s position is less radical and unappealing than it may seem.

First, note that many common arguments against duties to oneself do not apply to Kant’s conception of them. For example, assuming that all “duties *to*” are correlative with “rights *against*,” some argue that a duty to oneself would entail the absurdity of having a right against oneself. One cannot literally violate one’s own rights or, as the traditional slogan says, no one can do an injustice to himself. Kant himself accepts the point, reserving duties to oneself for ethics (*Tugendlehre*) rather than the domain of rights and justice (*Rechtslehre*) (MS 6:383). Also wide of the mark is the familiar disdainful association of owing something to oneself with doing oneself favors, for example, as a reward that one deserves or promised oneself. Kant’s duties to oneself are concerned instead with not damaging or degrading oneself and with working to improve one’s natural and moral powers. Some also object to Kant’s position because it implies that the last person on earth would still have duties to himself, and they find this counterintuitive. Intuitions about such remote examples are hardly decisive, but in any case Kant’s duties to oneself are primarily about the sort of person one should try to be as a person living in relations with others. Consider, for example, the duties to oneself not to lie or be servile and, more positively, to “make oneself a useful member of the world” and

develop the strength of will to fulfill *all* of one's duties (MS 6:429–31, 434–37, 446). To be sure, these are not derived from duties to others but they are duties to oneself concerned with how to be a rational agent living among others.

Consider next the familiar argument that a duty to oneself would entail the absurdity that one is bound by a duty from which one could release oneself at will. This seems to assume that the model for all duties *to* a person, in a strict sense, is a promise, contract, or other voluntary commitment (MS 6:417–18). The idea of a morally binding “promise to oneself” is indeed suspect, even though morally significant resolutions are sometimes described as such. Kant, however, does not invoke this idea. He seems to appeal instead to an older idea of an involuntary and irrevocable duty to someone by virtue of that person's inherent authority or awesome superiority. The idea apparently is that the ideal human being (with pure practical reason) in us imposes the obligations that bind us as the obligated moral agents. To make sense of this, as Kant suggests, we must conceive of ourselves in both of these roles but from different perspectives (MS 6:418).<sup>22</sup> The obligation is not voluntarily imposed or undertaken, as with promises, but is permanent and essential to being a moral agent. It is, in part, a metaphorical way of representing the thesis that to be under moral obligation *is* to be an imperfect rational agent who cannot but acknowledge the authority of the principles of rational autonomy that, in Kant's view, are expressed by the categorical imperative.

What some contemporary readers may find troubling is that Kant calls the relevant concerns *duties* at all, not that he classifies them as *to oneself*. If so, disagreement may run deep, but a few words of caution are in order. First, Kant's duties are not narrow requirements that are attached to social roles or meant to be enforced by busy-body neighbors. One's own reason, conscience, and self respect are the ultimate motivators for compliance with ethical (as opposed to juridical) duties. Duties are constraints, but they are not imposed arbitrarily by tradition, culture, or any person, human or divine. They are in a sense self-imposed by our own rational recognition of the principles to which we, as rational agents with autonomy of will, necessarily accept as authoritative. All duties potentially limit what we may do to satisfy our inclinations and pursue happiness, but common sense and virtually all ethical theories recognize some such limits.

<sup>22</sup> See also MS 6:379n, 386–87, 392, 395, 445–47.

A final note regarding the relevance of Kant's duties to oneself to contemporary ethics should at least be mentioned though it is too large a theme to develop here. Social contract theories from Hobbes to Rawls attempt to derive social and political obligations by asking what principles and institutions rational self-regarding agents would accept from a common deliberative perspective. Even T.M. Scanlon's theory, more focused on ethics, tries to determine what we owe to others by contractualist thinking based on intuitive "reasons" from different agents' perspectives.<sup>23</sup> In each case what matters is not just the kind of impartiality or reciprocity the contractualist deliberations involve but also the nature of rational self-regard.

Kant acknowledges that morally constrained prudence is rational and that, more generally, rational agents will take the necessary and available means to their subjective ends unless they deliberately abandon or suspend those ends (for example, for moral reasons).<sup>24</sup> Equally important, however, is the rational self-regard that, in Kant's view, underlies the ethical duties to oneself. To extrapolate, Kant thought that, apart from duties to others, finite rational agents will place a high priority on continuing to exercise and preserve rational capacities, on living among other rational agents as equals, and on communicating with others in a straightforward rational manner. They will also be rationally self-concerned to develop useful physical and mental powers, and to live under conditions favorable for maintaining their self-respect and following their consciences. These are not concerns they have primarily for the sake of others even though, as moral agents, they must respect the same rational interests of every other person.

These claims may be challenged in various ways, but their relevance to contemporary contractualist thinking should be obvious. The rational interests that each individual can bring to the table, as it were, for deciding on (further) principles are not merely self-preservation and desire satisfaction (Hobbes), or natural rights and property (Locke), or primary social goods (Rawls), but (arguably) those rational interests that lie behind Kant's duties to oneself. If so, this would partially explain how duties to oneself are fundamental in Kant's ethics.

##### 5 NEGLIGENCE AND SECOND-ORDER DUTIES

Kant's system of ethical principles includes first-order principles regarding actions, such as suicide, lying, and disobedience to law, as well as

<sup>23</sup> T.M. Scanlon, *What We Owe Each Other* (Cambridge, Mass.: Harvard University Press, 1998).

<sup>24</sup> G 4:414–20, and Editor's Introduction, in Kant, *Groundwork*, Hill (ed.), 51–57.

commitment to moral ends and attitudes. The principles primarily address the first-person questions “What ought I to do?” and “What should I especially value and strive for?” These are the questions we raise when we are morally alert, reflective, and ready to make our acts and attitudes conform to our best judgments. It seems, however, that the most pervasive, if not the worst, wrongs and vices are due to moral negligence, self-deception, and weakness. In these cases we do not raise the moral question, see the moral problem, or anticipate how hard it is to live up to our best intentions. In these contexts principles that merely tell us how to assess our intentions seem to offer little help. At a later time we may become aware of our negligence, self-deception, and weakness, and then we may blame and hold ourselves responsible – as perhaps others will. Such retrospective and third-person assessments of blameworthiness and responsibility, however, are at best a secondary concern in Kant’s theory. The first moral concern is not “Who is to blame?” but “What should I do and strive for now?”

On reflection, however, we can and should face the problems of moral negligence, self-deception, and weakness in our thinking about what we should do and strive for *now*. These are common human tendencies that underlie much wrongdoing, and we can know in general that we are liable to them. We may be unaware, or only half-aware, that we are being negligent, self-deceiving, or weak at the moment of decision, but our ever-present liability to these faults is well known. We should expect, then, that a complete system of principles for guiding first-person deliberation will include principles about preparing ourselves to reduce the influence of these propensities insofar as we can.

Kant’s system includes very general principles of this sort. The first stems from “a human being’s duty to himself as his own innate judge” (MS 6:437–40). This concerns conscience, which Kant imagines as an inner forum analogous to a criminal court. In it one must think of oneself as different persons, by turns the accuser, the defender, and the judge. As judge one hears evidence and then condemns or acquits oneself as the accused, with the best outcome (a clear conscience) being relief, not reward. The conscience is also seen as observing, warning, and even threatening us unbidden. We can “distract” or “stun” ourselves to avoid its message temporarily, but it is an innate predisposition that is inescapably bound up with our moral agency. One’s specific duty regarding conscience is to cultivate it and “sharpen one’s attentiveness to the voice of the inner judge” (MS 6:401). The metaphor of courtroom activity, however, also suggests that the duty is to take on seriously the active role of the judge, scrupulously weighing evidence before exonerating the defendant or rightfully

imputing the act as a violation of moral law (MS 6:448–49).<sup>25</sup> Importantly, this self-judging can only compare *our own understanding* of our acts (past, ongoing, or proposed) with *our own best judgments* about what ought to be done in the context (RGV 6:186; MS 6:401).<sup>26</sup> A clear conscience at best signals that we are blameless, not that the acts are objectively right.

Kant's metaphors of conscience may be questioned, but the main forward-looking message is hard to deny. That is, we should be careful to see that what we intend to do can survive our own critical moral scrutiny. This means that, at least for serious matters, we may need to investigate our facts further, examine our assumptions, rethink our moral reasoning, and consult others to challenge our understandings. This presumably is the *duty of due care* that, Kant suggests, the well-intentioned Spanish Inquisitors failed to heed when they burned heretics at the stake (RGV 6:186).

A further second-order principle is “the first command of all duties to oneself,” namely, “*know* (scrutinize, fathom) *yourself* ... in terms of your moral perfection in relation to your duty” (MS 6:441). Kant's concern here is with our deepest predispositions and commitments as well as more specific motives. Self-scrutiny can make us more aware of our respectable predisposition to the good (“the moral law within”), but it can also reveal a deep, actual will that is impure, contemptible, and evil (RGV 6:29–39; MS 6:441). Although it is “difficult to fathom” the depths of our hearts, self-scrutiny may also expose particular ways in which we deceive ourselves, for example, by taking “wishes empty of deeds” as proof of our good will. Deceiving ourselves about our true motives can also distort our moral deliberations and allow us to persist in wrongdoing. We need to expose self-deception not just to achieve purity of motive but to avoid wrongdoing and neglect of positive duties.

Finally, the duty to oneself to increase one's moral perfection includes not only striving for purity in one's dispositions and motives but striving to “fulfill all one's duties” and “attain completely one's moral end with regard to oneself” (MS 6:446). These duties correspond to the biblical commands “be holy” and “be perfect” (MS 6:446).<sup>27</sup> The latter requires

<sup>25</sup> In judging oneself, however, one must think of the judge as another person, even if “merely an ideal person that reason creates for itself” (MS 6:436). Earlier Kant implies that this ideal judge (“conscience”) is our practical reason (RGV 6:40). We do not have a duty to acquire a conscience but must not follow its initial promptings *blindly*.

<sup>26</sup> For a fuller discussion see “Four Conceptions of Conscience,” *Nomos* 40 (1998), 13–52. Reprinted in Thomas E. Hill, Jr., *Human Welfare and Moral Worth: Kantian Perspectives* (Oxford University Press, 2002), 277–309.

<sup>27</sup> See I Peter 1:16; Matthew 5:48; Philippians 4:8.

us to develop *virtue*, which Kant conceives of as an inner fortitude or strength of moral will with which we can overcome obstacles to duty more easily and in a good spirit. Because persons of good will can be weak, virtue is not merely a good will basically committed to duty above self-interest. Virtue is the opposite of weakness of moral will, which we must strive to overcome. This weakness – or lack of virtue – does not excuse wrongdoing, because, in Kant's view, we must presume we can do what we ought to do even if it is hard and unpleasant. We exercise our will by adopting and acting on maxims, so “weakness” and “strength” here cannot be understood literally as forces, physical or metaphysical. The good-willed person who does wrong from weakness of will, then, should apparently be seen as (voluntarily) acting at the moment on a bad maxim that is inconsistent with his or her own higher order good maxim to avoid wrongdoing. If so, the person would be responsible for the act and reveal an imperfect character (though not as bad a one as if the person had acted with a thoroughly bad will).

Kant may have thought instead that as a result of human frailty a person who has not yet developed virtue will sometimes find it impossible (not just difficult) to do her (normally binding) duty on a particular occasion. The important point here, however, is that he clearly thought that it is a duty to strive to develop virtue and that over time we can improve in that regard. This is an important second-order duty – the duty to increase our moral strength of will to fulfill our other duties despite obstacles.

In sum, Kant's system of ethical principles includes at least three second-order duties in recognition that failure to fulfill first-order duties is often due to moral negligence, self-deception, and weakness. These second-order duties require us, *first*, to be alert to the warnings of conscience and respond appropriately; *second*, to scrutinize our motives to expose our tendency to make excuses; and, *third*, to strive to develop virtue conceived as strength of will to do the right thing despite temptations. These are not criteria for assessing blameworthiness but principles for guiding first-person moral decisions about what to do now. Their ability to guide us, however, presupposes that, despite our failings, we are to some extent aware of their message, do not completely refuse to see their implications, and are not always too weak to follow them.

## 6 MORAL MOTIVATION

An important feature of Kant's system of ethical principles is its requirement of a moral incentive for avoiding wrongdoing and adopting the

obligatory ends. This feature is one of several that distinguish principles of Right from ethics as conceived in the *Tugendlehre*, and it also distinguishes Kant's ethics from most ethical theories prominent today. Kant's requirement of a moral incentive, however, raises questions of interpretation as well as objections.

Principles of Right, according to Kant, forbid or demand particular actions, not maxims or motives. Ethical duties, by contrast, require us to adopt maxims regarding ends to respect and promote.<sup>28</sup> This means that ethical duties essentially concern our attitudes, goals, and intentions, contrary to contemporary theories that relegate such concerns to theories of character. These theories reserve *duty* and *obligation*, *right* and *wrong*, for "the acts themselves" as described apart from such variable ("subjective") factors. In addition, Kant held that ethical duty, the product of "inner legislation," prescribes a moral incentive – respect for the moral law. Duty requires not only that we respect and pursue obligatory ends but that we do so for moral reasons. This demand contrasts sharply with mainstream contemporary theories. It is off-putting to many readers and, perhaps for this reason, is often de-emphasized by those who try to present Kant's theory in a favorable light.

How should we understand the *requirement of a moral incentive*? Is it that we must strive to make the thought of duty present and salient every time we do something to promote an obligatory end? Must I, for example, try to live so that each time, or even usually, when I help others, the immediate honest explanation for my doing so would be that duty requires me to promote the happiness of others? Must I strive to make the thought of duty my dominant and explicit reason for visiting a friend in the hospital, taking care of an aging parent, or not committing suicide today? To suppose so strikes many of us as implausible and even repugnant.

Perhaps, then, we need to think of duty explicitly only when we are disinclined to act as we should. Even if so, Kant still says that we must strive to develop *virtue*, the perfection of which requires "actions being done not only in conformity with duty but also *from duty*" (MS 6:446). Does it show a defect in virtue if any motive but duty keeps us on the right path when we are otherwise inclined? Is it, for example, a moral failing or defect for someone to overcome a temporary suicidal impulse from love of his family rather than from duty, assuming that duty would have

<sup>28</sup> All duties, including perfect duties to oneself and respect for others, are based on "the end of humanity" – that is, the imperative to treat humanity in each person as an end in itself (MS 6:395).



been by itself a sufficient motive? Similarly, suppose someone who cannot altogether expunge urges to eat or drink excessively would be moderate from duty alone if need be, but her desire to avoid the social consequences of gluttony and drunkenness is her immediate and sufficient motive. In this case again the thought of duty may be unnecessary, and to insist on it as required for virtue may be to insist on one thought too many. Assuming that we accept duty as a sufficient reason in general for avoiding suicide, drunkenness, and gluttony, it is at least counterintuitive to suppose that striving for virtue means trying to ensure that duty is always our immediate conscious motive.

An alternative way to understand the ethical duty to have a moral incentive might be this: we should strive always to count the moral law as a sufficient, overriding constraint and guide as we shape our plans, policies, and characters, and then we should try to strive to live our everyday lives according.<sup>29</sup> On this understanding the thought of duty need not be our immediate and conscious motive whenever we act as we should. It would suffice, for example, if the moral law serves as an incentive to adopt, and reaffirm as need be, the obligatory ends as a permanent part of our life-plan. The ends may be internalized, not as blind habits, but as deep-rooted commitments that play a role in a full explanation of our choices even though they are not always the salient factors we call “the reason” for each act of kindness, self-improvement, etc. When other permissible aims and affections converge with our deep moral commitment, then there may be no meaningful choice about “making” one or the other our immediate reason. We would then have *both* an overriding commitment to morality and more specific maxims to promote our permissible ends. When sufficient motives “cooperate” happily in this way, it could be misleading to say that either is “the reason” for which one is acting.

In *Groundwork* I Kant asserts that only acting from duty, as opposed to inclination and mixed motives, has “moral worth” (G 4:397–99). This may seem incompatible with the broader interpretation, but arguably is not. Kant’s aim is to reveal the principle of a good will, and so he focuses on actions from duty, which most directly express a good will. Such acts have special “moral worth” as manifesting the agent’s commitment to duty above self-love. Acts from inclination and mixed motives, by contrast, may reflect the ultimate commitments of a depraved, impure, or

<sup>29</sup> See Barbara Herman, “Making Room for Character,” in *Moral Literacy* (Cambridge, Mass.: Harvard University Press, 2007), 1–28, and Marcia Baron, *Kantian Ethics Almost without Apology* (Ithaca, NY: Cornell University Press, 1995), Part II, 117–94.

weak will, none of which consistently subordinates self-love to duty. In these cases (in contrast to cases of cooperating motives discussed above), acts from inclination and mixed motives would be morally defective, not merely expressive of a good person's morally constrained personal values. It is immoral to subordinate duty to self-love (and "evil" to do so systematically), but it is not necessarily wrong or less than virtuous to do the right thing from other immediate motives.

On the broader understanding of the moral incentive requirement, I have suggested, sometimes conforming to duty without the explicit thought of duty is not a moral failing or defect. This reflects common understanding about the examples reviewed earlier. In other cases, however, the immediate thought of duty does seem appropriate and a sign of virtue. Consider, for example, a teacher trying to grade a favorite (or disliked) student, jurors trying to reach a verdict on a notorious (or popular) defendant, or a lifeguard rescuing a distressed millionaire known for her generous rewards. If the moral incentive guides as well as constrains our plans, policy, and character development, it should lead us to be more alert to duty in these cases than in the cases considered previously.

The broader understanding may also help deflect some common objections. Consider, for example, the objection that we could not have a duty always to act as we should *from duty* because we cannot control our motives at will. Kant seems to recognize this problem when he says that the duty to increase one's moral perfection, though "narrow and perfect in terms of quality," is "wide and imperfect in terms of its degree, because of the frailty (*fragilitas*) of human nature" (MS 6:446). We cannot expect ourselves to be perfect but only to strive sincerely to make progress towards the ideal. Given the conditions of human life, however, is it really ideal to have duty always at the forefront of one's mind, even when overcoming inclinations to act badly? Would this be an essential feature of moral perfection for us even if we could achieve it? The broader understanding of the moral incentive opens the way for a more attractive ideal, which is to have all our plans, attachments, and character traits so shaped by our fundamental commitment to morality that we always conform to duty but think explicitly of duty only as most appropriate to the circumstances. This is still an ideal beyond the reach of most of us, an ideal of human perfection and not a mere second-best that falls short of virtue, and it does not demand that we can call up at will a sufficiently moving thought of duty each time we act as we should.

Another familiar objection is that Kant's motive of duty represents a joyless recognition of burdensome constraint. By contrast Aristotle's

ideally virtuous person seems fully engaged, ready, and pleased to do the right thing in each situation. The broader understanding of Kant's moral incentive may help with this objection too. Undeniably Kant did think that recognition of the moral law constrains us, humbles us, and strikes down our natural self-conceit. He did not believe that non-rational human nature is as well-disposed or malleable as Aristotle did, and so he could hardly imagine that we could dispense with the constraining thought of duty. How often and how prominently the explicit thought of duty should appear in our daily lives, however, is not clear from these general features of Kant's view. If we have a good will, we are committed to doing what is right whether or not we feel so inclined, but this commitment may not need to be explicitly invoked when we have no temptations to the contrary. Similarly, assuming a good will, perhaps we need not experience duty as a constraint if such temptations are easily rejected from other (permissible) concerns that are part of our life-plans and good character. Moral law, in Kant's theory, represents both what we have good reason to do and (in imperative form) what duty constrains us to do. The more we acknowledge and come to care about the moral reasons, the less work there may be for explicit thought of duty *as a constraint*.

Finally, consider the objection that the idea of a duty to act *from duty* is incoherent because it leads to an infinite regress. For example, if I have a duty to tell the truth from duty, then it seems I must have a duty to fulfill *that* duty from duty, and then a *further* duty to fulfill that second duty from duty, and so on and on. The regress could be blocked by stipulating that the second-order (motivational) duty is only to fulfill all first-order duties (e.g. regarding specific actions or ends) from a motive of duty. The mistake, one might argue, is to assume that literally every duty must be fulfilled "from duty." This reply, however, may at first seem arbitrary. What is so special, one might ask, about the duty to act rightly from a moral incentive? Why is a moral incentive not required for this as well?

The answer, I suspect, is that if we properly understand the requirement of a general moral incentive, the request for a further moral motive makes little sense. The issues here are complex but roughly the idea is this: to have duties for Kant is basically to have sufficient, overriding rational considerations for certain choices, often experienced as constraints because they can conflict with inclinations. For a *rational* person to recognize such considerations *is* to be rationally predisposed to be moved by them, at least to some extent. As imperfectly rational agents, however, we also have conflicting inclinations and a propensity to subordinate duty to self-love. Our most fundamental choice – for Kant the

choice between good and evil – is which dispositions to incorporate into our basic life-governing maxim (KpV 5:32–39; MS 6:441). A fully rational person would choose a fundamental maxim always to give priority to the moral predisposition, and as imperfectly rational agents it is our duty to choose as we would if fully rational. If so, it is a fundamental duty, in this sense, to choose the principle of a good will, to respect and follow the unconditionally rational moral law. In other words, it is a duty to incorporate the moral disposition (or incentive) into one’s life-governing maxim. To do this is to maintain the attitude that, for the sufficient reason that the moral law requires various specific choices, one is ready and willing to make them.

This understanding of the requirement of a moral motive does not treat it as an extra duty added on to each particular duty, for example, a duty always to remember to think of the moral law whenever one conforms to it. To the contrary, the requirement is to take up a fundamental and comprehensive attitude that should shape and permeate all one’s plans, policies, and choices. Although Kant presents the duty to strive for a pure moral motive (“be holy”) as one among many in his catalogue of duties of virtue, it has a special status as part of a prior higher-order duty. On this view, choosing and maintaining the attitude of a good will makes one ready and willing to fulfill specific duties, even though the thought of duty need not be ever-present. If one could fulfill the higher-order duty completely, conformity to specific duties would follow; but mere conformity to specific duties does not fulfill them *as ethical duties* if one lacks that comprehensive commitment to the priority of the moral law (a good will). Although it sounds paradoxical, this understanding gives a sense to “a (comprehensive) duty to do one’s (specific) duties from (one’s ultimate commitment to) duty.” To make that ultimate commitment for duty over self-interest (good over evil) is what a fully rational person would do and so by definition it is what we should do, but it makes no sense to suppose it is a still further duty to make *this* ultimate choice for duty over self-love *from a prior commitment to duty*.

## 7 CONCLUDING NOTE

The *Tugendlehre* is obviously ambitious and far-reaching in its scope, but in several ways it is also incomplete. It is incomplete partly because Kant’s aim was only to present the first principles of “the doctrine of virtue,” and perhaps too because he never managed to explain and illustrate even that system of first principles as fully as he may have wished. That aside, the

system is also meant to be incomplete in another way because Kant wisely realized that there are limits to what even the best ethical theory can do. Some ethical principles leave a wide latitude for choice, and applying any principle requires judgment and understanding of the particular context. There are many matters on which moral principles do not speak, and regarding these our consciences – and Kantian moral theories – should also remain silent.

## Bibliography

- Achenwall, G. *Ius Naturae*, 5th edn., Göttingen: Bossiegel, 1763  
*Prolegomena Iuris Naturalis*, 2nd edn., Göttingen: Bossiegel, 1763
- Allison, H.E. *Kant's Theory of Freedom*, Cambridge University Press, 1990
- Ameriks, K. *Kant's Theory of Mind*, Oxford University Press, 1982, 2002
- Aristotle. *De Anima*, J.A. Smith (trans.), in J. Barnes (ed.), *The Complete Works of Aristotle*, Princeton University Press, 1984, vol. I, 641–92  
*Nicomachean Ethics*, S. Broadie (ed.), C. Rowe (ed. and trans.), Oxford University Press, 2002  
*Nicomachean Ethics*, W.D. Ross (trans.), in Barnes (ed.), *The Complete Works of Aristotle*, vol. II, 1729–867
- Arnoldt, E. *Gesammelte Schriften*, 6 vols., O. Schöndörffer (ed.), Berlin: B. Cassirer, 1907–09
- Arntzen, S. “Kant on Duty to Oneself and Resistance to Political Authority,” *Journal of the History of Philosophy* 34 (3) (1996), 409–24
- Baranzke, H. “Tierethik, Tiernatur und Moralanthropologie im Kontext von §17 Tugendlehre,” *Kant-Studien* 96 (2005), 336–63
- Baron, M. *Kantian Ethics Almost without Apology*, Ithaca, NY: Cornell University Press, 1995
- Baron, M., and M. Seymour Fahmy. “Beneficence and Other Duties of Love in *The Metaphysics of Morals*,” in Thomas E. Hill, Jr. (ed.), *The Blackwell Guide to Kant's Ethics*, Oxford and Malden, Mass.: Wiley-Blackwell, 2009, 211–28
- Baxley, A.M. “Autonomy and Autocracy,” *Kant-Studien* 94 (2003), 1–23.
- Beck, L.W. *A Commentary on Kant's Critique of Practical Reason*, University of Chicago Press, 1960
- Brandt, R. “Das Erlaubnisgesetz, oder: Vernunft und Geschichte in Kants Rechtslehre,” in R. Brandt (ed.), *Rechtsphilosophie der Aufklärung: Symposium Wolfenbüttel 1981*, Berlin: Walter de Gruyter, 1982, 233–85  
“Das Problem der Erlaubnisgesetze im Spätwerk Kants,” in O. Höffe (ed.), *Immanuel Kant: Zum ewigen Frieden*, Berlin: Akademie Verlag, 1995, 69–86
- Broadie, A., and E.M. Pybus. “Kant's Treatment of Animals,” *Philosophy* 49 (1974), 375–83
- Busch, W. *Die Entstehung der kritischen Rechtsphilosophie Kants, 1762–1780*, Berlin: Walter de Gruyter, 1979

- Byrd, B.S. "Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution," *Law and Philosophy* 8 (1989), 151–200
- Byrd, B.S., and J. Hruschka. *Kant's Doctrine of Right: A Commentary*, Cambridge University Press, 2010
- Dancy, J. *Ethics without Principles*, Oxford University Press, 2004
- Denis, L. "Animality and Agency: A Kantian Approach to Abortion," *Philosophy and Phenomenological Research* 76 (2008), 117–37
- "Kant on the Wrongness of 'Unnatural' Sex," *History of Philosophy Quarterly* 16 (2) (1999), 225–48
- "Kant's Conception of Duties Regarding Animals: Reconstruction and Reconsideration," *History of Philosophy Quarterly* 17 (2000), 405–23
- "Kant's Ethics and Duties to Oneself," *Pacific Philosophical Quarterly* 78 (1997), 321–48
- Moral Self-Regard: Duties to Oneself in Kant's Moral Theory*, New York and London: Garland Publishing, 2001
- Donagan, A. "Moral Dilemmas, Genuine and Spurious: A Comparative Anatomy," *Ethics* 104 (1993), 7–21; reprinted in J.E. Malpas (ed.), *Philosophical Papers of Alan Donagan*, vol. II, University of Chicago Press, 1994, 153–68
- "The Relation of Moral Theory to Moral Judgment: A Kantian Review," in Baruch Brody (ed.), *Moral Theory and Moral Judgments in Medical Ethics*, Dordrecht: Kluwer Academic Publishers, 1988, 171–92; reprinted in Malpas (ed.), *Philosophical Papers of Alan Donagan*, vol. II, 194–216
- The Theory of Morality*, University of Chicago Press, 1977
- Dreier, R. "Zur Einheit der praktischen Philosophie Kants," *Perspektiven der Philosophie* 5 (1979), 5–37
- Engstrom, S. "The Inner Freedom of Virtue," in M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002, 289–315.
- Feinberg, J. "The Expressive Function of Punishment," *Monist* 49 (1965), 394–423
- "The Nature and Value of Rights," *Journal of Value Inquiry* 4 (1970), 243–57
- Fichte, J.F. *Fichtes sämtliche Werke*, I.H. Fichte (ed.), Berlin: Walter de Gruyter, 1970
- Foundations of Natural Right*, M. Baur (trans.), F. Neuhouser (ed.), Cambridge University Press, 2000
- Fleischacker, S. "Kant's Theory of Punishment," *Kant-Studien* 79 (4) (1988), 434–49
- Flikschuh, K. "Freedom and Constraint in Kant's *Metaphysical Elements of Justice*," in S. Byrd and J. Hruschka (eds.), *Kant and Law*, Hants: Ashgate, 2006, 87–108
- "Hobbes and Kant on Public Authority," *Hobbes Studies* (in press)
- "Ist das rechtliche Postulat ein Postulat der reinen praktischen Vernunft? Zum Endzweck der Kantischen Rechtslehre," *Jahrbuch für Recht und Ethik* 12 (2004), 299–330

- Kant and Modern Political Philosophy*, Cambridge University Press, 2000
- “Kant’s Kingdom of Ends: Metaphysical, Not Political,” in J. Timmermann (ed.), *Kant’s Groundwork of the Metaphysics of Morals: A Critical Guide*, Cambridge University Press 2009, 119–39
- “Reason, Right, and Revolution: Locke and Kant,” *Philosophy and Public Affairs* 36 (2008), 375–404
- Freud, S. “Formulations on the Two Principles of Mental Functioning,” in *The Case of Schreber, Papers on Technique and Other Works*, Standard Edition of the Complete Psychological Works of Sigmund Freud, vol. XII (1911–13), J. Strachey (ed.), London: Hogarth Press, 1958, 218–26
- “Formulierungen über die zwei Prinzipien des psychischen Geschehens,” in A. Mitscherlich (ed.), *Studienausgabe*, 8th edn., Frankfurt am Main: S. Fischer, 1997, vol. III, 13–24
- Frierson, P. “Kant’s Empirical Account of Human Action,” *Philosophers’ Imprint* 5 (7) (2005), 1–34
- Gregor, M.J. *Laws of Freedom: A Study of Kant’s Method of Applying the Categorical Imperative in the Metaphysik der Sitten*, Oxford: Basil Blackwell, 1963
- “Translator’s Note on the Text of the *Metaphysics of Morals*,” in Immanuel Kant, *Practical Philosophy*, M.J. Gregor (ed. and trans.), Cambridge University Press, 1996, 355–59
- Grenberg, J.M. *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue*, Cambridge University Press, 2005
- “Replies,” *Philosophy and Phenomenological Research* 75 (3) (2007), 640–54
- Grimm, J. and G. (eds.). *Deutsches Wörterbuch*, 33 vols, Munich: Deutscher Taschenbuch Verlag, 1984, vol. XXII.
- Grotius, H. *De Iure Belli ac Pacis*, B.J.A. de Kanter-van Hettinga Tromp (ed.), Louvain: Brill, 1939
- Guevara, D. *Kant’s Theory of Moral Motivation*, Boulder, Colo.: Westview Press, 2000
- Guyer, P. “Duties Regarding Nature,” in *Kant and the Experience of Freedom*, 303–34  
*Kant and the Experience of Freedom*, Cambridge University Press, 1993  
*Kant on Freedom, Law, and Happiness*, Cambridge University Press, 2000
- Hare, J.E. *The Moral Gap*, Oxford: Clarendon Press, 1996
- Hampton, J. “The Moral Education Theory of Punishment,” *Philosophy and Public Affairs* 13 (3) (1984), 208–38
- Herman, B. “Making Room for Character,” in *Moral Literacy*, 1–28.  
*Moral Literacy*, Cambridge, Mass.: Harvard University Press, 2007  
*The Practice of Moral Judgment*, Cambridge, Mass.: Harvard University Press, 1993
- “Rethinking Kant’s Hedonism,” in *Moral Literacy*, 176–202
- Hill, Thomas E., Jr. “Donagan’s Kant,” *Ethics* 104 (1993), 22–52; reprinted in Thomas E. Hill, Jr., *Respect, Pluralism, and Justice*, Oxford University Press, 2000, 119–51
- “Four Conceptions of Conscience,” *Nomos* 40 (1998), 13–52; reprinted in *Human Welfare and Moral Worth*, 277–309



- Human Welfare and Moral Worth: Kantian Perspectives*, Oxford University Press, 2002
- "Kant on Imperfect Duty and Supererogation," *Kant-Studien* 72 (1971), 55–76
- "Meeting Needs and Doing Favors," in *Human Welfare and Moral Worth*, 201–44
- "Wrongdoing, Desert, and Punishment," in *Human Welfare and Moral Worth*, 310–39; originally published as "Kant on Wrongdoing, Desert and Punishment," *Law and Philosophy* 18 (1) (1999), 407–75
- Hobbes. T. *Leviathan* [1651], E. Curley (ed.), Indianapolis, Ind.: Hackett, 1994
- Höffe, O. "Categorical Imperatives of Right According to Ulpian," in *Kant's Cosmopolitan Theory of Law and Peace*, A. Newton (trans.), Cambridge University Press, 2006, 119–33
- Categorical Principles of Law: A Counterpoint to Modernity*, M. Migotti (trans.), University Park: Pennsylvania State University Press, 2002
- "Der kategorische Rechtsimperativ. 'Einleitung in die Rechtslehre,'" in O. Höffe (ed.), *Immanuel Kant, Metaphysische Anfangsgründe der Rechtslehre*, Klassiker Auslegen XIX, Berlin: C.H. Beck-Verlag, 1999, 41–62
- "Kategorische Rechtsimperative nach Ulpian," in "*Königliche Völker*": *Zu Kants kosmopolitischer Rechts- und Friedenstheorie*, Frankfurt am Main: Suhrkamp Verlag, 2001, 147–60
- Kategorische Rechtsprinzipien. Ein Kontrapunkt der Moderne*, Frankfurt am Main: Suhrkamp Verlag, 1995
- Political Justice: Foundation for a Liberal Philosophy of Law and the State*, J.C. Cohen (trans.), Cambridge: Polity Press, 1995
- Politische Gerechtigkeit. Grundlegung einer kritischen Philosophie von Recht und Staat*, Frankfurt am Main: Suhrkamp Verlag, 1987
- "The Principle of Justice as Categorical Imperative of Law," in Y. Yovel (ed.), *Kant's Practical Philosophy Reconsidered*, Dordrecht: Kluwer Academic Publishers, 1989, 149–67
- Hohfeld, W. *Fundamental Legal Conceptions*, W. Cook (ed.), New Haven: Yale University Press, 1919
- Hruschka, J. "The Permissive Law of Practical Reason in Kant's Metaphysics of Morals," *Law and Philosophy* 23 (2004), 45–72
- Hume, D. *A Treatise of Human Nature*, 2nd edn., L.A. Selby-Bigge and P.H. Niddich (eds.), Oxford: Clarendon Press, 1978
- Hutcheson, F. *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*, A. Garret (ed.), Indianapolis, Ind.: Liberty Fund, 2002
- Ilting, K.-H. "Gibt es eine kritische Ethik und Rechtsphilosophie Kants?" *Archiv für Geschichte der Philosophie* 63 (1983), 325–45
- Ingensiep, H.W. "Tierseele und tierethische Argumentationen in der Deutschen philosophischen Literatur des 18. Jahrhunderts," *NTM: Internationale Zeitschrift für Geschichte und Ethik der Naturwissenschaften, Technik und Medizin* (N.S.) 4 (1996), 103–18
- Jellinek, G. *System der subjektiven öffentlichen Rechte*, Tübingen: Mohr, 1892, 1905

- Johnson, R.N. "Self-Improvement: An Essay in Kantian Ethics," unpublished manuscript
- "Value and Autonomy in Kantian Ethics," in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, vol. II, Oxford University Press, 2007, 33–48
- Kain, P. "Kant's Defense of Human Moral Status," *Journal of the History of Philosophy* 47 (2009), 59–102
- "Self-Legislation in Kant's Moral Philosophy," *Archiv für Geschichte der Philosophie* 86 (2004), 257–306
- Kant, I. *Die Metaphysik der Sitten in zwey Theilen. Erster Teil: Metaphysische Anfangsgründeder Rechtslehre; Zweiter Teil: Metaphysische Anfangsgründe der Tugendlehre*, Königsberg: Friedrich Nicolovius, 1797
- Erläuternde Anmerkungen zu der Rechtslehre*, Königsberg: Friedrich Nicolovius, 1798
- Groundwork for the Metaphysics of Morals*, T.E. Hill, Jr. (ed.), A. Zweig (trans.), Oxford University Press, 2002
- Metaphysik der Sitten*, 2nd edn., K. Vorländer (ed.), Leipzig: Verlag der Dürr'schen Buchhandlung, 1907
- Metaphysische Anfangsgründe der Rechtslehre: Metaphysik der Sitten, Erster Teil*, B. Ludwig (ed.), Hamburg: Felix Meiner Verlag, 1986, 1998
- Metaphysische Anfangsgründe der Tugendlehre: Metaphysik der Sitten, Zweiter Teil*, B. Ludwig (ed.), Hamburg: Felix Meiner Verlag, 1990
- Vorlesungen zur Moralphilosophie*, W. Stark (ed.), Introduction by M. Kuehn, Berlin: Walter de Gruyter, 2004
- Kelsen, H. *Introduction to the Problems of Legal Theory: A Translation of the First Edition of the Reine Rechtslehre or Pure Theory of Law*, B. Litschewski Paulson and S.L. Paulson (trans.), Oxford University Press, 1997
- Reine Rechtslehre*, Vienna: Auflage, 1960
- Kerstein, S.J. "Treating Oneself Merely as a Means," in M. Betzler (ed.), *Kant's Ethics of Virtue*, Berlin: Walter de Gruyter, 2008, 201–18
- Kersting, W. *Wohlgeordnete Freiheit*, Berlin: Walter de Gruyter, 1984
- Köhl, H. *Kants Gesinnungsethik*, Berlin: Walter de Gruyter, 1990
- Korsgaard, C.M. *Creating the Kingdom of Ends*, Cambridge University Press, 1996
- "Fellow Creatures: Kantian Ethics and Our Duties to Animals," in G.B. Peterson (ed.), *The Tanner Lectures on Human Values*, vol. xxv, Salt Lake City: Utah University Press, 2005, 77–110
- "Interacting with Animals: A Kantian Account," in T. Beauchamp and R.G. Frey (eds.), *The Oxford Handbook on Ethics and Animals*, Oxford University Press, in press
- The Sources of Normativity*, O. O'Neill (ed.), Cambridge University Press, 1996
- Kuehn, M. "Introduction," in Immanuel Kant, *Anthropology from a Pragmatic Point of View*, R.B. Loudon (ed. and trans.), Cambridge University Press, 2006, vii–xxix
- Kymlicka, W. *Liberalism, Community, and Culture*, Oxford University Press, 1991

- Mackie, J. "Can There Be a Rights-Based Moral Theory?" *Midwest Studies in Philosophy* 3 (1978), 350–59
- Meier, G.F. *Philosophische Sittenlehre*, Halle: Hemmerde, 1753–61, \$975
- Morris, H. "Persons and Punishment," *Monist* 52 (1968), 475–501
- Mulholland, L. *Kant's System of Rights*, New York: Columbia University Press, 1990
- Murphy, J.G. *Kant: The Philosophy of Right*, London: Macmillan, 1970
- Naragon, S. "Kant in the Classroom," online at [www.manchester.edu/kant/](http://www.manchester.edu/kant/)
- "Kant on Descartes and the Brutes," *Kant-Studien* 81 (1990), 1–23
- "Reason and Animals: Descartes, Kant, and Mead on the Place of Humans in Nature," unpublished doctoral dissertation, University of Notre Dame (1987)
- Nozick, R. *Anarchy, State, and Utopia*, New York: Basic Books, 1974
- Nussbaum, M. *Frontiers of Justice: Disabilities, Nationality, Species Membership*, Cambridge, Mass.: Harvard University Press, 2006
- Oberer, H. "Zur Frühgeschichte der Kantischen Rechtslehre," *Kant-Studien* 64 (1973), 88–102
- Oberhausen, M., and R. Pozzo (eds.), *Vorlesungsverzeichnisse der Universität Königsberg (1720–1804)*, 2 vols., Stuttgart-Bad Cannstatt: frommann-holzboog, 1999
- O'Neill, O. *Bounds of Justice*, Cambridge University Press, 2000
- Constructions of Reason: Explorations of Kant's Practical Philosophy*, Cambridge University Press, 1989
- Pincoffs, E. *The Rationale of Legal Punishment*, New York: Humanities Press, 1966
- Potter, N. "Does Kant Have Two Concepts of Freedom?" in G. Funke and J. Kopper (eds.), *Akten des Vierten Internationalen Kant Kongresses*, Berlin: Walter de Gruyter, 1974, 590–96
- "Duties to Oneself, Motivational Internalism, and Self-Deception in Kant's Ethics," in M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002, 371–89
- "Kant on Punishment," in Thomas E. Hill, Jr. (ed.), *The Blackwell Guide to Kant's Ethics*, Oxford: Wiley-Blackwell, 2009, 179–95
- Pybus, E., and A. Broadie. "Kant's Treatment of Animals," *Philosophy* 49 (1974), 375–83
- Rawls, J. *Lectures on the History of Moral Philosophy*, Cambridge, Mass.: Harvard University Press 2000
- "Themes in Kant's Moral Philosophy," in E. Förster (ed.), *Kant's Transcendental Deductions: The Three "Critiques" and the "Opus postumum"*, Stanford University Press, 1989, 81–113
- A Theory of Justice*, Cambridge, Mass.: Harvard University Press, 1971
- Reath, A. *Agency and Autonomy in Kant's Moral Theory: Selected Essays*, Oxford University Press, 2006
- "Self-Legislation and Duties to Oneself," in M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002, 349–70; reprinted in *Agency and Autonomy*, 231–49

- Regan, T. *The Case for Animal Rights*, London: Routledge & Kegan Paul, 1984
- Ripstein, A. "Authority and Coercion," *Philosophy and Public Affairs* 31 (1) (2004), 2–35
- Ritter, C. *Der Rechtsgedanke nach den frühen Quellen Kants*, Frankfurt am Main: Vittorio Klostermann, 1971
- Ross, W. D. *The Right and the Good*, Philip Stratton-Lake (ed.), Oxford: Clarendon Press, 2001
- Scanlon, T.M. *What We Owe Each Other*, Cambridge, Mass: Harvard University Press, 1998
- Scheid, D.E. "Kant's Retributivism," *Ethics* 93 (1983), 262–82
- Schleiermacher, F. *Kritische Gesamtausgabe*, Berlin: Walter de Gruyter, 1984
- Schmucker, J. Rezension, in *Die Ursprünge der Ethik Kants in seinen vorkritischen Schriften und Reflexionen*, Meisenheim am Glan: A. Hain, 1961, 365–69
- Schopenhauer, A. *Werke in fünf Bänden*, Ludger Lüdtkehaus (ed.), Ulm: Haffmanns Verlag, 1988, vols. 1 and III
- Sensen, O. "Dignity and the Formula of Humanity," in J. Timmermann (ed.), *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*, Cambridge University Press, 2009, 102–18
- Shell, S.M. "Kant on Punishment," *Kantian Review* 1 (1997), 115–36
- Sherman, N. *Making a Necessity of Virtue*, Cambridge University Press, 1997
- Singer, P. *Animal Liberation: A New Ethics for Our Treatment of Animals*, New York: Random House, 1975
- Thomasius, C. *Fundamenta Juris Naturae et Gentium*, 4th edn. [1718], Aalen: Scientia, reprint 1963
- Timmermann, J. "Kant on Conscience, 'Indirect' Duty, and Moral Error," *International Philosophical Quarterly* 46 (2006), 293–308
- "Kantian Duties to the Self, Explained and Defended," *Philosophy* 81 (2006), 505–29
- "When the Tail Wags the Dog: Animal Welfare and Indirect Duty in Kantian Ethics," *Kantian Review* 10 (2005), 128–49
- Timmons, M. (ed.). *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002
- Unna, Y. "Kant's Answers to the Casuistical Questions Concerning Self-Disembodiment," *Kant-Studien* 94 (2003), 454–73
- Vasilyev, V., C. Korsgaard, and P. Singer. "International Kant Interview, 2004–2005," online at [www.philos.msu.ru/community/staff/vasiliev/Kant\\_Interview/Kant\\_Interview.htm](http://www.philos.msu.ru/community/staff/vasiliev/Kant_Interview/Kant_Interview.htm)
- Vorländer, K. "Einleitung," in Immanuel Kant, *Metaphysik der Sitten*, 2nd edn., Leipzig: Verlag der Dürr'schen Buchhandlung, 1907, ix–li
- Walch, J.G. "Zulassung," in J.C. Hennings (ed.), *Philosophisches Lexicon*, 4th edn. [1775], Hildesheim: Olms, reprint 1968, vol. II, col. 1708
- Warren, M.A. *Moral Status: Obligation to Persons and Other Living Things*, Oxford University Press, 1997
- Wellman, C. *Real Rights*, Oxford University Press, 1995

- Westphal, K.R. "Do Kant's Principles Justify Property or Usufruct?" *Jahrbuch für Recht und Ethik* 5 (1997), 142–94
- Willaschek, M. "Which Imperatives for Right? On the Non-Prescriptive Character of Juridical Law in Kant's *Metaphysics of Morals*," in M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002, 65–87
- "Why the *Doctrine of Right* Does Not Belong in the *Metaphysics of Morals*," *Jahrbuch für Recht und Ethik* 5 (1997), 205–27
- Williams, H. *Kant's Political Philosophy*, London: St. Martin's Press, 1983
- Wood, A.W. "The Final Form of Kant's Practical Philosophy," in M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford University Press, 2002, 1–21
- "General Introduction," in M.J. Gregor (ed. and trans.), *Immanuel Kant, Practical Philosophy*, Cambridge University Press, 1996, xiii–xxxiii
- Hegel's Ethical Thought*, Cambridge University Press, 1990
- "Kant on Duties Regarding Nonrational Nature I," *Proceedings of the Aristotelian Society, Supplementary Volumes* 72 (1998), 189–210
- Kantian Ethics*, Cambridge University Press, 2008
- Kant's Ethical Thought*, Cambridge University Press, 1999

# Index

- Achenwall, Gottfried, 19, 103, 104, 105–06, 107
- acquisition  
external, principle of, 107, 109  
original, 107–10
- action  
and desire, 37–38  
determination of, 13, 21, 33, 54–55, 58, 99,  
135–36, 137, 140–41, 154  
from duty, 25–26, 60, 130, 251–53  
aesthetic preconditions of receptivity to duty,  
5, 137–51, 222–23
- agency  
conditions of, 127, 176–78, 181, 182–83, 247  
power of, 47, 49  
use of, 197–99
- Allison, Henry E., 49, 58
- Ameriks, Karl, 215
- amphiboly, 195, 210, 231–32
- anarchism, 81, 82, 83
- animality, 37
- animals  
duties regarding, 7, 195–96, 210–12, 231–33  
as duties to oneself, 195, 227  
moral status of, 7, 195–96, 220–21, 223–25,  
228–32  
nature of, 214–19, 222, 223–25
- anthropology, 7, 26, 219, 235  
and moral philosophy, 13, 19, 21–22  
and right, 76–78, 84
- antinomy of right, 81–82, 83
- apathy, 166, 168
- Aquinas, St. Thomas, 242
- Aristotle, 38, 46, 80, 165–69, 231
- Arnoldt, Emil, 17
- Arntzen, Sven, 256
- authority  
legislative, 59, 67–70, 74–75  
obligating, 78, 181, 185, 186–87, 245
- autocracy, 176, 178
- autonomy, 28, 50, 53–54, 61–62  
and duties to oneself, 183–84, 245
- and ethical lawgiving, 54  
as ground of morality, 171–73, 183–84  
as internal freedom, 58, 60  
and obligation, 53–56  
as self-legislation, 57–59  
as spontaneity of rational cognition, 50
- Baranzke, Heike, 211, 221
- Baron, Marcia, 147, 192, 241, 251
- Baumgarten, Alexander Gottlieb, 15, 17–20,  
22, 114, 125
- Baxley, Anne Margaret, 176
- beavers, 216
- Beccaria, Cesare, 128
- Beck, Lewis White, 13, 49, 58
- beneficence, 145–46
- Bentham, Jeremy, 242
- Bonnet, Charles, 216
- Brandt, Reinhard, 96, 100, 101
- Broadie, Alexander, 195, 211
- Buffon, Georges-Louis Leclerc, *Compte de*,  
217, 219
- Busch, Werner, 10
- Byrd, B. Sharon, 113, 121
- categorical imperative, 19, 23, 24, 25, 26, 65,  
196, 203, 213, 217, 219, 220, 240, 245  
formula of autonomy, 125, 171  
formula of humanity, 24, 84, 85, 122, 173,  
196–97, 212, 250  
formula of the realm of ends, 125  
formula of universal law, 53, 122  
formula of universal law of nature, 182  
of right, 80, 84
- causality, 180, 236  
and choice, 47–48  
and desire, 35–40, 42  
and feeling(s), 130–51  
and means, 196–97  
and practical knowledge, 31–34  
of representations, 35

- and will, 46, 49  
 censorship edict, 15–16  
 choice, 44, 47, 49, 101  
   power of, 46–50, 58–60, 98–99, 214  
   determination of, 48–50, 140–41  
 coercion  
   authorized, 80–84, 117–20  
   to enter civil society, 101–03  
   legitimate  
     opposing illegitimate, 83–84  
     preventive/restitutive, 83  
     punishment as, 117–22  
   and perfect duties to oneself, 174, 179  
   types of, 82–83  
 coexistence, 60–63, 76, 77, 78–80, 81, 82, 84, 90, 103  
 cognition, 32–34, 215–16, 231  
   discursive, 33, 34, 42, 44  
   formal unity of, 42  
   practical, 32, 33, 44–46  
   rational, 31, 33, 34, 38, 48, 50, 221  
   and self-consciousness, 44–45  
   theoretical, 31, 34  
   theoretical/practical, 42–43  
 concepts, faculty of, 39  
 concubinage, 187–88  
 conscience, 18, 143–45, 247–48  
 consequentialism, 19  
 contemplation, 164–66  
 contentment, 136  
 continence, 166–68
- Dancy, Jonathan, 239–40  
 death penalty, 112–13, 120  
 “Déclaration des droits de l’homme et du citoyen”, 88  
 deliberation, 47–48  
 Denis, Lara, 195, 222, 224, 225, 237, 243  
 desire, 2–3, 35–43  
   faculty of, 35–36, 39, 40, 215  
   determination of, 38, 42–43, 45, 47, 50  
   sensible, 37, 39–43  
 dignity, 87, 89, 177, 178, 183, 186, 211, 212  
   and self-degradation, 189–90  
 dogs, 217–18, 224–25  
 Donagan, Alan, 239  
 Dreier, Ralf, 53  
 dueling, 120  
 duties, *see also* animals, duties regarding; duties to oneself; duties to others  
   division of, 13, 17–18, 23–24, 173–74, 179  
   ethical, 24–26, 122, 250, 251–54  
     as duties to oneself, 183–84  
   ethical/juridical, 237  
   juridical  
   as indirectly ethical, 51–52, 56, 64, 185, 187–88  
   of love, 145–49  
   prima facie, 240  
   of right, 64  
   to/regarding, 6–7, 192–204, 207–09, 210, 231–32, 244  
   of virtue, 24–26, 64, 122, 240–41, *see also* obligatory ends  
 duties regarding animals *see* animals, duties regarding  
 duties to oneself, 14, 242–46  
   and autonomy, 245  
   and duties to others, 150, 173, 183, 199, 199–200  
   and obligation, 244  
   perfect, 174–91, 237  
   as ethical, 174  
   and freedom, 170–71, 173, 174–76, 177–78, 180–81, 182, 190–91  
   as inner duties of right, 179, 187–88  
   and obligation, 185, 186  
   primacy of, 6, 170–71, 173, 182–91  
   and virtue, 176–77, 182–83  
   perfect/imperfect, 176, 185  
   primacy of (as such), 183–84, 242  
   and prudence, 243  
   regarding animals, 195, 227  
   and right, 86–87  
 duties to others, 196  
   imperfect, 6–7, 194, 197–98, 204, 205–06, 207–09
- elephants, 217, 218–19, 225  
 ends, 32, 156, 160, 161, 196, 231, *see also* essential ends; obligatory ends  
   natural, 214, 241–42  
   of pure practical reason, 22–27  
 Engstrom, Stephen, 160, 165–66, 168  
 Epicurean ethics, 29  
 equality, 88  
   innate, 63, 88  
 Erhard, Johann Benjamin, 14  
 essential ends, 25  
   of human nature, 27  
   of humanity, 6, 174–76, 177, 181  
 ethics, 24–26, 56, *see also* duties, ethical, of virtue; lawgiving, ethical/juridical; obligation, ethical, ethical/juridical  
   and coercion, 61–62, 117  
   normative, 8, 235–37, 254–55  
   and right, 236–37  
 eudaimonism, 19  
 evil, propensity to  
   as enemy of virtue, 152–53, 155–60, 167, 169  
   and freedom, 155–60

- external objects of choice, 93, 100–01  
rights to, 4–5, 85–87, 93–94, 95–110,  
*see also* possession; rights, property
- fact of reason, 213
- faculty of desire in accordance with concepts,  
43, *see also* choice; will; wish
- Fahmy, Melissa Seymour, 241
- feeling(s), 5, 35, 36, 43, *see also* moral feeling;  
moral motivation, phenomenal  
cultivation of, 139–40, 146–48, 169, 223  
and duties to oneself regarding animals,  
222–27  
and moral motivation, 130–31, 132–34
- Feinberg, Joel, 202
- Fichte, Johann Gottlieb, 14, 121, 122
- Flikschuh, Katrin, 96
- freedom  
external, 59–60, 62–63, 64, 77  
and possession, 98–103  
restriction of, 78–80, 82–84, 117–20,  
*see also* coercion  
as fact of reason, 213  
as ground of morality, 3, 58, 171–73  
and humanity, 85, 88, 171–73  
inner, 6, 160–69, 176, 177, 178  
internal/external, 58, 60–62, 173  
medieval concepts of, 78  
and moral motivation, 131–32, 135, 138, 151  
and moral status, 213  
non-autonomous, 3, 59–60, 64  
and obligation, 173  
and perfect duties to oneself, 6, 170–71, 173,  
177–78, 180–81, 182, 190–91  
as a predisposition, 217, 220  
and propensity to evil, 155–60  
and rights, 88, 173  
and self-constraint, 180–81  
and self-legislation, 172  
and vice, 157–58, 161–65  
and virtue, 153–58, 160–69
- Freud, Sigmund, 81
- Frierson, Patrick, 215
- Gemüt*, 137, 139
- general will, 3, 4, 59–60, 62–63, 64, 68–69,  
94, 120  
possible, 107–10  
Rousseauan, 68
- Gesinnung*, 20, 26
- God  
and conscience, 144  
grace of, 159  
and the highest good, 124  
moral attributes of, 124–26  
Golden Bull of Andrew II of Hungary, 78  
goodness of will, 131, 242, 248–49, 251, 253–54  
and internal freedom, 52, 60–64  
as worthiness to be happy, 124
- grace, divine, 159
- gratitude, 146
- Gregor, Mary J., 10, 99–100, 137, 146, 170, 181,  
184, 186
- Grimm, J. and W., 72
- Grotius, Hugo, 103, 107
- Guevara, Daniel, 134
- Guyer, Paul, 171, 195, 223
- Hamann, Johann Gottfried, 12
- happiness, 47, 88, 163, 197–98, 216, 245
- Hare, John E., 159
- Hartenstein, Gustav, 23
- Herder, Johann Gottfried, 12, 14
- Herman, Barbara, 42, 43, 62, 148, 251
- Herz, Marcus, 12
- Hill, Thomas E., Jr., 113, 121, 192
- Hobbes, Thomas, 75, 241, 246
- Höffe, Otfried, 53
- Hohfeld, Wesley, 202
- holiness, 164, 167, 169
- homo noumenon/homo phenomenon*, 85, 172,  
178–79
- honeste vive*, 85–87
- honor, 4, 85–87, 120, 180
- horses, 224
- Hruschka, Joachim, 93, 96, 103, 108
- human rights, 4, 71–72, 74, 78, 79–80, 207  
history of, 78  
implicit in the innate right, 87–90  
“quasi”, 4, 92
- humanity, 4, 84–87, 178, 212, *see also*  
categorical imperative, formula of  
humanity; right of humanity in our  
own person  
authority of, 181, 185, 186–87, 188–89, 245  
dignity of, 87, 186, 189–90  
and freedom, 85, 171–73  
practical point of view and, 190
- Hume, David, 33, 241
- Hutcheson, Francis, 12, 133
- Iltting, Karl-Heinz, 10
- imagination, 37, 41, 215
- inclination, 40–41, 251–52  
as enemy of virtue, 152–53, 155–58, 161–62  
as natural force, 153, 155, 161–62
- infanticide, 120
- Ingensiep, Hans Werner, 219



- innate right to freedom, 4, 5, 79–80, 84, 87, 88, 93, 108–09, 118–19
- instincts, 40, 216
- intellect, 43
  - discursive, 30
  - infinite, 34
- interest, 40–42
- intuitionism, 239–40
- Jellinek, Georg, 71
- judgment, 44–46
- Jung-Stilling, Heinrich, 14
- jus*, 20, 178
- justice
  - acting justly/being just, 51–52, 64–65, 66, 69–70
  - Aristotelian, 51, 64–65, 66, 70
  - divine, 124–26
  - idea of, 80
  - just/unjust, 74
  - punitive *see* punishment
- Kantian liberalism, 52–53, 65–67, 69–70
- Kelsen, Hans, 75
- Kerstein, Samuel, 173
- Kersting, Wolfgang, 96
- knowledge, 48, 49
  - faculty of, 39, 50
  - practical, as self-knowledge, 32–33, 48
  - pragmatic, 217
  - rational, 29, 30–31, 242
  - theoretical/practical, 29, 31–34, 35, 36–37
- Köhl, Harald, 139
- Korsgaard, Christine, 57, 211, 219, 229–31
- Kymlicka, Will, 205
- Lambert, Johann Heinrich, 11, 12
- lawgiving
  - ethical/judicial, 54–56, 69–70, 77
  - judicial, 65, 66
- laws
  - ethical/judicial, 54–56, 74
  - of freedom, 29, 30, 31, 74
    - internal/external, 60–61
    - universal, 76–77
  - of nature, 29, 30, 31
- legal positivism, 75, 81, 82, 83
- Leibniz, Gottfried Wilhelm, 231
- lex talionis*, 112, 115, 116, 117, 118, 121, 128
- life, 36, 37
- life-power, 36, 39
- Linnaeus, 217
- Locke, John, 246
- logic, 29, 30
- love, 197
  - and duties regarding animals, 223–27
  - of human beings, 145–49
- Ludwig, Bernd, 9, 28
- Ludwig, Michael, 71
- lying, 89–90
- Mackie, John, 202
- Magna Carta Libertatum, 78
- magnanimity, 165–67
- maxims, 41, 48, 144, 184, 249
  - of action, 23, 24, 47, 155, 176, 182, 225
  - of ends, 47, 49, 122, 174, 176, 238
  - fundamental, 131, 135–36, 138, 151, 157, 158, 254
  - particular, 136, 138, 139, 145, 147–48, 151, 251
  - and right, 62, 65, 66–67, 70
- metaethics, 236
- metaphysics, 21, 235–36
  - of morals, 29, 76, 77, 236–37
  - of nature, 21, 29, 236
- Metaphysics of Morals*
  - Kant's conception of, 2, 21–22, 72–75, 234, 235–38, 254–55
- MitfreudelMitleid*, 146
- monkeys, 217
- Montesquieu, Charles-Louis Secondat, 127
- moral epistemology, 238–40
- moral feeling, 18, 140–42, 165
  - as causal, 138, 140, 142
  - and contemplation, 166
  - cultivation of, 142
  - respect as, 135, 136, 141–42
  - and virtue, 164
- moral health, 6, 176–78, 181, 182–83
- moral motivation, 5, 130–51
  - and moral incentives, 251–54
  - phenomenal, 131–51
    - and freedom, 131–32, 135, 138, 151
    - multistage model of, 138, 139–40, 142, 143, 144, 145, 147, 149, 150–51
- moral philosophy, 29–30
  - and anthropology, 21–22
  - as rational knowledge, 29
- moral principles
  - fundamental, 237–42, 254–55
  - second-order, 246–49
- moral prosperity, 176
- moral self-preservation, 6, 176–78, 181, 182–83, 225
- moral sense theory, 142
- moral status, 7, 172, 195–97, 198–99, 210, 212–13
  - animal, 220–21, 223–25, 228–32
  - human, 220
    - marginal cases, 210, 219, 220

- Mulholland, Leslie A., 187
- Naragon, Steve, 215, 217
- natural law theory, 10
- natural teleology, 214, 241–42
- negligence, 247
- Nietzsche, Friedrich, 27
- Nozick, Robert, 81
- Nussbaum, Martha, 211
- Oberer, Hariolf, 10
- object, 33, 41, 43, 91, 196, 198, 201, 213, 219,  
*see also* ends; external objects of choice;  
 perception; representation
- obligation, 196, 199–200, 212–13, 244  
 active/passive, 194–95, 212, 213  
 ethical, 3, 172–73, 178, 185  
 ethical/juridical, 54–56  
 and freedom, 173  
 and self-constraint, 180–81, 183–84, 186  
 of virtue, 24–26
- obligatory ends, 22–27, 122, 174, 179, 226, 237,  
 238, 240–41, *see also* duties, of virtue;  
 happiness; perfection  
 one's own perfection, 10, 150, 173, 176,  
 184–85, 222, 243, 248, 252  
 others' happiness, 6, 122–24, 194, 197–98,  
 200, 204, 205–06, 207–09
- O'Neill, Onora, 57, 62, 211
- original community of earth, 107–10
- passions, 157–58
- perception, 35, 36–37, 38–40, 42–43, 142
- perfection, 10, 34, 150, 173, 177, 248–49, 252
- permissive law  
 as justification, 96–103  
 postulate of practical reason as, 95–105  
 as power conferring norm, 103–05
- person(s), 74, 78, 217  
 concept of, 88, 172
- personality, 85, 212  
 predisposition to, 220
- philosophy  
 critical/precritical, 2, 10, 16–17, 20–21, 26–27  
 divisions of, 29–30  
 moral *see* moral philosophy  
 practical, 25, 154  
 as rational knowledge, 29
- pleasure, 13, 35–38, 39–43, 47, 134, 135, 140–41,  
 178, 215, 223–24  
 intellectual, 42  
 practical, 39
- pleasure principle, 81
- possession, 105  
 common/particular, 109–10  
 concepts of, 4, 94–95  
 empirical physical, 94–95, 96, 105, 106  
 and external freedom, 98–103  
 intelligible, 90, 95, 106–07  
 security in, 102  
 juridical, 106  
 pure concept of understanding, 95, 105  
 postulate of practical reason, 90–91, 95, 107  
 as justification, 4, 94, 96–103  
 as a permissive law, 95–105  
 as power conferring norm, 4–5, 94  
 postulate of public law, 101–03
- Potter, Nelson, 58, 177, 184
- practical point of view  
 and humanity, 190  
 and virtue, 153–58
- predispositions, 215, 216–17, 220
- property, 102, *see also* rights, property
- prostitution, 187–88
- prudence, 243, 246
- public legal order  
 right(s) to live in, 4, 92
- Pufendorf, Samuel, 10, 107
- punishment, 5, 111–29, *see also* retributivism  
 civil/natural, 116–17  
 as coercive protection of right, 117–22, 123  
 concepts of, 113–14, 116–17  
 justification of, 5, 114–22, 123  
 consequentialist, 111, 114–16, 118–19,  
 120–22, 127–28  
 retributivist, 111, 115–16, 117, 121–22,  
 126–29  
 law of, 112, 115, 118
- punishments  
 pragmatic, 113  
 corrective/exemplary, 114–15, 121, 126  
 retributive, 113, 115, 121, 126
- purity, 248
- Pybus, Elizabeth, 195, 211
- rational nature *see* humanity
- Rauscher, Frederick, 172
- Rawls, John, 42, 52, 81, 246
- reality principle, 81
- reason, 2–3, 22, 29–35  
 discursive, 34  
 logical conception of, 30  
 practical application of, 21, 42, 48, 50  
 primacy over desire, 49  
 as spontaneity, 31, 33–34, 38  
 theoretical application of, 235  
 theoretical/practical applications of, 12,  
 29–35, 236  
 unity of, 29–30, 35, 48, 235  
 and vice, 155–59

- and virtue, 153  
 as will, 42, 48–50, 58, 59
- Reath, Andrews, 42, 183, 195, 196, 201–02
- Recht*, 10, 56, 178, 236  
 and perfect duties to oneself, 178–79
- reflection, 31
- reflective equilibrium, 239
- Regan, Tom, 211
- representation, 35–43, 49, 135, 140, 216, 231  
 faculty of, 39
- respect, 5, 130–31, 149, 210, 250, *see also* self-respect  
 and animals, 228–31  
 as causal, 134–36, 144, 147, 150, 151  
 and conscience, 144  
 and contemplation, 164  
 as epiphenomenal, 133–34  
 and gratitude, 146  
 as moral feeling, 135, 136, 138, 141–42  
 and self-degradation, 189–90  
 and self-respect, 150  
 and sympathy, 147
- retribution, the law of *see lex talionis*
- retributivism, 111–14, 117, 121–29  
 as fundamental, 115–16, 117, 121–22, 126–29  
 as incompatible with divine justice, 124–26  
 as inconsistent with duties of virtue, 122–24
- right  
 antinomy of, 81–82, 83  
 categorical imperative of, 80, 84  
 and coercion, 57, 60–63, 65, 68–69, 84, 117–22  
 “doctrine” of, 72–75  
 and duties to oneself, 86–87  
 and ethics, 3, 89–90, 236–37  
 inner/outer, 187  
 moral concept of, 74, 75–80, 82, 91  
 moral status of, 3, 53–59, 63–66  
 natural/empirical, 72–75  
 priority of, 237  
 as public morality, 3, 52–53, 60–63, 65–70  
 and virtue, 53, 64, 250
- right of humanity in our own person, 6, 86–87, 178–81, 183–89, 237
- right of necessity, 120
- rights, *see also* human rights  
 collective, 7, 193, 205–09  
 contract, 93–94, 100–01, 104–05, 187–88  
 the “core” of, 204  
 as correlative with duties, 7, 78, 82, 202–09  
 domestic, 93–94, 100–01, 104–05  
 equivocal, 120  
 and freedom, 173  
 inalienable, 78  
 innate, 74, 78, 79, 80, 87–90
- property, 4–5, 78, 90–91, 93–94, 100–01, 104–05  
 provisional/peremptory, 101–03  
 to punish, 115, 117, 118
- Ripstein, Arthur, 62
- Ritter, Christian, 10
- Ross, W.D., 240, 241
- Rousseau, Jean-Jacques, 68
- Scanlon, Thomas M., 246
- Scheid, Don E., 113
- Schleiermacher, Friedrich, 9
- Schmucker, Josef, 10
- Schopenhauer, Arthur, 9, 10, 211
- Schweitzer, Albert, 211
- self-conception, 45–50  
 as agent/as knower, 45, 47–49
- self-consciousness, 31, 34, 44–46, 216
- self-constraint  
 and obligation, 172–73, 180–81, 184, 185, 186  
 and perfect duties to oneself, 174, 179, 180–81  
 and virtue, 58, 160–61, 164–65, 176
- self-deception, 162, 165, 177, 247
- self-esteem *see* self-respect
- self-knowledge, 32–33, 48, 177, 248
- self-legislation  
 and freedom, 172  
 as ground of morality, 28, 51, 52, 53–54, 57–59, 65, 184, 185
- self-recognition, 4, 86
- self-respect, 149–50  
 sense, 35
- Sensen, Oliver, 171
- sensibility, 34–35
- Shaftesbury, Anthony Ashley Cooper, third  
 Earl of, 133
- Sherman, Nancy, 147
- Sidgwick, Henry, 242
- Singer, Peter, 211
- slavery, 187–88
- Smith, Adam, 144
- sociability, 82
- social contract theory, 246
- souls, 215  
 rational, 219, 220
- speciesism, 74
- Spinoza, Baruch, 241
- Stark, Werner, 18
- state  
 authorities of, 117–18  
 duty to enter, 101–03  
 idea of, 117–18  
 possibility of, 102  
 purpose of, 5, 94, 117–20
- state of nature, 78, 79, 81, 92, 94, 101–03, 118–19

- Stoic dictum, 176  
 Stoic ethics, 29, 146  
 supreme principle of the doctrine of morals, 237–38  
 supreme principle of the doctrine of virtue, 176, 238, 242, 250  
 syllogism, 30, 33  
 sympathy, 146–49, 197  
   and duties regarding animals, 223–27
- thing  
   treating oneself as a, 174, 177, 186–87, 189–90  
 things, 172  
   and external objects of choice, 93, 100–01  
   and original acquisition, 104  
 Thomasius, Christian, 103–04  
 Timmermann, Jens, 170, 184, 195, 211, 225, 226–27, 230  
 transcendental idealism  
   and moral motivation, 131–32, 135, 138, 151  
 true natural needs, 40
- Ulpian formulae, 85  
 understanding, 30, 35, 39, 40–41, 43, 44, 74  
 “Universal Declaration of Human Rights”, 88, 89  
 universal law of Right, 61–63, 74, 76, 80, 97  
 universal principle of Right, 56, 61, 79–80, 238  
 Unna, Yvonne, 181  
 utilitarianism, 242, 243–44
- value, 198–99, 231, 238, 239  
 Vasilyev, Vadim, 211  
 vice, 156–60, 167, 169  
   as evil power of choice, 157  
   and freedom, 157–58, 161–65  
   and passions, 157–58  
   and self-deception, 162, 165  
 vices  
   opposed to duties to oneself, 177, 178, 183
- “Virginia Bill of Rights”, 88  
 virtue, 5–6, 22, 152–69  
   Aristotelian/Kantian, 6, 153, 165–69, 252–53  
   conditions of, 176–77, 178, 182–83  
   and contemplation, 164–66  
   enemy of, 5–6, 152–65  
     inclination as, 152–53, 155–58, 161–62  
     propensity to evil as, 152–53, 155–60, 167, 169  
   as inner freedom, 6, 153, 160–69, 176  
   practical point of view and, 153–58  
   as self-constraint, 160–61, 164–65, 176  
   as strength, 6, 153, 156, 160–69, 248–49  
 virtue ethics, 242, 243  
 Vorländer, Karl, 17
- Walch, Johann Georg, 103  
 Warren, Mary Anne, 210, 211  
 weakness, 247, 248–49  
 Wellman, Carl, 204  
 Westphal, Kenneth R., 100  
 will, 2–3, 35, 38, 39, 43–50, *see also* general will  
   as faculty of desire in accordance with  
     concepts, 43–46  
   as practical reason, 28, 48–50, 58–60  
   primacy over power of choice, 49–50  
   private, 64  
   public, 59, 62, 63, 68, 69, 70  
   public/private, 60  
   unilateral, 59, 62–63, 68–69  
 Willaschek, Marcus, 56, 61, 64  
 Wille, 48, *see also* will  
   juridical conception of, 3, 59–60, 64  
 Wille/Willkür, 3, 58–60, 64  
 willing  
   ethical/juridical, 69–70  
 Willkür, 46, *see also* choice, power of  
 wish, 38, 46–47, 49, 98–99, 101  
 Wolffian ethics, 10  
 Wood, Allen W., 11, 24, 27, 53, 56, 195, 198, 200–01, 211, 227, 228–29, 231, 234