THE UNIVERSITY OF CHICAGO

FORM AND THERAPY:

KANT ON THE PURPOSE AND LIMITS OF MORAL PHILOSOPHY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHILOSOPHY

BY

LAURENZ RAMSAUER

CHICAGO, ILLINOIS

AUGUST 2024

Table of Contents

Preface

When I began working on this project in the summer of 2020, the aim had been to get clear about the notions of form and matter in Kant's practical philosophy. Through the process of grappling with Kant's hylomorphism in practical reason, the project quickly turned into a delineation of the purpose and limits of moral philosophy. In the end, this book has become what I believe to be a Kantian reading of Kant's ethics. On a very high level of abstraction, Kant's original insight was to develop a philosophical system out of the recognition that it is the forms of our cognition that make our representations of objects and ends possible. And at the same time as they make representations possible, it is also those same forms that limit what can be expressed and thought. Consequently, trying to get a better sense of form and matter in Kant's practical philosophy was trying to get a better sense of the purpose and limits of Kantian moral philosophy. More importantly, however, I believe that appreciating the limits of Kant's moral philosophy also allows one to better appreciate what is genuinely valuable about it.

Despite – or perhaps precisely because of – being a slow partner in philosophical conversation, I believe that philosophy is only possible as a time-consuming, collaborative activity. Consequently, there is nothing in the following pages that I have not learned, in one way or another, from someone else. First and foremost, I owe an immense debt of gratitude to Candace Vogler. It is no exaggeration to say that much of this book is merely an attempt to elaborate, with all the footnotes required for a dissertation, a reading of Kant to which Candace has already introduced her students at the University of Chicago for decades. Without Candace's unwavering support and confidence, this book would never have been written. I owe equally much to Matthias Haase, whose ability to do philosophy in conversation and tease out insights from his interlocutors is truly

iv

List of Abbreviations

Unless otherwise stated, citations appear in the order of abbreviation, volume number, and page number from the Akademie Ausgabe (AA), *Kant's gesammelte Schriften*, edited by the Königlich Preußischen Akademie der Wissenschaften (vols. 1–16), the Preußischen Akademie der Wissenschaften (vols. 17–22), the Deutschen Akademie der Wissenschaften zu Berlin and/or the Akademie der Wissenschaften zu Göttingen (vols. 23–25 and 27–29), and the Berlin-Brandenburgischen Akademie der Wissenschaften (vol. 26), Berlin: Reimer, now: De Gruyter 1900 ff. The exception are references to the *Critique of Pure Reason*, which cite the page numbers of the first (A) and second (B) editions.

Unless otherwise stated, all translations come from The Cambridge Edition of the Works of Immanuel Kant, edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1992–).

Abbreviations of Kant's works

Anth   *Anthropologie in pragmatischer Hinsicht* (*Anthropology from a Pragmatic Point of View*), AA 07

Br   *Briefe* (*Letters*), AA 10-13

EEKU   *Erste Einleitung in die Kritik der Urteilskraft* (*First Introduction to the Critique of the Power of Judgment*), AA 20

GMS   *Grundlegung zur Metaphysik der Sitten* (*Groundwork of the Metaphysics of Morals*), AA 04

GSE   *Beobachtungen über das Gefühl des Schönen und Erhabenen* (*Observations on the Feeling of the Beautiful and Sublime*), AA 02

HN   *Handschriftlicher Nachlass*,* AA 14-23

KrV   *Kritik der reinen Vernunft* (*Critique of Pure Reason*)

KpV   *Kritik der praktischen Vernunft* (*Critique of Practical Reason*), AA 05

KU   *Kritik der Urteilskraft* (*Critique of the Power of Judgment*), AA 05

Log   *Logik* (*Jäsche logic*), AA 09

MAM   *Muthmaßlicher Anfang der Menschengeschichte* (*Conjectural beginning of human history*), AA 08

MAN   *Metaphysische Anfangsgründe der Naturwissenschaft* (*Metaphysical Foundations of Natural Science*), AA 04

---

* A selection of Kant's handwritten remains is included in the 2005 volume *Notes and Fragments* of the Cambridge Edition of the Works of Immanuel Kant. Where available, English translations of Kant's notes are cited after the Cambridge edition. Notes not contained in the CUP edition I translate myself, as indicated in the text. Since the CUP *Notes and Fragments* volume does not follow the Academy Edition for Kant's notes on the *Observations*, I quote Kant's handwritten notes on the *Observations* with both the AA number as well as the page number of the Cambridge edition *Notes and Fragments* and the page number of Rischmüller's 1991 edition (which was used as the source of the Cambridge edition's translation).

MS   *Die Metaphysik der Sitten* (*The Metaphysics of Morals*), AA 06

MSI   *De mundi sensibilis atque intelligibilis forma et principiis* (*Inaugural Dissertation*), AA 02

Päd   *Pädagogik* (*Lectures on pedagogy*), AA 09

PG   *Physische Geographie* (*Physical Geography*), AA 09

Prol   *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können* (*Prolegomena to any future metaphysics*), AA 04

Refl   *Reflexionen*, AA 14-19

RezHerder *Recensionen von J. G. Herders Ideen zur Philosophie der Geschichte der Menschheit* (*Review of J. G. Herder's Ideas for the philosophy of the history of humanity. Parts 1 and 2*) AA 08

RGV   *Religion innerhalb der Grenzen der bloßen Vernunft* (*Religion within the boundaries or mere reason*), AA 06

SF   *Der Streit der Fakultäten* (*The Conflict of the Faculties*), AA 07

TP   *Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis* (*On the common saying: That may be correct in theory, but it is of no use in practice*), AA 08

ÜGTP   *Über den Gebrauch teleologischer Principien in der Philosophie* (*On the use of teleological principles in philosophy*), AA 08

VAProl *Vorarbeit zu den Prolegomena zu einer jeden künftigen Metaphysik*, AA 23

VAZef   *Vorarbeiten zu Zum ewigen Frieden*, AA 23

V-Anth/Mensch *Vorlesungen Wintersemester 1781/82 Menschenkunde* (*Lectures on Anthropology, Menschenkunde*), AA 25

V-Lo/Blomberg *Logik Blomberg* (*Blomberg logic*), AA 24

V-Lo/Dohna *Vorlesungen Sommersemester 1792 Logik Dohna-Wundlacken* (*The Dohna-Wundlacken logic*), AA 24

V-Lo/Wiener *Wiener Logik* (*Vienna logic*), AA 24

V-MS/Vigil *Vorlesungen Wintersemester 1793/1794 Die Metaphysik der Sitten Vigilantius* (*Kant on the metaphysics of morals: Vigilantius's lectnre notes*), AA 27

V-NR/Feyerabend *Naturrecht Feyerabend (Winter 1784)*, AA 27

VRML   *Über ein vermeintes Recht, aus Menschenliebe zu lügen* (*On a supposed right to lie from philanthropy*), AA 08

WA   *Beantwortung der Frage: Was ist Aufklärung?* (*An answer to the question: What is enlightenment?*), AA08

WDO   *Was heißt: Sich im Denken orientiren?* (*What does it mean to orient oneself in thinking?*), AA 08

ZeF   *Zum ewigen Frieden* (*Toward perpetual peace*), AA 08

Abbreviations of Hegel's Works

Unless otherwise stated, all German citations come from the Suhrkamp edition of Hegel's collected works in 20 volumes (Berlin: Suhrkamp, 1986–)

E   *Enzyklopädie der philosophischen Wissenschaften* (*Encyclopedia of the Philosophical Sciences*)

JS   *Jenaer Schriften* (*Jena Writings*)

PG   *Phänomenologie des Geistes.*

        The Phenomenology of Spirit, translated and edited by Terry Pinkard, Cambridge: Cambridge University Press, 2018.

R   *Grundlinien der Philosophie des Rechts. Oder: Naturrecht und Staatswissenschaft im Grundrisse.*

        Elements of the Philosophy of Right, translated by H. B. Nisbet, edited by Allen Wood, Cambridge: Cambridge University Press, 1991.

Introduction

> *Kant's Joke* – Kant wanted to prove, in a way that would dumbfound the whole world, that
> the whole world was right: that was the secret joke of this soul. He wrote against the scholars
> in favor of popular prejudice, but for scholars and not for the people.
>
> -- Nietzsche, *The Gay Science*, §193

This book is a study of Kant's practically oriented view of moral philosophy. The conception of moral philosophy ascribed to Kant in the following chapters embodies an approach that represents a minority view of ethics today, but one that in its broad outlines arguably can be found throughout philosophers from the ancient to the early modern period. The interest of Kant's own ethics lies in his refinement and rigorous presentation of this traditional approach to ethics – one that deserves to be taken seriously in its own right.

While I do not share Nietzsche's sentiment expressed in his brief remark on Kant in *The Gay Science*, an ironic aspect of studying Kant's practical philosophy is that Nietzsche may well have appreciated a core aspect of Kant's ethics more than many self-styled Kantians and Neo-Kantians today. Explaining how this ironic turn of events was possible involves calling to mind three things: a contemporary assumption, an old philosophical tradition, and a Kantian insight.

One way of characterizing ethics at large is as that subject, or that part of human reflection, concerned with the question how one should live. A virtue of this way of characterizing the subject of ethics is that it immediately foregrounds the practical relevance of moral philosophy. However, there is a way of understanding the question that overly settles the type of inquiry. That is because one way to hear the question is as limiting the possibilities of what moral philosophy is about. On

this way of hearing the question, ethics is primarily about figuring out how one should live, i.e., figuring out what is truly right and wrong. By putting the question in this way, we assume a specific purpose of ethical as well as philosophical reflection: to determine the ends that we should aim for. Henry Sidgwick introduced one of the most influential works of anglophone moral philosophy in just this way:

> The boundaries of the study called Ethics are variously and often vaguely conceived: but they will perhaps be sufficiently defined, at the outset, for the purposes of the present treatise, if a 'Method of Ethics' is explained to mean any rational procedure by which we determine what individual human beings 'ought' – or what it is 'right' for them – to do, or to seek to realise by voluntary action.[1]

Important as this aspect of asking the question 'how should one live?' is, it also risks foreclosing other ways of ethical inquiry. By settling the purpose of ethical reflection as the search for the ends we should aim for, this way of asking the question also settles the purpose of moral philosophy as the search for a 'method of ethics.' As long as we do not presume that ethical and philosophical reflection is exhausted by the search and application of such a method respectively, the search for a method of ethics is an entirely reasonable project. However, much contemporary philosophy asks this question in a way that forecloses not merely interest in other ways of asking the question, but in a way that makes it difficult to appreciate that not all authors at all times in the field of ethics were engaging in the same search for a method of ethics. Much of this book will be arguing that this presumption about the purpose of moral philosophy makes it difficult to appreciate what is singularly insightful about Kant's ethics, and also that this presumption leads, when applied to authors who do not share the same preoccupation with a search for a method, to avoidable exegetical difficulties.

---

[1] Sidgwick 1907 p. 1

In contemporary scholarship on Kant's ethics, this assumption about the purpose of ethical and philosophical reflection – which I will, for the purpose of this book, call Sidgwickean[2] – has become almost ubiquitous. An early expression of this Sidgwickean reading of Kant's ethics has been provided by John Rawls's lectures on the history of moral philosophy. There, he coined the term 'decision-procedure,' which I will continue to refer to throughout this book. In these lectures, Rawls suggested that an important reason for studying Kant's 'CI-procedure' is that 'one is to use it as a way of generating the content – the first principles along with the essential rights, duties, permissions, and the rest – of a reasonable moral doctrine.'[3]

According to Rawls's doctoral student, Christine Korsgaard, Kant's 'approach is to raise practical questions as they are faced by the reflective moral agent herself. Moral philosophy is the extension and refinement of ordinary practical deliberation, the search for practical reasons.'[4] On her reading of Kant's *Groundwork*, '[t]he argument of Groundwork I is an attempt to give what I shall call a "motivational analysis" of the concept of a right action, in order to discover what that concept applies to, that is, which actions are right. […] Kant's achievement is to argue from this feature of right actions to a substantive moral principle which identifies which actions are right.'[5] And again, in her introduction to the Cambridge University Press's edition of Kant's *Groundwork*, Korsgaard writes: 'when we know what makes actions morally good, we will be able to determine which actions are morally good, and so to determine what the moral law tells us to do. This is what Kant means when he says he is going to "explicate the concept of a good will" (AK 4:397): that he is going to

---

[2] This is not meant as a genealogical claim, although I suspect that the most influential readers of Kant sharing this Sidgwickean assumption received a philosophical education (especially at Oxford and Harvard) in which Sidgwick's moral philosophy figured prominently.
[3] Rawls 2000 p. 163
[4] Korsgaard 1996a p. xii
[5] Korsgaard 1996a p. 47

find out what principle the person of good will acts on, in order to determine what the moral law tells us to do.'[6] Similarly, David Velleman has characterized Kant's ethics as centered around the idea 'that we can figure out what morality requires by analyzing the very idea of being morally required to do something.'[7]

Perhaps the most explicit way of giving a somewhat Sidgwickean reading of Kant's ethics can be found in Jerome Schneewind's work. According to Schneewind, 'Kant presented a method for discovering what morality requires us to do in any situation and claimed that it is a method everyone can use. The method consists in testing one's maxim against the requirement stated in the formulations of the categorical imperative. There has been endless discussion of the adequacy of Kant's method in giving moral guidance…'[8] One of the most recent – and equally explicit – expressions of this Sidgwickean reading has been provided by Martin Sticker, who has claimed that 'Kant describes pre-theoretical moral deliberation explicitly as an agent asking herself questions pertaining to the universalizability of her intended course of action, i.e. as an active use of rational capacities. […] Common agents in their everyday moral evaluations make use of a common universalization test which involves checking if one can will a maxim as universal law.'[9]

That Kant's texts are difficult to square with such a reading has not gone unnoticed. But the result can seem more of a motte-and-bailey game than a genuine reorientation of our reading of Kant's ethics. Thus, when pressed on the readings just quoted, Kant scholars sometimes retreat to less ambitious claims – which, however, still tend to presume the Sidgwickean conception of what moral philosophy is for. For instance, Kant scholars have drawn a distinction between the idea of

---

[6] Korsgaard 1998 pp. xii
[7] Velleman 2012 p. 343
[8] Schneewind 1991 p. 289
[9] Sticker 2015 p. 980

'deriving' or 'deducing' material maxims from the moral law and 'appraising' actual maxims in light of the moral law. As Paul Dietrichson has put it, 'what Kant is saying is, not that we can *deduce* material maxims from the formal principle called the moral law, but that we can, and should, *appraise* our actual material maxims and our contemplated material maxims in terms of that formal principle. There is a big difference between deducing and appraising.'[10] Similarly, Kant scholars can draw distinctions between 'algorithms,' which they reject, and 'derivations,' which they endorse. For instance, Kyla Ebels-Duggan's gloss on Kantian ethics emphasizes that '[n]o one supposes that the Categorical Imperative provides a mechanical algorithm that delivers all by itself a complete account of what we ought to do in any situation. But if it can be shown to have some attractive content, for instance, to rule out some plausibly immoral ways of acting, then the task of driving moral substance from the form of law would be accomplished.'[11]

Despite such occasional attempts to tone down the Sidgwickean project ascribed to Kant's ethics, the notion that Kant was in search of a method of ethics remains uncontroversial enough to form the background for much Kant scholarship. Thus, scholars continue to mention – and leading journals publish – passing remarks plainly presupposing the Sidgwickean project. For instance, a recent article on causality in Kant's *Groundwork* introduces its topic through the claim that 'Kant's formula of universal law (FUL) is standardly understood as a test of the moral permissibility of an agent's maxim: maxims that pass the test are morally neutral, and so permissible, while those that do not are morally impermissible.'[12] Moreover, even if one were to tone down the supposed Sidgwickean project ascribed to Kant's ethics – whether in terms of deriving rather than algorithms, or appraising rather than deducing – we would thereby still effectively foreclose one way in which

---

[10] Dietrichson 1964 p. 147, original emphasis.
[11] Ebels-Duggan 2011 p. 176
[12] Wolf 2023

ethical and philosophical reflection can matter to how we should live. Importantly too, as I will argue in this book, we will remain stuck with misleading assumptions that complicate our appreciation of Kant's texts.

Another possible purpose of moral philosophy is that moral philosophy ought to be medicine for the soul, to help free us from our own forms of self-entrapment. This understanding of moral philosophy has a venerable pedigree too. Its most famous proponents are not Victorian scholars but the big schools of Hellenistic philosophy.[13] 'Empty is that philosopher's argument', Epicurus tells us in a fragment, 'by which no human suffering is therapeutically treated. For just as there is no use in a medical art that does not cast out the sickness of bodies, so too there is no use in philosophy, unless it casts out the suffering of the soul.'[14] On this conception of what moral philosophy is ultimately about, we may be less in need of a method for figuring out what is truly right and wrong and instead require resources for a therapeutic way of reflection: a reflection that can help us get closer to the virtues appropriate to beings like us; a kind of reflection that can help us overcome our self-incurred forms of unenlightenment or self-entrapment, and bring us closer to our true nature as finite, rational beings.

Nothing in this book relies on genealogical claims about the Hellenistic philosophers' influence on Kant. Indeed, I believe that Kant's conception of the practical purpose of moral philosophy is adequately characterized as therapeutic completely regardless of any older tradition of moral philosophy as therapy, and would be adequately so characterized even if Kant had never read a single line of either Stoic or Epicurean thought. However, it is worth pointing out that Kant is not alone in this regard. If anyone, it us today who are being overly parochial by posing the question of

---

[13] For a detailed study of the therapeutic aspects of Hellenistic ethics, see especially Nussbaum 2018.
[14] Us. 211, Porphyrius ad Marcellam 31 p. 209, 23 Nauck. Translation by Martha Nussbaum (2018 p. 102).

ethics in a way that effectively precludes this practically oriented view of moral philosophy from forming part of our philosophical and exegetical repertoire – a practically oriented view of moral philosophy that, for much of western history, has had as much (if not more) influence on intellectual and cultural developments as the works of Aristotle and Plato.

Incidentally, Kant owed big intellectual debts to Stoic and Epicurean thought. We know as much from his work itself,[15] from his biographers,[16] and we know that his personal library included at least works of Cicero and Seneca.[17] The influence of Stoic cosmopolitan thought on Kant is well documented,[18] and there is ongoing discussion of the similarities and dissimilarities between Kantian and Stoic ethics in general.[19] It is thus safe to presume that Kant was familiar with at least some parts of Hellenistic ethical thought. Consequently, we should be open, in reading Kant's ethics, to the possibility that his way of asking the questions 'how should we live?' and 'what should I do?' was not meant to limit ethical inquiry to what has become the comfort zone of western moral philosophy since Sidgwick, namely the search for a method of ethics. Instead, on a closer look, Kant's moral philosophy may turn out to have much more in common with the therapeutic aims of the Hellenistic schools: philosophy that sees its practical purpose in providing a medicine for the soul, and intends to bring us closer to the virtues appropriate to the finite, dependent and rational beings like us. On the reading I present in this book, the unique insight of Kant's moral philosophy lies in his refinement and rigorous presentation of this therapeutic conception of moral philosophy.

---

[15] A concise summary of Cicero's influence on Kant's ethical writings can be found in Kuehn 2001b.
[16] See especially Kuehn 2001a.
[17] Warda 1922
[18] See Nussbaum 1997
[19] Schneewind 2010 (chapter 15), Annas 1993.

What makes Kant's refinement of the therapeutic approach to ethics stand out from the older proponents of this tradition is the way he develops his own therapeutic ethics out of his specific conception of form and matter – that is, his specific, hylomorphic conception of practical reason. In his *Critique of Pure Reason,* Kant glosses form as something determining (*die Bestimmung*) as opposed to the determinable or determined (*Bestimmbaren*), and provides two ways of characterizing form: as that which makes the nature or quality (*Beschaffenheit*) of things and makes for their 'specific difference;' and as the structural determination of things by their relation. Importantly, Kant makes clear that 'form' is something contributed by reason, not something to be posited in things in themselves. Thus, general logic is the mere 'form' of thinking, where we abstract from the objects of thought (KrV A54/B78). Similarly, space, time and the categories are 'form' whereas the manifold is matter which are accessible to us because of the form: "I call that in the appearance which corresponds to sensation its matter, but that which allows the manifold of appearances to be intuited as ordered in certain relations I call the form of appearance" (KrV A20/B34). In Kant's practical philosophy, this specific understanding of form and matter resurfaces already in his discussion of the form of pure practical reason in its formula of a universal law of nature in the *Groundwork*. There, Kant is concerned with the Categorical Imperative as that which determines (i.e., makes possible through its form) pure practical willing:

> Since the universality of law in accordance with which effects take place constitutes what is
> properly called nature in the most general sense (as regards its form) - that is, the existence
> of things insofar as it is determined in accordance with universal laws - the universal
> imperative of duty can also go as follows: act as if the maxim of your action were to become
> by your will a universal law of nature. (GMS 4:421)[20]

---

[20] See also GMS 4:437

Intuitively, thinking about the Categorical Imperative as the form of practical knowledge may suggest that this form, the Categorical Imperative, could be used as a method of ethics. We may, for instance, think that we can determine what makes an oak tree deficient by comparison with its 'form,' i.e., with what it is for the life of this kind of tree to go well. Similarly, we might take the Categorical Imperative to serve as a method of ethics in the sense that we know which intentional actions are wrong and which ones are right by comparing their maxims with their 'form,' with the Categorical Imperative. However, there is something misleading about this comparison. In the example of the oak tree, the 'form' (what it is for the oak tree to flourish) is supposedly compared with content (the actual state of the tree before us), which fails to live up to its form. By contrast, in the case of the maxim and the Categorical Imperative, we would not compare form with content in this way; we compare the form (of moral judgment) with another form (of the respective maxim). On Kant's view, the formulas of the Categorical Imperative do not contradict the content of actions; they contract the *form* of maxims. And this form of our maxims is determined by us, by our will.

Kant's particular hylomorphism limits the possible uses the form of moral judgment – and its formulaic expressions in philosophy – can be put to. For the kind of thing that would be 'tested' (our maxim, i.e., our representation of an action as to be done) does not exist out there independently of our will. Our way of knowing the form of our maxim is not anything like our way of knowing the actual state of some particular oak tree in front of us. That is because at the center of Kant's ethics is an account of practical self-consciousness. When we will something we are necessarily aware of what we are willing – to will something requires us to have a representation of an action or end as to be achieved, and that entails that we are conscious of this representation. If things go well, such conscious representations contain a sense of transparency: it is transparent to us whether we represent the action or end in question as intrinsically good or whether we represent it

9

as an object to be brought about for the sake of something else.[21] The structure, or form, of such

representations of actions or ends as intrinsically good is the Categorical Imperative. Thus, while we

may not be familiar with the philosophical terminology capturing that form – and while our

philosophical terminology may be mistaken – we must be familiar with whether or not our

representation of an action or end is an instance of representing something as intrinsically good. On

this picture, there is no practical use for any philosophical expression of the Categorical Imperative

in telling us whether or not we actually represent anything as intrinsically good, because this form is

always already implicit in our reflective awareness in practical reason – either by constituting the

form of our actual representation of an end as intrinsically good, or in our implicit awareness of our

not representing such a moral object because we are representing it as merely good instrumentally.

Whatever 'test' the Categorical Imperative would be put to in figuring out what is truly right and

wrong, it would come too late.

However, let me come back to the caveat: *if things go well.* There is another option – more

proto-Freudian than Sidgwickian – on which making explicit the form of the good will can indeed

be practically relevant. This option is not the puzzled agent, who just wants to know what is truly

right and wrong; it is the struggling agent, who has managed to rationalize away what would

otherwise have been transparent to them, or even what actually *was* transparent to them before they

started to rationalize. This agent indeed has a practical use for moral philosophy. Not as teaching

them something new, but as bringing them back to who they really are. For Kant, we are all this

---

[21] NB that on Kant's view it is *not* always transparent to us whether in acting in accordance with such a representation –
that is, when we bring about the action or end we have judged to be intrinsically good – we have acted *from* this
representation, or whether our motivation was secretly selfish. As Kant's famous example of the shopkeeper in the
Groundwork emphasizes, there can be actions that are both intrinsically and instrumentally good. In such cases, we may
represent the action as the right thing to do but may be unconsciously motivated to perform the action because of the
foreseen benefit rather than because of its intrinsic goodness. For discussion of Kant's view of moral self-opacity, see
Berg 2020.

agent. And according to Kant's view of moral philosophy, ethical reflection orients us towards achieving our true identity as finite, rational beings.

The frequent desire of contemporary philosophy to find or defend a method of ethics is entirely reasonable. As I have argued elsewhere (and emphasize again in chapter 8), we have good reasons to be ambitious in moral philosophy – more ambitious than I believe Kant ever was.[22] However, this desire can also distract from what is singularly valuable about Kant's ethics, namely his view of moral philosophy as helping us overcome specific forms of self-entrapment. Moreover, the tendency to read Kant as focused on defending a method of ethics also makes it difficult to appreciate Kant's specific hylomorphism of practical reason, leading to numerous (and in my view, mostly avoidable) exegetical puzzles. The ironic result of this skewed focus in our reception of Kant is that Nietzsche's joke arguably captures both the purpose and the limits of Kant's moral philosophy better than many Neo-Kantian ethics today.

Chapter outline

This book is divided into three parts. Part I gives an account of the practical purpose of Kant's moral philosophy. The first chapter, 'Conceptions of Moral Knowledge, and the Practical Purpose of Moral Philosophy,' argues that Kant's distinct conception of practical knowledge corresponds to a specific conception of the practical purpose of moral philosophy. Kant believed that common moral knowledge is distributed equally among all rational beings, that it is immediately present without need for complicated reflection, and that the task of philosophy is to excavate the principles of practical reason that underlie common moral cognition. However, Kant also believed

---

[22] Ramsauer 2023

that this immediate moral cognition can be clouded by our tendency to rationalize against the requirements of morality in favor of out inclinations. The practical purpose of moral philosophy is a kind of therapy: to help us overcome this rationalizing tendency by making the underlying principles of our practical reason explicit.

In 'The Role of Reflection in Ethical Life,' I discuss in more detail how Kant's practical therapy works, and what role the resources we get from moral philosophy can play in practical deliberation. The chapter starts with a seeming exegetical puzzle: on the one hand, Kant describes moral cognition as immediate; on the other hand, Kant says that all judgment requires reflection. I argue that conscious reflection on the source of our moral cognition provides the therapy of unclouding moral cognition that has been clouded by our tendency to rationalize away the requirements of morality.

Chapter three, 'The Formalism of the Categorical Imperative,' turns to the flipside of the account sketched in the earlier chapters: what Kant's practical therapy cannot do, namely provide the test or decision-procedure sought by many modern readers. In this chapter, I argue that Hegel's famous empty-formalism charge is best understood as a criticism of the (comparatively) limited resources of Kant's practical therapy. Through an exegesis of Hegel's criticism, this chapter also shows how the limits of Kant's moral philosophy go hand in hand with his insights.

Part II of this book takes a closer look at Kant's project of a metaphysics of morals from the early *Groundwork* to the late *Doctrine of Virtue*. In this part, I turn to a series of exegetical questions that can be resolved in light of the relation, discussed in Part I, between Kant's views about the practical purpose of moral philosophy and his hylomorphism about practical reason. Implicitly, part

II also provides a partial defense of Kant's practical therapy against the Hegelian charge by showing its comparatively modest aims.

Chapter four, 'Kant's Derivation of Imperatives of Duty,' shows that Kant's 'transition to a metaphysics of morals' in the *Groundwork* was less ambitious than many readers have assumed. I argue that Kant did not intend the infamous 'derivation of imperatives of duty' in section II of the *Groundwork* as an example of a test or decision-procedure in ethics. Instead, the main aim of Kant's 'derivation of imperatives of duty' was to show how his analysis of the form of moral judgment is presupposed in all four types of moral imperatives that philosophers of his time recognized. This is supported by an exegesis of the first two sections of the *Groundwork*, its historical-philosophical context, as well as Kant's response to an often-misidentified (and still untranslated) review of the *Groundwork* by G. A. Tittel.

Chapter five, 'A Sudden Change of Mind?' turns to Kant's idea of a 'metaphysics of morals.' I argue that the reading of Kant developed in the previous chapters allows us to see Kant's project of a 'metaphysics of morals' as consistent throughout his critical period. While many readers have assumed that Kant's *Groundwork* is trying to be 'pure' in a way that the later *Metaphysics of Morals* is not, I argue that the attribution of such a break to Kant's project relies on a skewed understanding of the purpose of Kant's moral philosophy.

In chapter six, 'A System of Moral Knowledge,' I argue in detail against the common view that Kant's late *Metaphysics of Morals* would be an attempt to systematically derive concrete duties from a principle of morality. Instead, the *Metaphysics of Morals* attempts to unify an already accessible heap of moral cognition into a systematic body of knowledge by help of the a priori form of

13

practical reason. This chapter concludes the defense of Kant's approach to moral philosophy against the Hegelian criticism by pointing to Kant's relatively limited ambitions in his late work.

Chapter seven, 'Kant's Casuistical Questions,' brings the focus of the book back to the therapeutic aspect of Kant's moral philosophy. Many standard interpretations of Kant's *Metaphysics of Morals* run into exegetical difficulties regarding Kant's sections on 'casuistical questions.' In chapter seven, I argue that the therapeutic purpose of Kant's moral philosophy helps us understand the casuistical questions as highlighting the complexity of moral life. In order to provide a kind of reflection that brings us closer to our true nature as rational beings, Kantian moral philosophy must be sensitive to the specifically human difficulties we face in our ethical life.

Part III of this book returns to the limits of Kantian therapy and asks what reasons we might have to remain unsatisfied with the specific ambitions of Kant's moral philosophy. Chapter eight, 'Kant's racism as a philosophical problem,' articulates a formalism worry about Kant's moral egalitarianism. In contrast to most contemporary readers, I argue that it is not obvious how Kant's racism would be inconsistent with his highly abstract universalism; and consequently, that Kantian therapy cannot provide the antidote to his own racism. Therefore, we might want to be considerably more ambitious, and demand more than the kind of therapy Kant's ethics provides.

Finally, the concluding chapter turns the Kantian approach to ethics onto itself. Kant's moral vocabulary has become part of the cultural repertoire through which we make sense of our ethical and political lives. Consequently, to follow Kantian ethics without any appreciation of its limits would be to turn Kant's practical therapy into a collective neurosis: a pathological way of making incomplete sense of our ethical lives. Against such collective neuroses too, moral philosophy ought to be therapy.

PART I

KANT'S PRACTICAL THERAPY

Chapter 1

Conceptions of Moral Knowledge
and the Practical Purpose of Moral Philosophy

Introduction

Kant believed that the form of practical reason constitutes the fundamental principle of

morality and that this principle is thus implicit in the practical knowledge of every rational being.

And not only did Kant believe that the fundamental principle of morality is already within us all

along, he also explicitly claimed that ordinary agents have no need for any kind of philosophy in

order to know what is right and wrong. Consequently, one might wonder what – if anything – could

possibly be the practical point of Kant's moral philosophy. After all, if we share Kant's view about

the nature and distribution of moral knowledge, it might seem as if moral philosophy is just a

speculative exercise with little to no practical relevance. At times, Kant himself appeared skeptical

about the practical purpose of philosophy. In an emotional passage from his notes for the

Observations on the Feeling of the Beautiful and Sublime, Kant tells us of the near embarrassment he felt

when comparing the purposefulness of his own vocation as a scholar with that of the common

laborer.[1] And yet, Kant also claims that moral philosophy is not merely an idle, theoretical exercise.

Quite to the contrary, practical knowledge needs moral philosophy "in order to provide access and

durability for its precepts." (GMS 4:405).

What, then, is the practical purpose of Kant's moral philosophy? Although few readers and

commentators have troubled themselves to find an explicit answer to this question, the literature on

---

[1] HN 20:44 and 20:102

Kant's ethics provides two intuitive suggestions: on one popular view, the point of Kant's ethics might be to help us figure out what is right and wrong, by giving us a decision-procedure that we can follow in our normal life or apply in moral dilemmas.[2] On another possible view, the practical purpose of Kant's ethics might be to respond to skeptical worries: it might aim to refute the theoretical skeptics who deny our freedom to act (and thus morality) or the pragmatic skeptics who doubt that there is any reason to be moral.[3] Unfortunately, both these views are prima facie unappealing. If Kant really believed that the fundamental principle of morality is simply the form of practical reason and thus within us, and also that ordinary agents have no need for any kind of philosophy in order to know what is right and wrong, then the *practical* purpose of Kant's ethics could hardly be to help us figure out what morality truly requires. And if Kant's moral philosophy is supposed to become practical only in difficult moral dilemmas, then a further exegetical argument would be needed to explain what these difficulties are and how exactly moral philosophy could help. Perhaps, then, the practical purpose of Kant's moral philosophy lies in dispelling skeptical worries? On the face of it, this option seems equally unattractive. If our concern was a theoretical skeptic who denies the possibility of freedom, the worry is supposedly resolved by the *Critique of Pure Reason*, not by Kant's moral philosophy. If our concern was a pragmatic skeptic who denies any reason to be moral, Kant's moral philosophy cannot be the answer either, for he repeatedly points out that the bindingness of the moral law cannot be proven theoretically.[4] Moreover, Kant famously argues that

---

[2] See for instance Schneewind 1991 p. 289, Korsgaard 1996a pp. xii and 47, Korsgaard 1998 p. x-xii, Rawls 2000 p. 163 and Velleman 2012 p. 343.

[3] See for instance Korsgaard 1998 p. x. In contemporary terms, the type of skepticism presumed on this view usually appears to be either skepticism about moral truth or practical moral skepticism. An exception here is the 'misologist,' a type of skeptic not usually included in contemporary discussions of moral skepticism (compare Sinnott-Armstrong 2019). The 'misologist' started out trusting that following the commands of reason would lead to happiness, and is so disappointed by reason not leading to happiness, that they conclude that humans' proper end must be attained through non-rational faculties. For a discussion of Kant's response to the misologist, see Callanan (2019).

[4] More specifically, there cannot be a further justification of morality. There is only a defense of its possibility, which is a critique of spurious metaphysics in Kant's *Critique of Pure Reason*, and an explicit discussion of this possibility in section

practical belief in the postulates of God and the immortality of the soul cannot be justified theoretically but is instead justified by our immediate certainty of the moral law. (KrV A631/B659-A642/B670)[5] Put programmatically, "Critik macht die Religion frey von der speculation" (VAProl 23:59). Such a practical justification of belief in God would be useless if there needed to be a further theoretical justification of morality; and it would certainly not 'free religion from speculation.'[6]

As Kant tells us at the beginning of his first systematic work on moral philosophy, practical reason needs moral philosophy because of a "propensity to rationalize against those strict laws of duty and to cast doubt upon their validity […] and, where possible, to make them better suited to our wishes and inclinations" (GMS 4:405), which Kant calls the 'natural dialectic.' Both Kant's theoretical and his practical philosophy are concerned with different kinds of 'dialectics' as misuses of reason, a "logic of illusion" (KrV (A293/B349). Against the misuse of theoretical reason, the remedy was to establish its strict limits (a *Critique* of pure reason), and thus to prevent philosophers from engaging in spurious metaphysics. As Kant tells us at the beginning of the *Groundwork*, practical reason too has its dialectic which demands a philosophical remedy. Here, the remedy will be to defend the claims of pure practical reason, in order to prevent philosophers from engaging in spurious limitations of practical reason by tying it to our natural inclinations. Both endeavors of Kant's critical philosophy are aptly called therapeutic. By reading Kant's practical philosophy in this light, we can better appreciate both intuitive suggestions above regarding the point of Kant's ethics:

III of the *Groundwork* and more extensively in the *Critique of Practical Reason*. For a recent and extensive discussion of this topic see especially Ludwig 2020 pp. 15-26, see also Neiman 1994 chapter 3, p. 107.

[5] See especially KrV A633f/B661f and KrV A811/B839.

[6] NB that this is not the same as the question whether freedom 'grounds' morality or morality 'grounds' freedom. Although there is some controversy over what exactly the relation between freedom and morality is in Kant, and whether or not Kant changes his view about this relation between writing the *Groundwork* and writing the second *Critique*, this debate is orthogonal the question whether the practical purpose of Kant's ethics was to respond to the moral skeptic.

Kant responds to the speculative threat against morality through his therapeutic endeavor in the *Critique of Pure Reason*, and thereby paves the way for a therapy of practical reason. The specific difficulties that common agents face arise from the natural dialectic which can cloud their (better) moral judgment.

So far, the relatively few remarks about the practical purpose of Kant's moral philosophy usually consider the topic in isolation from its historical and philosophical context. However, the therapeutic purpose of Kant's moral philosophy can only be fully appreciated by considering Kant's specific conception of moral knowledge and how it differed from other earlier and contemporary views. The practical purpose of Kant's moral philosophy goes hand in hand with Kant's specific conception of moral knowledge and his view of enlightenment as overcoming a darkness within us. Kant believed that the form of practical reason constitutes the fundamental principle of morality, and that this principle is thus implicit in the practical knowledge of every rational being. Moreover, Kant had a remarkably egalitarian view about the distribution of moral knowledge. For Kant, there is simply no room for moral experts who know better what is right or wrong than the common moral agent. At the same time, human beings are also fallible and vicious, because of our natural tendency to place natural desires above the demands of reason and to rationalize away the requirements of morality. This 'radical evil' part of our finite nature keeps us in self-incurred darkness. Because both the fundamental principle of morality as well as its main obstacle are already within us, Kant's moral philosophy attempts to help us overcome the self-incurred darkness by restoring trust in our rational, moral capacities. In brief, the practical purpose of Kant's ethics is enlightenment through therapy.[7]

---

[7] To my knowledge, the only authors who have taken a therapeutic reading of at least *parts* of Kant's ethics are John Callanan (2019), Ido Geiger (2010), Samuel Fleischacker (1991) – and, potentially, Lucas Thorpe (2006) and Jens

In the first section of this chapter, I briefly outline three important philosophical distinctions, which provide a helpful background for understanding Kant's specific views about the nature and distribution of moral knowledge in section two. In section three, I argue that Kant's conception of moral knowledge entails a particular view about the purpose of moral philosophy: enlightenment through therapy. I conclude with some tentative remarks on the further upshots for reading Kant's ethics that follow from his views about moral knowledge and the therapeutic purpose of moral philosophy.

Three Distinctions about Moral Knowledge

In order to get Kant's conception of moral knowledge properly in view, it is helpful to consider three pairs of distinctions which run orthogonal to the currently common ways of categorizing views in and about ethics. The first distinction is between what I propose to call the *scientistic* and the *archeological* models of moral knowledge.[8] On the scientistic view, beliefs about right

_____

Timmermann 2007b (especially at p. 182) – although neither of them has chosen to explicitly describe their reading as 'therapeutic' and only Callanan 2019 emphasizes the connection with Kant's notion of the natural dialectic. Thus, the closest position to the view presented here is presented by John Callanan, who has argued that understanding the 'natural dialectic' and the threat of mysology in section I of the *Groundwork* is crucial for understanding the *Groundwork* (Callanan 2019). While I agree with Callanan's insightful discussion of Kant's *Groundwork*, I argue that the practical purpose of Kant's moral philosophy *in general* is therapeutic. Ido Geiger has also focused on the *Groundwork*, and on Kant's four examples of derivations of imperatives of duty in section II in particular. Geiger compellingly argues that the practical use of the Universal Law formulation of the categorical imperative in these four examples is not in telling us what is right and wrong but in its 'heuristic value' that supposedly helps us overcome selfish temptations (Geiger 2010). Fleischacker (1991) has insightfully compared Kant's and Adam Smith's views about the purpose of moral philosophy in the *Groundwork* and the *Theory of Moral Sentiments* respectively. Only Thorpe (2006) considers the purpose of Kant's moral philosophy more generally, and argues that it is to provide theoretical understanding of morality, which will in turn make it easier to pay attention to the demands of morality. However, none of the authors mentioned above give much detailed attention to Kant's particular conception of moral knowledge, nor do they consider the purpose of moral philosophy in the context of Kant's views about self-incurred un-enlightenment. This chapter remedies this lacuna.

[8] Although the description of this distinction as between 'archeological' and 'scientistic' views might be new, the distinction itself is, of course, not. Candace Vogler has drawn the distinction by another name, and from a slightly different perspective, in her lectures in the history of ethics as between two purposes of moral philosophy. Vogler distinguishes between accounts that take the purpose of ethics to be the development of a 'theory of right action' (i.e., determining what is really right and wrong) and those that take the purpose of ethics to be the development of a 'theory

and wrong are akin to beliefs about natural phenomena in that it would be theoretically possible for everyone to be in complete error about the most fundamental facts of morality – just like everyone might be in complete error about the best explanation for some natural phenomenon. Just like a vast number of people once erroneously, but honestly, believed that the sun travels around the earth, so people might be prone to errors about what is right and wrong in action. Moreover, on the scientistic view we cannot claim to have anything more than mere opinion on moral matters unless we are in possession of a persuasive theory justifying what constitutes right and wrong. Thus, the search for moral principles – if philosophers do so – is a search for something that can transform our mere opinions into justified beliefs about right and wrong. Note that this scientistic model of moral knowledge is not restricted to either metaethical 'cognitivists' or 'realists.' One may believe that, ultimately, there is no such thing as truth in moral matters (and thus that we can never achieve anything more than empty opinion, and that normative ethics is, consequently, pointless), or one might believe that what makes people apply the label 'right' to an action is merely a function of their emotional states (and thus that knowledge about ethical matters is ultimately just knowledge of human psychology), and in both cases one could hold a scientistic view of moral knowledge. More importantly, the scientistic model of moral knowledge implies a particular view about the purpose of moral philosophy: if moral philosophy was possible at all, its purpose would be to determine what is

---

of moral judgment' (i.e., how we make judgments about what is right and wrong). (Vogler, Lectures on Kant's Ethics 2019, unpublished) James Griffin has also drawn this distinction in passing by another name, referring to "two quite different, incompatible conceptions of the relation of philosophy to ethics." Griffin distinguishes between a conception of philosophy as justifying, sanctioning and systematizing otherwise unreliable beliefs about what is right and wrong, where to reason morally necessarily is to reason philosophically, and a conception of philosophy as taking up an independently existing human practice, where the philosopher does not have special standing regarding claims about right and wrong (Griffin 1996 p. 131). For a similar distinction in meta-epistemology see Sosa 1991 pp. 158-159. Finally, Jerome Schneewind discusses the distinction as relating to different conceptions about the history of the moral philosophy, distinguishing between two origin stories about moral philosophy: the 'Socrates-story' and the 'Pythagoras-story' (Schneewind 1998 pp. 533-554 and Schneewind 2010 pp. 107-126).

really right and wrong, just like the purpose of empirical sciences is to determine what is really the best explanation of some natural phenomenon.[9]

By contrast, on the archeological view, knowledge about what is right and wrong could not be subject to such widespread error, because the basic determinant of moral judgment is already within us. Consequently, the purpose of moral philosophy cannot be to find the true and potentially new way of justifying moral beliefs. Instead, the purpose of moral philosophy would be to unearth something that each of us – or, at least, most healthy rational beings – already carry within. Just like the scientistic model, the archeological conception of moral knowledge is not restricted to any contemporary metaethical theory.[10] Rather, what matters is that the archeological model of moral knowledge comes with a specific view about the purpose of moral philosophy: its purpose is to illuminate the basic principles underlying the use of our evaluative faculties. In this history of western philosophy, the dominating conception of moral knowledge has certainly been the archeological model. According to the Christian story, influenced heavily by Plato, all rational beings have been endowed by God with reason, and thus also with access to the basic principles of natural law.[11] If all human beings are endowed with such access to the basic principle of natural law, then the task of the moral philosopher is to unearth these principles (and potentially also to explicate all their minute consequences).[12]

---

[9] Paradigmatic proponents of this scientistic view throughout the history of western philosophy are John Stuart Mill (especially chapter I of *Utilitarianism*) and Henry Sidgwick (see especially chapter I of *The Methods of Ethics*). Compare also Rorty 2004.

[10] Although I take it that the majority of proponents of the archeological model in the history of western ethics were scholastic, natural law theorists.

[11] See e.g., Pufendorf I.iii.12 (1991 p. 37)

[12] Paradigmatic proponents of the archeological model include Thomas Aquinas (see ST I.79.12) Adam Smith and Immanuel Kant.

A second important distinction is that between *epistemic elitist* and *epistemic egalitarian* conceptions of moral knowledge. On epistemic elitist views, only the wise or virtuous have moral knowledge. This might be because of their alleged superior rational capacities, or because of their better upbringing or their more developed emotional sensitivities. In its extreme form, an elitist view of moral knowledge takes knowledge of right and wrong to be akin to aesthetic sensibilities, and claims that knowledge of right and wrong just isn't for everyone.[13] Less extreme versions of elitism about moral knowledge have taken the view that only the trained philosophers (or theologians) can properly access the principles of natural law and determine its consequences or derivative laws. Consequently, epistemic elitist views usually endorse the idea of moral experts who can tell laypeople how to properly lead their lives. And such epistemic elitism is by no means restricted to Ancient and scholastic ethics. As late as the 1970s, famous utilitarian ethicists have publicly argued that laypeople have much to learn from expert moral philosophers and their rational assessment of difficult ethical questions.[14]

By contrast, egalitarians about moral knowledge believe that the very idea of moral experts is dubious, if not simply false. This might be because they believe that all (healthy) people have the same emotional sensitivities that enable them to distinguish right from wrong, or because they believe that learned cultural norms are reasonable vehicles of moral knowledge, or because they

---

[13] Famously, such arguments were part of ancient defenses of slavery. See e.g., Aristotle *Politics* I.13.1260a12 and *Politics* book I, chapters 4-8 more generally.

[14] An excellent example of this epistemic elitism is Peter Singer's 1974 article in *New York Times Magazine* 'Philosophers Are Back on the Job'*,* in which Singer claims that academic moral philosophers have superior knowledge of 'moral concepts' and the special 'logic of moral arguments,' and consequently that both the wider public and key professionals like medical doctors can supposedly learn much from the philosophers' superior rational assessment of such difficult ethical topics as euthanasia and abortion. Singer also argued that "it would be surprising if moral philosophers were not, in general, better suited to arrive at the right, or soundly based, moral conclusions than non-philosophers." Importantly, Singer here goes so far as to claim that "Indeed, if this were not the case, one might wonder whether moral philosophy was worthwhile." (1972 p. 1117) As I will argue in detail below, Kant believed that moral philosophy was indeed worthwhile, even though he endorsed an egalitarian conception of moral knowledge.

23

believe that all rational beings have equal access to the basic principles of practical reason. (Note that epistemic elitism is not necessarily restricted to academic elitism. One may deny that there could be moral experts in an academic sense, but still take an elitist view of moral knowledge based on e.g., a person's physical and psychological characteristics or their upbringing and education, as e.g., Aristotelian and Neo-Aristotelian accounts have done respectively.)

On first sight, it is tempting to assume that the scientistic model of moral knowledge predisposes one toward epistemic elitism, while the archeological model predisposes toward epistemic egalitarianism. For according to the archeological model of moral knowledge, all moral agents have some access to the basic determinant of moral evaluation. And according to cognitivist versions of the archeological view, all moral agents have some knowledge of the fundamental principles of practical reason. However, archeological views about moral knowledge are just as open to epistemic elitism as scientistic views. In fact, for much of western philosophical history, the dominating view of moral knowledge was both archeological and elitist. A concise expression of this (once standard) view can be found in Johannes Althusius work *Politica methodice digesta atque exemplis sacris et profanis illustrate* from 1603. There, Althusius explains that:

> There are different degrees of this knowledge [by which we know the moral law] and inclination. For law is not inscribed equally on the hearts of all. The knowledge of it is communicated more abundantly to some and more sparingly to others, according to the will and judgment of God. Whence it is that the knowledge of this law may be greater in some than in others.[15]

Wolff's ethics – still popular during Kant's lifetime – similarly argued that in order to understand properly what is right and wrong, one must first gain deep metaphysical knowledge. Because such metaphysical knowledge was, unsurprisingly, beyond the reach of most non-academic philosophers,

---

[15] Althusius 1964 pp. 135-136.

Wolff believed that society should rely on moral experts who can instruct people on how to live and what rules to follow.[16] Thus, elitism about moral knowledge was the standard view in philosophy for most of the previous millennium. But as Jerome Schneewind has shown in his studies of the history of modern moral philosophy, this picture started to change in the 18th century.[17] One reason for taking a more egalitarian view about the distribution of moral knowledge among the jurists and philosophers of the early modern period was a belief in universal accountability for one's actions before God. Consequently, 18th century thinkers argued that even the common moral agent must have an intuitive knowledge of the requirements of morality, in order to explain people's accountability for their actions before God. Joseph Butler's expression of his egalitarianism about moral knowledge is particularly memorable:

> The inquiries which have been made by men of leisure after some general rule, the conformity to, or disagreement from which, should denominate our actions good or evil, are in many respects of great service. Yet let any plain honest man, before he engages in any course of action, ask himself, 'Is this I am going about right, or is it wrong? Is it good, or is it evil?' I do not in the least doubt but that this question would be answered agreeably to truth and virtue, by almost any fair man in almost any circumstance.[18]

As quaint as Butler's formulation of his epistemic egalitarianism might sound today, it plausibly captures an almost universally shared cultural premise of modern, liberal societies. (Many of our

---

[16] Wolff 1736 §§148-150 (pp. 82-84). For English translation see Wolff 2003 p. 341. For an extensive overview of epistemic elitist views in the history of western moral philosophy, see Schneewind 2010 pp. 87-92. Unsurprisingly, elitist views of moral knowledge often come together with arguments for the alleged authority of moral experts to which laypeople ought to defer. This natural outcome of elitist views has led Bernard Williams to reject the elitist conception (1995 p. 205). However, belief in moral experts does not *necessarily* come with belief in moral authority of experts. For instance, Hasting Rashdall as argued that moral experts serve a role analogous to the role of a judge in a jury trial – directing the jury about the implications of their reasoning and the facts – rather than an authority to which they should defer in moral matters (Rashdall 1894). Sarah McGrath has recently suggested that Rashdall's view is in fact the predominant self-conception of moral philosophers (McGrath 2014 p. 198-199).

[17] For a detailed discussion of the development of egalitarian views about moral knowledge see Schneewind 1991 pp. 289-294. For the claim that egalitarianism about moral knowledge is one (out of three) defining feature of modern moral philosophy – and also of our own, contemporary assumptions – see Schneewind 2010 pp. 92-96.

[18] Butler 2017 p. 34

most basic institutions – such as universal suffrage, jury duty, public trials, freedom of expression etc. – presume the potential of every adult member of the body politic to come to a practically reasonable decision in moral and political matters.)

The third distinction to get in view concerns moral judgment: as either *procedural* or *immediate*. Perhaps the most obvious example for views that adopt a procedural conception of moral judgment are act-maximizing utilitarian accounts of morality. If knowledge about what is right and wrong depends on (correctly) assessing the likely consequences of our actions, and an exercise of calculating and/or balancing outcomes, then a correct moral judgment will require a conscious process of deliberation. Otherwise, our intuitive judgement about what is right and wrong would be a mere intuition, and far from genuine moral cognition. Caricature-views of Kant's ethics also represent procedural conceptions of moral knowledge: if our knowledge of what is right and wrong in a particular situation depends on us correctly applying e.g., a universalization test, then moral judgment requires a process of deliberating about our maxims. And even sympathetic readers of Kant's ethics have taken the view that the categorical imperative serves primarily an active deliberative function in practical reasoning.[19]

By contrast, what we can tentatively call 'immediate' conceptions of moral judgment do not require a conscious process of weighing, balancing or universalizing. Extreme versions of such a view would be variants of ethical intuitionism. On immediate conceptions of moral judgment, the representation of an action as to be done would be more akin to perception than to calculating consequences: my taking what I see before me to be a coffee cup does not require a conscious thought-process of applying certain concepts to a heap of sense-data, as if such taking-something-

---

[19] See for instance Theunissen's remarks on this 'standard reading' 2013 p. 109, and Sticker 2015. Most famously, this view has been attributed to Kant by O'Neill 1975 and Rawls 2000.

to-be-something necessarily consisted in assertoric judging or conceptual sorting. Even if we believe – with Kant – that such taking-to-be requires the application of categories of the understanding to a manifold, we need not also take this exercise of conceptual capacities as a conscious thought-process of assertoric judging or conceptual sorting. Similarly, we might think that our representation of an action as good will involve a practical judgment that applies certain conceptual capacities, but such a representation of an action as good might not imply a conscious process of deliberation or maxim-testing preceding it.[20] (Indeed, starting with Sidgwick, and continuing at least until the middle of the twentieth century, the standard classification of Kant was as an ethical intuitionist.[21])

Kant's Conception of Moral Knowledge

Kant was an archeologist in moral philosophy, an egalitarian about moral knowledge, and thought moral judgment was immediate. When Kant analyses common moral knowledge in order to find the supreme principle of morality, he is after a rational (a priori) principle that was there all along, and which makes moral judgment possible for finite, rational beings. This supreme principle of morality is the form of practical reason, thus we always have this principle 'before our eye,' even when we are ignorant of it in its explicit linguistic expression like the formulations of the categorical imperative. And since all rational beings have this form of practical reason always 'before their eye,' even the least educated person can know what morality requires without any technical or philosophical reflection. (Indeed, as we will see below, Kant's remarks about the nature of moral knowledge often simultaneously refer to both its archeological as well as its egalitarian aspect.) And,

---

[20] For a discussion of how to conceive of such an 'immediate' judgment in both practical and theoretical cognition – and how they relate to Kant's notion of spontaneity – see Pippin 2013.
[21] See for instance Matson 1954 p. 858.

finally, Kant also believes that moral judgment is immediate as opposed to procedural: we need not go through any conscious thought-process about reasons for acting morally, or about the possible contradictions (logical or practical or whatnot) that we might get ourselves into by acting immorally. Rather, Kant thought that rational beings are aware of the requirements of morality immediately, since these are determined by the very form of practical reason.[22]

Kant's Archeology

Kant's archeological methodology perspicuously structures his first 'critical' book on moral philosophy: the first section of the *Groundwork* concerns the 'transition from common to philosophical moral rational cognition.' Although at this stage of the project we do not yet know if morality is more than a phantom of the brain, Kant makes clear that *if* there is such a thing as moral judgment by finite, rational beings, then we must find its highest principle in the presuppositions of common moral judgment. And moving from common to *philosophical* moral cognition would be to find the principle inherent in common moral cognition. Only then can we move to a 'metaphysics of morals' (i.e., to analyze the rational principles that underlie our moral judgment and their systematic interconnection) and a critique of practical reason (i.e., to determine if reason is in fact capable of acting in the way our analysis of practical judgment showed).

Kant again states his archeological view of moral philosophy in the *Critique of Practical Reason*. There he argues that

---

[22] Although I am sympathetic to Kant's conception of moral knowledge, I do not here attempt to defend it philosophically, nor do intend to provide a historical explanation of Kant's own reasons for this approach. Here, I am only interested in getting clear about Kant's conception of moral knowledge in order to draw conclusions about the possible practical purpose of moral philosophy.

[…] pure reason, without the admixture of any empirical determining ground, is practical of itself alone: this one had to be able to show from the most common practical use of reason, by confirming the supreme practical principle as one that every natural human reason cognizes – a law completely a priori and independent of any sensible data – as the supreme law of its will. It was necessary first to establish and justify the purity of its origin even in the judgment of this common reason before science would take it in hand in order to make use of it, so to speak, as a fact that precedes all subtle reasoning about its possibility and all the consequences that may be drawn from it. (KpV 5:91)

Just shortly after, Kant asks how the causality of a transcendentally free being can be conceived, and comes to discuss the categorical imperative as an answer to this question. In this context, Kant remarks: "Now, this principle does not need to be searched for or devised; it has long been present in the reason of all human beings and incorporated in their being, and is the principle of morality." (KpV 5:105) And also in his handwritten notes to the *Observations on the Feeling of the Beautiful and the Sublime*, published after his death, Kant memorably expresses his archeological view of moral philosophy. In a famous note, Kant says:

Newton saw for the first time order and regularity combined with great simplicity, where before him was found disorder and barely paired multiplicity; and since then comets run in geometrical courses. *Rousseau discovered for the first time beneath the multiplicity of forms human beings have taken on their deeply buried nature and the hidden law by the observation of which providence is justified.* Before that the objection of Alphonsus and Manes still held. After Newton and Rousseau, God is justified and Pope's theorem is true. (HN 20:58-59, emphasis added)[23]

And finally, we find references to Kant's archeological view again in his work on theology and rational religion. In the *Religion within the boundaries of mere reason*, Kant draws a distinction between pure moral faith, which is based on reason alone and can thus attain universality among human beings, and faith of revelation, which is based on historical events and thus cannot claim the same sense of universality as moral faith unless it is itself based on the latter. Importantly, Kant

---

[23] Also in Rischmüller 1991 p. 48, translation from *Notes and Fragments* p. 9

argues that historically revealed faith must always be interpreted in light of morality (and even suggests that 'thoughtful teachers' have always already done so in all major world religions)[24]. And he continues: "That this, however, can be done without ever and again greatly offending against the literal meaning of the popular faith is due to the fact that, long before this faith, *the predisposition to moral religion lay hidden in human reason.*" (RGV 6:111)[25] Thus, consistently throughout his major works, Kant's moral philosophy is the work of an archeologist trying to unearth something that was there all along.[26]

Kant's Epistemic Egalitarianism

Kant first expresses his egalitarianism about moral knowledge in the *Critique of Pure Reason*. In his 'General remarks on the transcendental aesthetic,' Kant outlines his view about the impossibility of representing things as they are in themselves, rather than merely as they appear to us, and argues that the distinction between appearances and things in themselves does not apply to practical cognition. In this context, he says:

> Without doubt the concept of right that is used by the healthy understanding contains the very same things that the most subtle speculation can evolve out of it, only in common and practical use one is not conscious of these manifold representations in these thoughts. Thus one cannot say that the common concept is sensible and contains a mere appearance, for right cannot appear at all; rather its concept lies in the understanding and represents a constitution (the moral constitution) of actions that pertains to them in themselves. (KrV A43/B61-A44)

---

[24] RGV 6:110-111

[25] For more passages concerning the Kant's view about interpreting religion morally see also RGV 6:115, 6:121 and SF 7:37, 7:44, 7:64 and 7:69.

[26] This is also the consensus among Kant scholars. See especially Greenberg 2013 and 2015, and also Geiger 2015. For a qualified detraction from this consensus-view see Sticker 2016.

Kant again expresses this epistemic egalitarianism toward the end of the first *Critique*. At the very end of the subsection of the Canon of Pure Reason, in a subsection concerned with the distinction between having an opinion, believing and knowing, Kant discusses the possibility of practical belief in God. Although Kant's immediate focus here is belief based on our moral cognition rather than moral cognition itself, the final passage of the subsection demonstrates Kant's deep respect for common moral knowledge.

> But is that all, one will say, that pure reason accomplishes in opening up prospects beyond the bounds of experience? Nothing more than two articles of belief? This much common understanding could also have accomplished without taking advice from the philosophers! […] But do you demand then that a cognition that pertains to all human beings should surpass common understanding and be revealed to you only by philosophers? The very thing that you criticize is the best confirmation of the correctness of the assertions that have been made hitherto, that is, that it reveals what one could not have foreseen in the beginning, namely that in what concerns all human beings without exception nature is not to be blamed for any partiality in the distribution of its gifts, and in regard to the essential ends of human nature even the highest philosophy cannot advance further than the guidance that nature has also conferred on the most common understanding. (KrV A830/B858-A831/B859)

These passages in Kant's *Critique of Pure Reason* already demonstrate his firm commitment to egalitarianism about moral knowledge – but perhaps the most explicit (and well-known) passage concerning the adequacy of common moral knowledge occurs towards the end of section I of the *Groundwork*, in the section entitled *Transition from common rational to philosophical moral cognition*. There, Kant says:

> Thus, then, we have arrived, within the moral cognition of common human reason, at its principle, *which it admittedly does not think so abstractly in a universal form but which it actually has always before its eyes* and uses as the norm for its appraisals. Here it would be easy to show how common human reason, with this compass in hand, knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty or contrary to duty, if*, without in the least teaching it*

anything new, we only, as did Socrates, make it attentive to its own principle; and that *there is, accordingly, no need of science and philosophy to know what one has to do in order to be honest and good*, and even wise and virtuous. We might even have assumed in advance that cognizance of what it is incumbent upon everyone to do, and so also to know, would be the affair of every human being, even the most common. (GMS 4:403-404, emphasis added)[27]

This passage in the *Groundwork* not only demonstrates Kant's trust in the practical knowledge of common moral agents. It also demonstrates the strong connection between Kant's archeological methodology and his egalitarianism about moral knowledge. As he points out here, his concern is to make human reason 'attentive to its own principle,' without 'teaching it anything new.' This thought is also in line with Kant's description of the methodology of the *Groundwork*. In the final paragraph of the preface, Kant tells us that with his method we "proceed analytically from common cognition to the determination of its supreme principle, and in turn synthetically from the examination of this principle and its sources *back to the common cognition in which we find it used.*" (GMS 4:392, emphasis added) Importantly, Kant here indicates that his method does not take us to a philosophical method of assessing what's good and bad, but starts with common moral knowledge and *takes us back* to the cognition with which we started. In other words, there is no special philosophical way of figuring out what is the right thing to do.

Finally, we find a memorable expression of Kant's egalitarianism about moral knowledge in his *Theory and Practice* essay. There, Kant responded to Christian Garve's criticism that his moral philosophy might be correct in theory, but of no use in practice.[28] In his response, Kant claimed that

---

[27] See also GMS 4:454

[28] To claim that Kant's moral philosophy was no use in practice was entirely unoriginal. For what is likely the first version of this type of criticism, see Tittel 1786. I discuss Tittel's ciriticism in detail in 'Kant's Derivation of Imperatives of Duty.' For a detailed discussion of Kant's exchange with Garve in the Theory and Practice essay see Timmermann 2007b.

"everything in moral philosophy that is correct for theory must also hold for practice." To this, he adds:

> Everyone in his capacity as a human being, a being subjected by his own reason to certain duties, is accordingly a man of affairs; and since, as a man, he never outgrows the school of wisdom, he cannot with proud contempt, as someone supposedly better instructed by experience about what a human being is and what can be required of him, send the adherent of theory back to school. (TP 8:288)

Although Kant's use of the term 'man of affairs' (Geschäftsmann) might sound strange to our ears, the term here simply refers to someone who is competent in their business. Thus, Kant's moral philosophy is trying to show that in moral matters there is no hierarchy. Philosophers do not know better what is right and wrong than the layperson – they are both competent in the business of ethical life.[29] Indeed, Kant's insistence on the adequacy of common moral knowledge suggests that the philosopher may even be in a worse position than the common moral agent. In the passage from the *Groundwork* quoted above, Kant emphasizes that "there is, accordingly, no need of science and philosophy to know what one has to do in order to be honest and good." (GMS 4:404) And while the common moral agent is merely tempted away from their immediate moral judgment by personal interest, the philosopher can be further tempted by misleading theory.

In his notes on his copy of the *Observations on the Feeling of the Beautiful and the Sublime*, Kant tells us that it was Rousseau who convinced him of epistemic egalitarianism. In a famous passage from his handwritten remains, he says:

> I am myself by inclination an investigator. I feel a complete thirst for knowledge and an eager unrest to go further in it as well as satisfaction at every acquisition. There was a time when I believed that

---

[29] See also KpV 5:36. There is, of course, an obvious potential problem for this conception of moral knowledge: either the universally shared moral knowledge is entirely formal or abstract (in which case it risks becoming meaningless); or it is supposedly substantive, in which case the view is (at least in theory, and likely also in practice) open to empirical falsification. Much of the attractiveness of Kant's moral philosophy depends on whether one believes that it can avoid both horns of this apparent dilemma. Hegel's empty-formalism charge famously ceases on the former horn of the dilemma. I discuss Hegel's charge in detail in 'The Formalism of the Categorical Imperative.'

this alone could constitute the honor of mankind, and I had contempt for the rabble who know nothing. Rousseau brought me around. This blinding superiority disappeared, I learned to honor human beings […] (HN 20:44)[30]

If moral knowledge is distributed equally among finite, rational beings, then the practical purpose of moral philosophy can hardly be to help people figure out what is really moral and immoral. Not only is moral philosophy an exercise in rational archeology, trying to unearth a principle of practical reason that was there all along – this view Kant shares with the (frequently elitist) writers of the Natural Law tradition. In addition, Kant's epistemic egalitarianism is strikingly substantive.[31] As the passages surveyed here suggest, it seems that moral philosophy is not even in the business of drawing out the minute and complicated consequences of this supreme principle of practical reason so that common people may know which concrete rules or actions embody this first principle – they are allegedly well equipped without philosophy to know 'what one has to do' (GMS 4:404).[32] On Kant's view, common moral knowledge is already adequate as long as it remains unclouded by our tendency to rationalize against it in favor of opposing inclinations.[33]

---

[30]Also in Rischmüller 1991 pp. 37-39, translation from *Notes and Fragments* p. 7. Unfortunately, in the secondary literature this passage (HN 20:44) is sometimes wrongly cited as 2:216–17. This is likely due to the citation method of the CUP edition, which does not mention the place in the AA where the actual note was published but instead mentions the AA number of the passage of the *Observations* to which Kant's notes were added (as in this case GSE 2:216-17). In this, the CUP edition follows Rischmüller 1991. Unfortunately, there is no clear convention yet for numbering Kant's notes, and the CUP edition decided not to follow Rischmüller's numbering but started their own. I therefore omit numberings of the notes to avoid confusion and instead cite according to the AA edition in the text. For another relevant note regarding Rousseau and equality see also HN 20:176 (Rischmüller 1991 p. 130 and *Notes and Fragments* p. 23). For Rousseau's positive influence on Kant more generally see further Ameriks 2012, Cassirer 1983, Henrich 1992, Shell 2009, Shell and Velkey 2017, Velkley 1989, Zammito 2002.

[31] For the purpose of this paper, I leave aside the question whether Kant is at all entitled to this substantive epistemic egalitarianism. I discuss Hegel's complaint that Kant' analysis of the form of practical judgment does not entitle him to this epistemic egalitarianism in 'The Formalism of the Categorical Imperative.'

[32] This consequence of Kant's egalitarian conception of moral knowledge is also demonstrated by his aversion to casuistry. I discuss Kant's views of casuistry and casuistical questions in detail in 'Kant's Casuistical Questions'.

[33] For a recent defense of Kant's ethics as a vindication of felt phenomenological experience of moral life – both as philosophical starting point and ultimate grounding relation – see Grenberg 2013.

The Immediacy of Moral Judgments

According to Kant, judgments about what is morally right or wrong are immediate in the sense that they do not require a conscious process of deliberation.[34] Kant's emphasis on this aspect of moral judgment becomes especially apparent when he invokes the imagined response of children and students to moral questions (which I discuss in more detail below).[35] Speaking in more abstract terms, in the *Critique of Practical Reason*, Kant claims that "What is to be done in accordance with the principle of the autonomy of choice is seen quite easily and *without hesitation* by the most common understanding […]." (KpV 5:36, emphasis added) Similarly, when Kant considers the example of an ordinary agent who is tempted to embezzle a deposit, he writes: "if he asks himself what his duty is in this matter, he is not at all perplexed about what answer to give but *certain on the spot* what he has to do." (TP 8:287, emphasis added) These passages suggest that Kant thought of moral judgment as immediate in the sense in which we 'immediately' judge that something on the table is red. Kant's own, preferred analogy is more self-referential. In the 'doctrine of method of pure practical reason,' Kant writes that judging right from wrong is, to common human reason, like telling left from right.

> But if one asks: What, then, really is pure morality, by which as a touchstone one must test the moral content of every action? I must admit that only philosophers can make the decision of this question doubtful, for it is long since decided in common human reason, not indeed by abstract general formulae but by habitual use, like the difference between the right and the left hand. (KpV 5:155)

A more extensive expression of Kant's view of the immediacy of moral judgment comes up in Kant's *Theory and Practice* essay, in his response to Christian Garve. Garve had claimed that although he admitted the correctness of Kant's view 'in his head,' his heart just could not follow it.

---

[34] On Kant's view, the form of practical reason does not just enable non-procedural or immediate cognition of what is right and wrong. Kant also believed that this cognition can itself be motivating. Kant explains how this pre-reflective state of mind can be both rational and immediately motivating through his account of 'respect' (Achtung). For the purpose of this essay, I do not discuss how moral judgments can be motivating through respect.

[35] See e.g., TP 8:286 and MS 6:480-81. I discuss these passages in more detail in 'Reflection in Moral Knowledge.'

To this, Kant responded that Garve was more likely to have rationalized away his 'heart's'

knowledge. Kant's response to Garve is illuminating on several counts, and is worth quoting

(almost) in full.

> I now come to the point that really concerns us here, namely to illustrate with examples and to test
> the supposed conflicting interests of theory and of practice in philosophy. Garve gives the best
> example of it in his treatise cited above. He says first (speaking of the distinction I find between a
> doctrine of how we are to become happy and one of how we are to become worthy of happiness):
> "For my own part, I confess that I very well conceive this division of ideas in my head, but that I do
> not find this division of wishes and strivings in my heart. […]

> As for Garve's avowal, just cited, that he does not find such a division (strictly speaking, separation)
> in his heart, *I have no hesitation in contradicting his self-accusation outright and in championing his heart against his*
> *head. He, a man of integrity, has actually found this separation in his heart every time (in his determination of will),*
> *only it would not be reconciled in his head* - for the sake of speculation and of comprehending what is
> incomprehensible (inexplicable), namely the possibility of categorical imperatives (such as those of
> duty are) — with the usual principles of psychological explanation (all of which have the mechanism
> of natural necessity as their basis). (TP 8:284-285, emphasis added)

This passage provides a beautiful illustration of Kant's earlier statement in the *Groundwork* that

common human reason has its principle always implicitly before its eye, and that "there is,

accordingly, no need of science and philosophy to know what one has to do in order to be honest

and good." (GMS 4:404) On Kant's view as presented here, if Garve had not rationalized away his

immediate moral judgment with reflections about his finite, dependent nature as a moral being, then

Garve himself would have remained perfectly aware of the strength of the requirements of practical

reason. Moreover, Kant even goes so far as to suggest that Garve *still*, in a way, knows the

requirements of practical reason 'in his heart,' however much he might have rationalized against

them.[36] In this way, Kant's reply shows not only that he believed that moral judgment can be

---

[36] This immediacy of moral knowledge also plays a crucial role in Kant's justification of the postulates in the *Critique of*
Practical Reason. After all, if conscience of the moral law were not immediate, the moral law would need further

immediate and precede conscious deliberation, but that moral cognition can even *survive* misguided deliberation and reflection. (Just like one might think that my cognition that there is a car approaching can survive my skeptical rationalizing (*vernünfteln*) about the reliability of sense-experience.)

Of course, the immediacy of moral judgment does not imply that moral judgment is simply "given" to entirely passive moral agents. Quite to the contrary, on Kant's view it is the form of practical reason that makes moral cognition so much as possible. But Kant's remarks also make clear that it would be a mistake to think of the contribution of practical reason as active in the sense of a conscious process of deliberation. Instead, the activity of practical reason in enabling immediate moral judgment appears more analogous to the activity of the understanding in enabling empirical cognition.[37]

The Practical Purpose of Moral Philosophy: Enlightenment through Therapy

If moral philosophy is supposed to uncover the practical principles that we all have implicitly 'before our eyes,' and if there is no need for moral experts to help us figure out what is really right and wrong, what, then, is moral philosophy good for? In light of Kant's conception of moral knowledge, we might think that moral philosophy is simply a theoretical exercise. It might give us a more sophisticated understanding of *why* some things are moral or immoral, but there may not be any practical purpose to moral philosophy. On such a reading, the purpose of Kant's ethics would

---

justification. But in this case, Kant would not be able to argue for a 'practical belief' in God and the immortality of the soul on the basis of morality. For a detailed discussion of the importance of the immediacy of moral knowledge for Kant's justification of the postulates see Ludwig 2020 pp. 15-26.

[37] I discuss this analogy in more detail in 'Reflection in Moral Knowledge.'

be entirely metaethical. Indeed, some commentators have either ascribed such a view to Kant or to his interpreters.[38] However, Kant's own remarks about moral philosophy strongly suggest that Kant believed moral philosophy to be more than a merely theoretical exercise. To the contrary, practical wisdom "needs science, not in order to learn from it but in order to provide access and durability for its precepts." (GMS 4:405) For Kant, moral philosophy is thus not just philosophically interesting, it is practical.

But if Kant neither believed that his task was to win over the moral skeptic, nor to determine what morality really requires for the first time, what practical purpose might moral philosophy serve? After noting that practical wisdom 'needs science' in order to find 'durability for its precepts,' Kant claims that between the demands of our more selfish inclinations and the demands of morality, there can arise a natural dialectic, which leads people to rationalize away the requirements of morality:

> There is something splendid about innocence; but what is bad about it, in turn, is that it cannot protect itself very well and is easily seduced. Because of this, even wisdom - which otherwise consists more in conduct than in knowledge - still needs science, not in order to learn from it but in order to provide access and durability for its precepts. The human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect - the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name happiness. Now reason issues its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with disregard and contempt for those claims, which are so impetuous and besides so apparently equitable (and refuse to be neutralized by any command). But from this there arises a natural dialectic, that is, a propensity to rationalize against

---

[38] Larry Krasnoff, for instance, has claimed that this metaethical reading is the "most familiar" reading of Kant's ethics (2004 p. 135) and the "conception standardly ascribed to Kant" (ibid. p. 136). However, Krasnoff does not actually cite any Kant scholarship for this claim, but merely refers to two works in contemporary ethics: Christine Korsgaard's *Sources of Normativity* (1996b) and Michael Smith's *The Moral Problem* (1994). Lucas Thorpe (2006) comes close to the metaethical reading (and Callanan 2019 ascribes such a metaethical reading to Thrope 2006). Although Thrope claims that Kant's ethics does serve a practical purpose, he argues that this simply comes down to 'clarifying an ideal' and providing 'theoretical understanding.' By providing theoretical understanding, Thorpe argues that Kant's ethics intends to make it easier for human beings to pay attention to the demands of morality. However, Thorpe neither connects this to Kant's conception of enlightenment nor his account of radical evil and the 'natural dialectic.' Thus, his account remains closer to the metaethical reading of Kant's ethics than to the therapeutic reading.

those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity - something that even common practical reason cannot, in the end, call good. In this way common human reason is impelled, not by some need of speculation (which never touches it as long as it is content to be mere sound reason), but on practical grounds themselves, to go out of its sphere and to take a step into the field of practical philosophy, in order to obtain there information and distinct instruction regarding the source of its principle and the correct determination of this principle in comparison with maxims based on need and inclination, so that it may escape from its predicament about claims from both sides and not run the risk of being deprived of all genuine moral principles through the ambiguity into which it easily falls. So there develops unnoticed in common practical reason as well, when it cultivates itself, a dialectic that constrains it to seek help in philosophy, just as happens in its theoretical use; and the first will, accordingly, find no more rest than the other except in a complete critique of our reason. (GMS 4:405, see also 4:411n)[39]

Since our inclinations appear so 'apparently equitable', and since reason does not 'promise anything to them,' a natural dialectic arises which tempts us to rationalize away the requirements of morality. Importantly, this is not just the naïve thought that we might be overpowered by inclinations to do what we believe to be wrong. In other words, the natural dialectic is not simply a conflict between inclination and reason; rather, this natural dialectic is itself a rational process. As finite beings, we necessarily take an interest in our own happiness. And our happiness is a genuinely valuable end – which makes pursuing our happiness also appear so 'equitable.' And the experience of our judgment about what is intrinsically good and what is furthering our own happiness coming apart is a naturally uncomfortable – if not unnerving – experience. If we take an interest in what we judge intrinsically good as well as in what furthers our own happiness, then we necessarily also take an interest in achieving both. To decide between them and to forfeit one is deeply uncomfortable in itself – that

---

[39] As John Callanan has compellingly argued, in the passage on the natural dialectic, Kant is implicitly responding to Rousseau. In particular, Kant disagrees with Rousseau that science, art and philosophy are corrupting the common moral knowledge of the natural state of humans. Kant agrees with Rousseau's view about the correctness of common moral knowledge, but does not share his disdain towards philosophy. (Callanan 2019)

is, it is deeply uncomfortable independently of the discomfort of frustrating our given inclination. And since it is hardly an option for finite, dependent beings like us to shape our inclinations at will, we take a pathological interest in justifying them. In other words, we are naturally prone to rationalizing in favor of our inclinations and against our moral judgment *not* simply because we are suffering from overly strong inclinations, but precisely because we are finite, rational beings who take an interest in what is reasonable (as our happiness is).[40] In order to overcome this natural dialectic, we need the resolve to trust our capacity for pure practical reason and pursue what is intrinsically good even when it conflicts with our happiness. Kant believed that philosophy could help people find this resolve to rely on their rational capacities and to trust their immediate moral judgments, instead of rationalizing away the requirements of morality and relegating their decisions to inclinations. Moreover, while bad moral theory risks reinforcing our rationalizations, moral philosophy done well deprives us of such tools and may even help us at catching ourselves out.[41]

Unsurprisingly, the practical need for moral philosophy in the natural dialectic also relates to Kant's view of the root of all human evil. In the *Religion within the boundaries of mere reason*, Kant provides an explanation of how it can be that free, rational beings like us, with a natural inclination towards moral goodness, nevertheless can fail to live up to the requirements of reason. Sometimes, human beings can feel impulses toward some aspect of our own happiness so strong that they overpower our impulses from the requirements of pure practical reason. In this way, humans are simply frail, finite, dependent beings. But according to Kant, human beings are not only frail in this banal sense. They also have a pathological propensity to subordinate the requirements of pure

---

[40] For related discussion of Kant's notion of 'rationalizing' (vernünfteln) see also Sticker 2021, especially pp. 29-30.
[41] For recent discussion of Kant's account rationalizing and self-deceit, see Papish 2018 and. Wehofsits 2020.

practical reason to their inclinations: 'radical evil.'[42] Although Kant does not explicitly draw on the

Groundwork's discussion of the natural dialectic in his discussion of radical evil, they clearly must be

connected if radical evil is to be the root of all evil. One natural way of understanding their relation

is that radical evil is the pathological propensity to fall on one side of the natural dialectic: for our

inclinations and against pure practical reason.

On Kant's view, moral philosophy is practical because it can counteract the natural dialectic

within us and our propensity to 'radical evil.' In contemporary terms: moral philosophy done well is

therapeutic. Moral philosophy done well brings us closer to who we really are as (noumenal) rational

beings; done well, it helps us counteract the destructive tendencies we bring with us as finite,

desiring beings (our 'radical evil'); done well, it gives us the resources to better understand ourselves,

our struggles and confusions about our what we desire and what is good (the 'natural dialectic').

Thus, we can also see how moral philosophy is situated within the enlightenment project: as

other scholars have previously noted, our 'radical evil' is also what keeps us in a 'self-incurred' state

of unenlightenment.[43] In the most general way, the process of enlightenment is the clearing of a

darkness.[44] All enlightenment thinkers saw themselves as fighting the 'darkness' of prejudice,

---

[42] A significant upshot of this view about the sources of darkness that keep us in a state of un-enlightenment is that Kant – unlike many of his predecessors and contemporaries – does not think that the primary problem is bad moral philosophy. Although their theories might ultimately promote the rationalization of our selfish inclinations and our subordination of the moral law to our self-interest, the most significant source of un-enlightenment is within all of us, as a part of human nature. Thus, unlike many other writers in the history of western moral philosophy, Kant also goes to some length in order to praise his opponents' moral character (e.g., David Hume and Joseph Priestley at KrV A745-6/B773-4).

[43] See Schneewind 2010 pp. 296-318

[44] Also the German writers of the Aufklärung (literally: clearing up, or elucidation) frequently invoked the metaphor of the sun dispersing clouds as a symbol for the philosophical-political project of their time. For instance, both Christian Wolff's *Vernünftige Gedancken von der Menschen Thun und Lassen* and his *Vernünftige Gedancken von Gott, der Welt und der Seele des Menschen* show a cover page where beams of sunlight dispel a thicket of dark clouds. For discussion of the sunlight metaphor see Schneewind 2010 p. 296. For a critical discussion of the origins of the term 'Enlightenment' as the designation of a particular philosophical era, see especially Schmidt 2003 and 2006. For a complementary discussion of the German term 'Aufklärung' see Pütz 1978 pp. 12-15 and Schmidt 1992.

superstition and error. But regarding the sources of this darkness, opinions varied. John Locke, for

instance, thought that people's laziness and stupidity were the problem, but also believed that this

could be remedied through education and authoritative instruction;[45] while Hume thought the

problem was religious superstition, which would eventually be dispelled by the growing knowledge

of science.[46] By contrast to many of his predecessors and contemporaries, Kant did not think that

religion per se was the source of darkness that prevented people from trusting their own reason.[47]

Rather, Kant claimed that enlightenment is the exit of our *self-incurred* immaturity (selbst *verschuldete*

Unmündigkeit) (WA 8:35). In other words, Kant thought that the sources of darkness that keep us

in a state of un-enlightenment are within us. But Kant also clearly did not believe that it was the

alleged stupidity of laypeople that kept them in a state of un-enlightenment. For we have already

seen that Kant, following Rousseau, had tremendous respect for common (moral) knowledge. On

Kant's famous definition, the 'immaturity' or 'minority' that characterizes a state of un-

enlightenment is the "inability to make use of one's own understanding without direction from

another." (WA 8:35)[48] Thus, immaturity is a self-incurred failure to make proper use of one's rational

faculties. Kant uses the term 'selbstverschuldet' three times in his *Enlightenment* essay when defining

his notion of enlightenment, and clearly believed this qualification to be important. Thus, critics of

---

[45] In his *Essay concerning human understanding*, Locke puts his view in stark terms: "Now that there is such a difference between men, in respect of their understandings, I think nobody, who has had any conversation with his neighbours, will question: though he never was at Westminster Hall, or the Exchange on the one hand; nor at alms-houses, or Bedlam on the other! Which great difference in men's intellectuals, whether it rises from any defect in the organs of the body, particularly adapted to thinking; or in the dullness or untractableness of those faculties, for want of use; or, as some think, in the natural differences of men's souls themselves; or some, or all of these together, it matters not here to examine: only this is evident, that there is a difference of degrees in men's understandings, apprehensions, and reasonings, to so great a latitude, that one may, without doing injury to mankind, affirm, that there is a greater distance between some men, and others, in this respect, than between some men and some beasts." (Locke 1997 pp. 625-626) For a detailed discussion of this topic in Locke see Spellman 1988 and Marshall 1994.
[46] See for instance Hume 2007 pp. 134-136.
[47] Of course, Kant was strongly opposed to what he called religious fanaticism. However, Kant was clearly not opposed to religion in general.
[48] See also Anth 7:208-10 and Refl 15:230

the enlightenment quickly picked up on it too. In a famous letter to the philosopher Christian Jacob Krauss, Johann Georg Hamann complained about the "damned" description. In his letter from the 18th December of 1784 – the same month in which Kant's *Enlightenment* essay had earlier been published – Hamann protested that "there lies a fundamental error (πρωτον Ψευδος) in the damned adjective *self-incurred*."[49] Because, Hamann argued, "inability is really no fault, like our Plato himself recognizes, but only turns into guilt through the will and its lack of courage and resolution."[50] How this 'inability' might be self-incurred through a lack of courage and resolution, Kant does not answer in his short essay on enlightenment.[51] In order to understand how immaturity can be self-incurred, we have to keep in mind Kant's discussion of the natural dialectic in the *Groundwork* and his discussion of radical evil in the *Religion within the boundaries of mere reason*.[52] The darkness from which we need enlightenment, the self-incurred immaturity, is our propensity to rationalizing away the requirements of reason. Moral philosophy is therapeutic when it helps us overcome this source of darkness and helps us in our process of enlightenment.

Moral philosophy as therapy toward enlightenment is also an essentially 'critical' project. At the end of Kant's description of the practical purpose of moral philosophy in the *Groundwork* (4:404), he claims that the therapeutic purpose of moral philosophy is just one aspect of the

---

[49] Hamann 1965 p. 289, translation mine.
[50] Hamann 1965 p. 289
[51] In his letter, Hamann ceased on this missing explanation with much passion: "With what conscience can a raisonneur and speculator, behind stove and in his pajamas, accuse the patronized of cowardice when their unthinking guardian has a well-disciplined and vast army to vouch for his infallibility and orthodoxy." (Hamann 1965 p. 290, translation mine)
[52] Much of the literature on Kant and enlightenment does not focus on the question of why the darkness before enlightenment is self-incurred. Rather, the literature often concerns itself with the political institutions that have helped maintain the darkness, and with the (political) means against these institutions, such as free speech. (See for instance Bartuschat 2009. See also Brandt 2018, who takes a Straussian reading and argues that Kant's enlightenment essay contains a subversive and encrypted message against feudalism.) Since Kant's most famous essay on enlightenment deals specifically with free speech and the 'public' vs 'private' uses of reason, this focus is not surprising. (For an overview of the historical and political context of Kant's famous essay on enlightenment see Schmidt 1989.) However, this should not distract from the importance of Kant's other writings on the reasons for the self-incurring of un-enlightenment. For a notable exception, which directly links Kant's understanding of enlightenment and his account of radical evil in the *Religion*, see Schneewind 2010 pp. 296-318.

'complete critique of our reason,' i.e., including speculative reason. In other words, just like Kant's theoretical philosophy, his practical philosophy is ultimately a defense of the faculty of reason against the dialectics within us that keep us in a state of un-enlightenment. By keeping in mind the limits of the proper use of both theoretical and practical reason, we might defend reason against vicious dialectics that arise from it. In the theoretical realm, the threat is spurious metaphysics; in the practical realm, the threat is spurious rationalizations. For Kant, this general theme of a defense of reason through critique is the essence of enlightenment, which is not just concerned with the practical sphere but the use of reason – with thinking going well for beings like us – in general:[53]

> One readily sees that while enlightenment is easy *in thesi*, *in hypothesi* it is a difficult matter that can only be accomplished slowly; for while not being passive with his reason but always being legislative for himself is something that is very easy for the person who would only be adequate to his essential end and does not demand to know that which is beyond his understanding, nevertheless, since striving for the latter is hardly to be forbidden and there will never be lacking many who confidently promise to be able to satisfy this desire for knowledge, it must be very difficult to maintain or establish the merely negative element (which constitutes genuine enlightenment) in the manner of thinking (especially in that of the public). (KU 5:294)

To read human frailty and radical evil as one of the sources of darkness that keeps people from enlightenment – i.e., as the sources of self-incurred immaturity – also finds support in Kant's remarks about immaturity in his *Anthropology from a pragmatic point of view*. There, Kant distinguishes between different kinds of immaturity, the most important of which are the natural immaturity of children and the legal immaturity of people in their civil affairs. The more interesting kind of immaturity, however, is self 'made' immaturity. And, importantly, the reason Kant gives for people's

---

[53] See WA 8:35 and WDO 8:146n.

common tendency to make themselves immature is the comfort this state provides. Consequently, this tendency can also be exploited by those in power:

> [T]o make oneself immature, degrading as it may be, is nevertheless very comfortable, and naturally it has not escaped leaders who know how to use this docility of the masses (because they hardly unite on their own); and to represent the danger of making use of one's own understanding without the guidance of another as very great, even lethal. (Anth 7:209)

Thus, the practical purpose of Kant's moral philosophy is to play a role in the enlightenment's defense of reason. And because the clouds which this enlightenment is supposed to dispel are within us, a natural dialectic tempting us to rationalize against the requirements of practical reason, the practical purpose of Kant's moral philosophy is best characterized as therapeutic. Kant's moral philosophy can attempt this therapeutic project precisely because his conception of moral philosophy is archeological, he is an epistemic egalitarian about moral knowledge, and believed that moral judgment can be immediate without recourse to procedural deliberation. We do not need to be taught or instructed what morality requires; we all have equal, immediate access to the requirements of practical reason. Even when we rationalize against them in our head (like Kant believed Garve had done), they are still in our 'heart,' because the form of practical reason is something that we necessarily always have 'before our eyes' metaphorically speaking. As a therapeutic exercise, moral philosophy does not need to instruct people authoritatively about what to do. Rather, moral philosophy is an attempt to help people see something within them that was there all along, and help them overcome the struggle of the natural dialectic.[54]

---

[54] This is not to say that our finite human nature, which leads us into a natural dialectic and is the sources of radical evil, is the only potential hinderance to moral action. Kant also leaves room for a kind of despair, where agents come to fear that all their moral efforts are in vain, given the ubiquity of vice and indifference in the world, or in light of the seemingly insurmountable structural challenges they face. Kant devotes a good amount of effort in several essays on arguing for the stringency of a belief in the moral progress of humanity. However important Kant believed the possibility of moral progress (and our hope in a better future) to be *for* his moral philosophy, his work on this topic does not constitute a discussion of the purpose *of* his moral philosophy. For Kant's most explicit discussion of the importance

Seeing Kant's ethics in this light finally brings us back to the posthumously published note mentioned at the beginning of this chapter. After voicing his heartfelt skepticism about the point of doing philosophy in general (cited above in the introduction), Kant briefly hints, rather enigmatically, at one such purpose.

> I am myself by inclination an investigator. I feel a complete thirst for knowledge and an eager unrest to go further in it as well as satisfaction at every acquisition. There was a time when I believed that this alone could constitute the honor of mankind, and I had contempt for the rabble who know nothing. Rousseau brought me around. This blinding superiority disappeared, I learned to honor human beings, and I would find myself far more useless than the common laborer if I did not believe that this consideration could impart to all others a value in establishing the rights of humanity. (HN 20:44)[55]

To do philosophy is to provide something valuable for others that will help them establish the rights of humanity. Kant makes a similar remark in his notes on anthropology, scribbled to the margins of his copy of Baumgarten's *Metaphysica* sometime between 1772 and 1777: "It is well possible that human beings are ruled as slaves or minors through coercion, status and prejudice; but all these evils must come to an end one day, and philosophy must provide the principles for this, if it should have any use at all." (Refl 15:230, translation mine). If our ('radically evil') pathological response to the natural dialectic keeps us in self-incurred immaturity, then moral philosophy can provide a therapy toward enlightenment and thereby help 'establishing the rights of humanity'.[56]

---

of hope, see his *Orientation* essay. For a recent discussion of the importance of the ability to reconcile moral ideals with lived experience in order to avoid despair and a breakdown in moral agency see Brender 2004.

[55] Also in Rischmüller 1991 pp. 37-39, translation from *Notes and Fragments* p. 7.

[56] As Callanan has pointed out, this threat of the natural dialectic is broader than moral skepticism, since the natural dialectic *could* lead to declaring morality a phantom of the brain. But it could e.g., also lead to us paying lip service to morality while making exceptions for ourselves. (Callanan 2019 p. 7) Thus, while Kant is not interested in responding to the moral skeptic on their own terms (by answer the question 'why be moral?') he arguably identifies an alleged source of moral skepticism in the natural dialectic and a potential therapy against it.

If moral philosophy can be practical in this way, then its therapy also shares our final end. On Kant's view, this final end is not merely 'being virtuous' but a union of moral perfection and complete happiness. This final end is the Highest Good, an idea of reason that is unattainable in our current life but still provides the end for which we may hope and toward which we may work. As we saw, to reach this final end we need no test or decision-procedure for figuring out what is truly right and wrong, but we need therapy that brings us closer to our noumenal self. And the collective, everyday work toward the Highest Good, the union of moral perfection and complete happiness, is the process of enlightenment.

How Should We Read Kant's Therapeutic Ethics?

Kant's view about the practical purpose of moral philosophy is intimately linked with his conception of moral knowledge. Since moral philosophy is supposed to be archeological, moral knowledge distributed equally, and moral judgment immediate, it would be misleading to characterize the practical purpose of Kant's moral philosophy as aiming to "discover what that concept [of a right action] applies to, that is, which actions are right."[57] Moral philosophy might help us gain more explicit and detailed understanding why some actions are right and others wrong, but in that sense moral philosophy is primarily of theoretical interest. By contrast, Kant's conception of moral knowledge opens the door for a therapeutic purpose of moral philosophy. If both the form of practical reason as well as its counterweights are within us, then moral philosophy can be an attempt to help us overcome a natural dialectic that arises from our nature as finite, rational beings, and to

---

[57] Korsgaard 1996a p. 47. Korsgaard's remark is about the *Groundwork*. If my analysis is correct, then that is neither the purpose of the *Groundwork* nor of any other of Kant's works on moral philosophy.

help us follow that inner compass that we always have 'before our eye' through the form of practical reason.

Taking Kant's conception of moral knowledge and the therapeutic purpose of his moral philosophy seriously has significant upshots for how we should read Kant's ethics. For instance, it lends further support to interpretations of Kant's infamous four 'derivations of imperatives of duty' in section II of the *Groundwork* that do not read these examples as providing a decision-procedure for figuring out what the right thing to do would be in a given situation. As Ido Geiger has pointed out, the four examples of duties in section II of the *Groundwork* all concern an agent struggling with the demands of morality. Thus, if the formulas of the categorical imperative are to serve some practical purpose, they may more plausibly be read as "provid[ing] a heuristic tool for moral self-criticism and so for fighting temptation."[58]

While this upshot of Kant's conception of moral knowledge and the therapeutic purpose of moral philosophy might make Kant's ethics appealing, taking its therapeutic purpose seriously also creates serious challenges. First and foremost, how is this therapy supposed to work? Specifically, what role does practical deliberation play in moral knowledge as well as in the therapy that Kant's moral philosophy is supposed to make possible?[59] And how far can such therapeutic reflection take us? Even if one is sympathetic to Kant's conceptions of moral knowledge and his analysis of practical reason, one might seriously doubt that moral philosophy could have any significant therapeutic impact on its readers. Christian Garve, for instance, clearly did not feel the therapeutic

---

[58] Geiger 2010 p. 286. Similarly, Samuel Fleischacker has pointed out that: "Never does Kant take as the paradigm moral question, "what should all people do?," or "what should so-and-so over there do?" The question is always, "what should *I* do?," or, more strictly, "how can I convince myself of what I ought to do?" For the agent in the examples is never really in doubt as to the nature of the appropriate moral laws […]" (Fleischacker 1991 p. 263)
[59] I discuss this question in detail in "The Role of Reflection in Ethical Life" (chapter 2).

value of Kant's moral philosophy. And not all anecdotes are as funny as two academics confounding each other's hearts and heads. In August of 1791, 22-year-old Maria von Herbert wrote to Kant for solace. Having read much of Kant's work, von Herbert had become convinced that the morally right thing was to disclose to a man she loved that she had been in a previous relationship (upon which he abandoned her). Von Herbert deeply admired Kant's work, and turned to the aging philosopher for counsel in this personal matter in three moving letters between 1791 and 1794.[60] Kant's reply, written in spring 1792, is both careful and empathetic. To von Herbert's last two letters, Kant did not respond. Perhaps, he was unable to properly deal with the desolation and depression which both letters display. And perhaps, he was all too aware of the limited use a Kantian philosopher could be as counsel and consoler. Having put her affairs in order and after a breakfast with friends at her family home in Klagenfurt, Maria von Herbert drowned herself on the 23rd of May 1803.[61]

Of course, not even the staunchest defender of Kant's ethics would argue that a moral philosopher should have been able to save Maria von Herbert from suicide by writing a letter. Indeed, Kant's initial response – and his inability to respond a second time – is perhaps the clearest indication we have of his own, sober estimation of the potential practical use of moral philosophy, as well as its limits. But whatever we may think about the promise of form and therapy at this point, Kant's moral philosophy deserves a closer look before we reach a definitive judgment. Keeping in mind the practical purpose of Kant's moral philosophy as enlightenment through therapy, the

---

[60] Maria von Herbert's letters are particularly moving in the original, since von Herbert did not write consistently in standard German but also used Carinthian (southern Austrian) vernacular, which makes her formulations particularly pithy and endearing to those familiar with the dialect.
[61] For literature on Maria von Herbert and Kant, see especially Ritter 2021, Langton 2012, Langton 1992, and Mahon 2006. For a short biographical note see also Zweig 1999 pp. 578-9.

following chapters are an attempt to delineate both the practical purpose as well as the limits of

Kant's moral philosophy.

Chapter 2

The Role of Reflection in Ethical Life

What Is the Point of Reflection?

What is the point of reflecting about what to do? In a way, nothing could seem more obvious: it is figuring out *what to do.* However, when we take a first look at Kant's ethics, things become much less obvious. On the one hand, Kant often portrays moral judgment as immediate and non-deliberative. These judgments allegedly occur without hesitation, i.e., without the need for a temporally extended reflection.[1] Kant's frequent allusions to the immediate moral judgment of children provide a particularly memorable example. In his 'Fragment of a moral catechism,' Kant describes an imaginary conversation about morality between a teacher and a young student. Kant's imaginary teacher asks the student to imagine that it were entirely up to them to distribute all happiness in the world, and asks the student: "would you keep it all for yourself or would you share it with your fellow human beings?" The student immediately replies: "I would share it with others and make them happy and satisfied too." Soon afterwards, Kant's imaginary teacher asks their student to imagine a situation "in which you could get a great benefit for yourself or your friend by making up a subtle lie that would harm no one: What does your reason say about it?" Again, Kant's imaginary student replies without hesitation: "That I ought not to lie, no matter how great the benefits to myself and my friend might be. Lying is mean and makes a human being unworthy of happiness." (MS 6:480-81) Kant paints a similar picture in his response to Christian Garve in the

---

[1] When speaking of moral judgments as 'immediate and non-deliberative' in this paper, I mean the temporal sense of immediate. Here I am solely interested in the relation between the things a moral agent can say immediately, without hesitation, and the things they can articulate only on reflection. Conversely, by 'immediate and non-deliberative' I do *not* mean 'without thought' or 'without cognitive activity.'

Theory and Practice essay. There, Kant provides one of his most memorable versions of the deposit-example:

> Take the case, for example, that someone is holding in trust something belonging to another (depositum), the owner of which has died, and that the owner's heirs know nothing about it and can never come to know of it. […] and add that the holder of this deposit suffers at this very time (through no fault of his own) a complete reversal of his fortune and sees around him a miserable family of wife and children oppressed by want that he could relieve in a moment by appropriating this deposit; we add further that he is philanthropic and beneficent whereas those heirs are wealthy, hard-hearted and, besides, so thoroughly given to luxury and wastefulness that adding anything to their resources would be equivalent to throwing it into the sea. And we now ask whether, under such circumstances, it can be considered permissible for him to put this deposit to his own use. (TP 8:286)

Instead of answering ourselves, Kant asks us to imagine putting this question "even to a child some eight or nine years old." And he claims: "The one being questioned will undoubtedly answer, No! and, in place of any grounds, will be able to say only, It is wrong! — that is, it conflicts with duty." (TP 8:286)[2] Thus, on Kant's view, usually no conscious process of deliberation is required in order to make moral judgments.

On the other hand, Kant also describes instances of conscious reflection about what to do. And there, Kant seems to suggest that some kinds of conscious reflection are indeed practically valuable. These passages give some intuitive reasonableness to the recurring assumption that, in Kant's ethics, the practical point of reflection is figuring out what to do, i.e., to figure out what is *really* right and wrong.[3] Thus, some authors tend to downplay Kant's insistence on the immediacy of

---

[2] For similar passages see also KpV 5:36, 5:79-80, 5:161-162, SF 7:63, TP 8:284-287.
[3] See for instance O'Neill 1983 p. 404, Brewer 2002 pp. 542 and 549, Herman 1993 pp. 75-78, Sticker 2015 p. 980, Sticker 2016 pp. 87-89. This view also seems to be implicit in Wuerth 2011 p. 161. In her chapter "The Practice of Moral Judgment" (1993), Barbara Herman is trying to make sense of the alleged requirement to reflect and universalize for moral judgment. Like Sticker (2015), she ultimately wants to argue that deliberation is not always in order but only when things become complicated. On her view, agents have implicit moral knowledge of the morally salient features of situations which flag the relevance (or irrelevance) or a categorical-imperative procedure for testing maxims.

moral judgment and take Kant to hold that common moral agents usually discover moral knowledge through a categorical-imperative procedure. The most recent expression of this view is Martin Sticker's, who claims that "Kant describes pre-theoretical moral deliberation explicitly as an agent asking herself questions pertaining to the universalizability of her intended course of action, i.e. as an active use of rational capacities. […] Common agents in their everyday moral evaluations make use of a common universalization test which involves checking if one can will a maxim as universal law."[4] And indeed, one can find a handful of passages that might seem to support this reading. A particularly memorable one comes up in the *Critique of Practical Reason*, in Kant's discussion of the 'typic' of pure practical judgment. There, Kant attempts to explain how actions can be represented as standing under laws, even though these laws are not laws of nature but laws of freedom. And he argues that the idea of a law of nature serves as a type for the appraisal of actions, although merely as regards its *form* (KpV 5:69). Kant then claims that this 'type' of the moral law (namely the one which represents moral laws analogous to laws of nature) can also play a role in common moral judgment.

> The rule of judgment under laws of pure practical reason is this: ask yourself whether, if the action you propose were to take place by a law of the nature of which you were yourself a part, you could indeed regard it as possible through your will. Everyone does, in fact, appraise actions as morally good or evil by this rule. Thus one says: if everyone permitted himself to deceive when he believed it to be to his advantage, or considered himself authorized to shorten his life as soon as he was thoroughly weary of it, or looked with complete indifference on the need of others, and if you belonged to such an order of things, would you be in it with the assent of your will? Now everyone knows very well that if he permits himself to deceive secretly it does not follow that everyone else does so, or that if, unobserved, he is hardhearted, everyone would not straightaway be so toward him; accordingly, this comparison of the maxim of his actions with a universal law of nature is also

---

[4] Sticker 2015 p. 980

not the determining ground of his will. Such a law is, nevertheless, a type for the appraisal of maxims in accordance with moral principles. (KpV 5:69)[5]

In which cases (if any), then, is practical deliberation necessary in order to figure out what to do? And when a conscious process of deliberation is not required for moral judgment, what kind of cognitive act is involved in making such immediate judgments?

One seemingly tempting way to think about Kant's account of moral judgment would be to follow the one-sided interpretations like Sticker's above, and take immediate moral judgment to be a blind exercise of cultural prejudice or habit – which may sometimes lead to the correct outcomes, but for the wrong reasons – and to take the further requirement to reflect to be the way we know right from wrong. This reading is especially tempting for those sympathetic to the procedural model of moral judgment, according to which active deliberation is necessary for judging right from wrong.[6] But no matter how tempting, this view is difficult to square with Kant's deep respect for common moral knowledge and the immediacy of moral judgment.

Kant's seemingly contradictory insistence on both the immediacy of moral judgment and the requirement to reflect might be more appealing than it first appears. For there is an equally tempting view of moral knowledge that is directly opposed to the procedural model just mentioned. On this alternative view, our lives contain moments in which a moral action is good precisely because it does not involve stepping back in order to reflect about one's maxim or 'one thought too many.' Moreover, in some situations stepping back in order to reflect about one's maxim might not just

---

[5] One might also have thought of three passages in the *Groundwork* ("I ask myself…" (4:403), "I ask myself only…" (4:403), "But he still asks himself…" (4:423) as further support for this view. As I argue in detail in 'Kant's Derivation of Imperatives of Duty,' these remarks in the *Groundwork* cannot unproblematically be taken to demonstrate an alleged universalization procedure for the practical deliberation of common moral agents. See also Kant's remark about the role of philosophy as legislator of reason at Anth/Busolt 25:1482.
[6] I outline the distinction between what I call 'procedural' and 'immediate' models of moral judgment in 'Conceptions of Moral Knowledge and the Practical Purpose of Moral Philosophy.'

make the action less admirable but even demonstrate a failure of practical reason. And if practical reason should be able to make such immediate moral judgments without 'one thought too many,' we might be curious to consider what kind of faculty this practical reason would be, what its activity was, and what its form or laws might be that would enable such immediate practical judgments. At the same time, moral deliberation is part of our ethical lives too. Whatever way there is for making room for immediate moral judgment, it must also be able to give a place to thoughtful reflection. And if one wanted to make room for both immediate moral judgment and thoughtful reflection, then one might also want to give some considerations to the relation between these activities of practical reason.

Moreover, the relation between the immediacy of moral judgment and a requirement for moral reflection is not merely an exegetical puzzle but also points to a fundamental question about virtue. Namely, what is the relation between immediate moral judgment and a virtuous disposition to do what is right? In order to make sense of virtue, we should ideally be able to make sense of three related, and individually plausible, notions: (i) complete virtue is exceedingly difficult to achieve – only the most megalomaniac among us would claim to be close to attaining complete virtue – and yet, (ii) the non-virtuous still have enough access, of one sort or another, to moral knowledge in order to be held accountable for their actions, and (iii) moral knowledge is non-contingently connected to being virtuous – genuine virtue cannot merely be a separate capacity added (so to say) accidentally to common moral knowledge and which allows one to endure moral requirements grudgingly or to do what is right merely out of blind habit or custom. Instead, the virtuous person has a special kind of appreciation of the universal and immediate moral knowledge that all of us, somehow, already have some access to. By the end of this chapter, it will be clearer how these three

aspects of virtue relate in Kant's moral philosophy. As I will argue, Kant's account of reflection in

moral knowledge provides one possible way of making sense of virtue.

What Is Reflection, Anyway?

Both Kant's practical and his theoretical philosophy are replete with references to reflection.

On the most general level, 'reflection' in Kant is "an act of consciousness that turns upon itself,

rather than directly upon objects, and involves the comparison or holding together of either

different representations or of representations and their corresponding cognitive powers."[7] Beyond

this rather broad description, Kant uses the term 'reflection' (as well as the German *Überlegung*,

nachdenken and Latin *reflexio*) and its cognates quite indiscriminately.[8]

In his theoretical philosophy, Kant frequently uses 'reflection' to refer to the activity of the

understanding.[9] 'Reflection' also figures in Kant's text as the activity that is common to all

cognition,[10] and more famously, Kant also refers to a special type of 'transcendental reflection' in the

amphiboly chapter of the *Critique of Pure Reason*. In 'transcendental reflection,' one compares a

representation with its cognitive origin. "The action through which I make the comparison of

representations in general with the cognitive power in which they are situated, and through which I

distinguish whether they are to be compared to one another as belonging to the pure understanding

---

[7] Gorodeisky 2021 p. 374.

[8] For different ways of distinguishing between Kant's uses of 'reflection' see Merritt 2015 and 2018 (distinguishing between six uses), Birken-Bertsch 2015 (distinguishing between four) and Gorodeisky 2021 (also distinguishing between four uses). Nothing in the argument of my paper depends on the exact number of uses we identify.

[9] MSI 2:393-4, Prol 4:288, Refl 425, 15:171, Refl 3917, 17:342, KrV A310/B366-7, Refl 2834, 16:536; Refl 230, 15:88; Anth 7:138, Anth 7:142, V-Lo/Wiener 24:909, V-Lo/Wiener 24: 904-909, V-Lo/Wiener 24:914, Refl 2389 16:340; Refl 2860, 16:549; Refl 2865, 16:552; Refl 2876, 16:555; Refl. 2878, 16:556, KrV A260/B316, Log 9:93.

[10] Refl 650, 15:287, V-Lo/Blomberg 24:130, Refl 2851 16:546.

or to pure intuition, I call transcendental reflection." (KrV A261/B316–317)[11] On Kant's view, such

transcendental reflection is required to prevent metaphysical blunders because it allows us to

distinguish between claims about objects of the senses (phenomena) and objects of the intellect

(noumena). In a similar vein, Kant also speaks of 'reflection' as a normative requirement on thought

more generally. In his lectures on logic, Kant refers to such a normative requirement on thought

when he claims that "more mature reflection [Überlegung], makes a judgment that is more correct

and more fitting concerning the probability and improbability of a cognition." (V-Lo/Blomberg

24:145) In the same vein, Kant speaks of 'reflection' as an antidote to prejudice. (V-Lo/Blomberg

24:161ff and 24:167-71)[12] Indeed, Kant repeatedly characterizes prejudice as judging without (or

prior to) reflection [Überlegung] (V-Lo/Blomberg 24:162 and V-Lo/Blomberg 24:165).[13]

Kant also speaks of reflection (often in its synonym 'Überlegung') in his practical

philosophy. There, he talks of reflection (Überlegung) as a requirement for becoming virtuous.

Unsurprisingly, this use of the term occurs primarily in the *Doctrine of Virtue:* "Now the capacity and

considered resolve [überlegte Vorsatz] to withstand a strong but unjust opponent is fortitude

(fortitudo) and, with respect to what opposes the moral disposition within us, virtue (virtus,

fortitudo moralis)." (MS 6:380)[14] Here, our judgment of what is right and wrong and our moral

disposition are already in place – but a further, reflected resolve is required to withstand temptations

to act contrary to those demands of morality.

---

[11] See also KrV A262-3/B318-9, KrV A260–262/B316–319, KrV A268f/B325f, and KrV A 294f/B351.
[12] Compare also V-Lo/Dohna 737ff.
[13] In its most colloquial (and contemporary) sense, 'reflection' and its cognates also occur as the activity of thinking something through. (See e.g., V-Lo/Blomberg 24:20.) In this use, 'reflection' seems to have a broader meaning than when Kant speaks about reflection as a normative requirement as an antidote to prejudice. However, Kant uses the actual "reflexio" or "reflexiones" rarely in this context, and prefers to use the German 'überlegen' or 'Speculation,' (see e.g., Eleutheriology, 8:454) which are usually translated as 'reflecting' and 'reflection' in English editions.
[14] See also MS 6:383-4, MS 6:409 and KpV 5:118.

And in at least two further instances, Kant also mentions 'reflection' specifically as the cognitive attention to our own moral judgment. Thus, the section title of §16 in the *Doctrine of Virtue* reads "On an amphiboly in moral concepts of reflection [der moralischen Reflexionsbegriffe], taking what is a human being's duty to himself for a duty to other beings" (MS 6:442). In general, in an amphiboly one is misinterpreting something that lends itself to ambiguity. For Kant more specifically, an 'amphiboly' is an ambiguity about the cognitive source of our concepts of reflection. A 'concept of reflection' (Reflexionsbegriff) is a concept describing the relation between our representations when we reflect about these representations – rather than describing a relation between sensible objects and their properties. 'Duty' is such a moral 'concept of reflection,' because it refers not to a sensible object but to a representation, namely the representation of an action as morally right. Thus, in this section of the *Doctrine of Virtue*, Kant argues that "[a] human being can therefore have no duty to any beings other than human beings; and if he thinks he has such duties, it is because of an amphiboly in his concepts of reflection, and his supposed duty to other beings is only a duty to himself." (MS 6:442) In such an amphiboly of moral concepts of reflection, one misinterprets the cognitive sources of one's representations, i.e., which cognitive faculty they derive from. Here, we misinterpret our 'moral concepts of reflection.' Our 'moral concept of reflection' *duty* is applied to our representation of treating animals well as morally right. But this moral concept of reflection, duty, is then mistakenly interpreted as arising from the status of the animal rather than from our own moral status.[15]

In what way all these cognitive acts of 'reflection' are precisely related is a question that I want, for the most part, to leave aside for now. Instead, I want to emphasis the following two

---

[15] Kant also talks about reflection in connection with 'reflective judgment' in the third *Critique* (see especially KU 5:179, EEKU 20:211-218, EEKU 20:249 and KU 5:190f). For the purpose of the present discussion, we can leave this type of reflection aside.

modest points. First, all these kinds of 'reflection' are instances of "an act of consciousness that turns upon itself" – whether in 'moral concepts of reflection,' 'reflected resolve' or 'transcendental reflection' – rather than directly onto an external object. In this way, 'reflection' is intrinsically linked to self-conscious activity. As Stephen Engstrom notes, 'reflection' in Kant's must not be understood as a cognitive act that presupposes a prior understanding of the distinction between the self and the self's object. Instead, "[r]eflection in Kant's sense is better regarded as a specific type of contemplation, or attention, through which, in the activity of conscious representing, we become explicitly conscious of that activity itself and its identity, notwithstanding the diversity of representation in which it figures."[16] And second, as Melissa Merritt has perspicuously pointed out, these kinds of reflection outlined above fall into roughly two categories: constitutive and normative.[17]

This distinction between a constitutive element of reflection and a further normative requirement to reflect is fairly straightforward regarding Kant's theoretical philosophy. Insofar as 'reflection' captures the kind of activity of the understanding, reflection is a constitutive element in our cognition of objects. Specifically, Kant claims that all concepts rest on reflection in respect of their universality. In this sense, all employment of concepts constitutively involves reflection. If this kind of reflection fails, there simply is no perception of appearances and no thought about objects. By contrast, transcendental reflection and reflection as an antidote to prejudice are examples of

---

[16] Engstrom 2015 p. 24. Melissa Merritt has similarly noted this aspect of Kant's sense of 'reflection' but includes it as another (fourth, in her classification) use of 'reflection' in Kant: namely as the use of 'reflection' referring to "the self-consciousness that is internal to the activity of thinking or makes it 'possible'." (Merritt 2015 p. 479 and 2018 p. 15.) However, Merritt does not cite any specific passage that would indicate such use of the term in Kant, and I tend to agree with Engstrom that the link with self-consciousness should be understood as essential to Kant's notion of reflection in general, rather than a separate use or meaning of the term in his work.
[17] Merritt 2015 and 2018 labels them 'reflection-c' and 'reflection-n' respectively. Merritt also argues compellingly that Kant's distinction between pure and applied logic tracks the different requirements of reflection. While pure logic is concerned with the laws or principles that are constitutive of thinking, applied logic is concerned with the laws of how to make good use of one's thinking. See especially Merritt 2015 pp. 478-489.

normative requirements of thought. If those kinds of reflection fail, one has not thought well. In the amphiboly chapter of the first *Critique*, Kant complains about Leibniz's failure to engage in 'transcendental reflection,' which allegedly led him into significant metaphysical errors – but Leibniz undoubtedly had intelligible thoughts, otherwise there would not have been much to criticize. Similarly, those who are subject to prejudices and thereby judge without reflecting are thinking something and can be criticized for what they are thinking. In these instances, reflection is normatively required, because attention to our own cognitive capacities and their exercises can prevent us from making mistakes.

Intuitively, we might be tempted to think about reflection in theoretical cognition merely as something that occurs only as a thought process of thinking-something-through, when we are induced (in some way or another) to question whether our passive taking something to be so was in fact well-founded. However, Kant's critical philosophy – and indeed, the tradition following Kant – thought it was a momentous mistake to conceive of any type of cognition as purely passive. Thus, from a Kantian perspective, the taking-something-to-be-so and a consciously reflective act of question whether this taking-to-be was well-founded are *both* acts involving the same cognitive capacities. At the same time, it is of course not the case that a simple taking-something-to-be-so is a conscious act of reflecting, or making a deliberate judgment. This is the peculiar sense of 'spontaneity' that Kant ascribes to the exercise of our rational capacities – "a modality of conceptual activity other than assertoric judging or acts of conceptual sorting ("This is an A, this is a B, etc."), or deliberate rule-following."[18] Thus, not only is there a distinction between a constitutive act of reflection in cognition (of taking this thing over there to be so) and a further, normative requirement

---

[18] Pippin 2013 p. 386

to reflect (of reconsidering, in some circumstances, whether we should really judge things to be so); there is also an intimate relation between these two acts of reflection in cognition: the further, normatively required act of reflection involves the same cognitive capacities and may even consciously take into account the form, or presuppositions, of the exercise of those capacities.

If Kant's theoretical philosophy is concerned with two different kinds of reflective acts (constitutive acts of reflection and normatively required acts of reflection), why not think that something similar should be the case in Kant's practical philosophy? After all, Kant believed that theoretical and practical reason are merely different applications of "one and the same reason" (KpV 5:121). Moreover, Kant himself draws an analogy between the formal contributions in practical and theoretical cognition: "this [a will pure and practical of itself, which contains the supreme condition, in accordance with reason] is roughly like the way in which concepts of the understanding, which by themselves signify nothing but lawful form in general, are added to intuitions of the world of sense and thereby make possible synthetic propositions a priori on which all cognition of a nature rests." (GMS 4:454) Practical cognition, like theoretical cognition, relies on a specific form that makes it so much as possible. Kant again repeats this analogy in this *Critique of Practical Reason*.

> These *categories of freedom* [good and evil] – for this is what we are going to call them in contrast to those theoretical concepts which are categories of nature – have an obvious advantage over the latter inasmuch as the latter are only forms of thought which, by means of universal concepts, designate only indeterminately objects in general for every intuition possible for us; the former, on the contrary, are directed to the determination of a free choice […]; hence, instead of the form of intuition (space and time), which does not lie in reason itself but must be drawn from elsewhere, namely from sensibility, these, as practical elementary concepts, have as their basis the *form of a pure will* as given within reason and therefore within the thinking faculty itself […]. (KpV 5:65-66, original emphasis)

As I argue the remainder of this essay, once we get the relation between the constitutive and normative elements of reflection in view, we will be able to see how there can be both immediate

moral judgment as well as a normative requirement to reflect, such that virtue is a reflected (überlegter) resolve to do what is morally required, and also what this further requirement to reflect amounts to.

Two Kinds of Reflection in Moral Knowledge

Despite Kant's use of the analogy between practical and theoretical cognition, there is also an obvious disanalogy between them: the former concerns external objects or *what is*, whereas the latter does not concern any such external objects but what *ought to be.* The former kind of cognition plausibly requires some constitutive act of reflection, an act of consciousness that allows us to have a unified representation of appearances, before we can start consciously reflecting about what it is that we are perceiving. Thus, it makes sense to parse our activity of theoretical reason into two kinds of reflection – an immediate act of reflection making a unified representation of appearances possible, and a further, higher-order requirement to reflect on the sources and nature of this act. But does this distinction really have any analogy in the practical? On a popular procedural view of moral judgment, our knowledge of an action as morally required would just result from our thinking-it-through. On such a procedural view, there is no room for an adequate, immediate moral judgment and a further higher-order requirement to reflect; there is only either blind prejudice or deliberating about what to do. However, as we saw, Kant was remarkably confident about the adequacy of immediate (common) moral judgment. And when Kant describes the need for a priori laws that make representations of what ought to be so much as possible, he often does so in a way that resembles the immediacy of empirical judgment much more than a deliberate thinking-something-through.

Now, since the concepts of good and evil, as consequences of the a priori determination of the will, presuppose also a pure practical principle and hence a causality of pure reason, they do not refer originally to objects (as, say, determinations of the synthetic unity of the manifold of given intuitions in one consciousness), as do the pure concepts of the understanding or categories of reason used theoretically; instead, they presuppose these objects as given; they are rather, without exception, modi of a single category, namely that of causality, insofar as the determining ground of causality consists in reason's representation of a law of causality which, as the law of freedom, reason gives to itself and thereby proves itself a priori to be practical. However, since actions on the one side indeed belong under a law which is no law of nature but a law of freedom, and consequently belong to the conduct of intelligible beings, but on the other side as also events in the sensible world yet belong to appearances, the determinations of a practical reason can take place only with reference to the latter and therefore, indeed, conformably with the categories of the understanding, but not with a view to a theoretical use of the understanding, in order to bring a priori the manifold of (sensible) intuition under one consciousness, but only in order to subject a priori the manifold of desires to the unity of consciousness of a practical reason commanding in the moral law, or of a pure will. (KpV 5:65)

Given Kant's confidence in immediate moral judgment and his characterization of the a priori laws making moral judgment possible, he appears to endorse both a constitutive as well as a higher-order act of reflection in moral knowledge. How, then, should we best describe these acts of reflection?

The constitutive element in moral judgment is the form of our representation of actions as intrinsically good: the categorical imperative.[19] Practical agents represent actions as intrinsically good through subjective principles (which Kant calls 'maxims'), and these principles representing an action as intrinsically good always have a specific formal structure. Any representation of a determinate object has both matter and form, without which no determinate object could be so much as thought. Thus, any representation of an action or end as intrinsically good also has a specific form, which we find in the structure of the subjective principle for action (the maxim).

---

[19] Kant frequently emphasizes that the categorical imperative is just the form of pure practical reason. GMS 4:400, 403-4, 416, 421, 426, 428, 436-7, 444, 458, 461-2.

According to Kant, the special form of the maxims representing actions and ends as intrinsically good is universality. In other words, Kant's thought is that only by representing actions through a principle with a specific form (namely, the form of universality) can we represent them as intrinsically good. Without this formal structure, we would simply not have maxims through which we could represent actions as to be done from duty. If the categorical imperative is just the form of pure practical reason, then representing actions in accordance with the categorical imperative is the kind of reflection that is constitutive of moral judgment. Conversely, the immediate moral judgment that we find depicted in so many passages of Kant's ethics is the (pre-theoretical) reflective awareness of representing an action as intrinsically good. We know 'in our hearts' that certain maxims are moral because having such maxims just *is* to represent an action or end as intrinsically good. Consequently, our representation of an action as intrinsically good is also inseparable from the universalizability of its maxim: the universal form of the maxim is constitutive of our representation of the action as intrinsically good. Thus, theoretical awareness of the universalizability of some maxim is just a higher-order version of the immediate awareness of the maxim's morality already contained in the representation of the respective action as intrinsically good.[20]

If Kant thought that the form of pure practical reason enables us to make the immediate moral judgments that his imagined conversations with students and children so vividly depict, why bother about any additional, higher-order requirements to reflect? In order to see what good such further reflection is for, consider what could go wrong. Kant tells us most explicitly about the

_____

[20] This leaves open the question how we know which actions should be represented through maxims of a universal form. To put it bluntly, having a good will does not mean that one also ends up doing any good. As I argue in chapter 3, Hegel's empty-formalism charge consists in arguing that Kant's ethics does not give us the resources to answer this question; a (partial) defense of Kant will consist in arguing that Kant simply did not have the ambition to answer this question (as I argue in chapters 4-7).

possibility of moral judgment being clouded, and how we might lighten it up again, at the beginning of the *Groundwork*. In a passage emphasizing the practical need for moral philosophy, he explains:

> The human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect - the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name happiness. Now reason issues its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with disregard and contempt for those claims, which are so impetuous and besides so apparently equitable (and refuse to be neutralized by any command). But from this there arises a natural dialectic, that is, a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity […]. (GMS 4:405)

According to Kant, our inclinations are both adamant and 'so apparently equitable.' Thus, what clouds our moral judgment is not simply a desire but something that is in fact reasonable, to some extent. Indeed, Kant has absolutely no qualm with happiness per se – after all, the highest good is *happiness* in accordance with virtue. In this sense, Kant is far from being a Stoic. Quite to the contrary, he thinks happiness is good, and the happiness of others (as well as our own) is even something we have a moral duty to promote. Since our inclinations appear so 'apparently equitable', and since reason does not 'promise anything to them,' a natural dialectic arises which tempts us to rationalize away the requirements of morality. This is not simply to say that we can be overpowered by inclinations to do what we believe to be wrong; rather, the natural dialectic is itself a rational process. Our happiness is a genuinely valuable end, which makes pursuing our happiness also appear so 'equitable.' Moreover, the experience of our judgment about what is intrinsically good and what is furthering our own happiness coming apart is naturally disconcerting. This is not merely the same as the frustration of failing to satisfy some desire; it is the implicit realization that something is wrong about the way I desire, that what I judge intrinsically good and what I am inclined to have come

apart. And since it is hardly an option for finite, dependent beings like us to shape our inclinations at will, we take a pathological interest in justifying them. In other words, we are naturally prone to rationalizing in favor of our inclinations and against our moral judgment not because we would be suffering from overly strong inclinations, but precisely because we are *rational* beings who take an interest in what is reasonable.[21]

If we take the problem posed by the natural dialectic seriously, then it also becomes clearer how a higher-order requirement to reflect might help us overcome it. By turning out attention to the source of our moral judgment (i.e., to the form of pure practical reason), conscious reflection might help us overcome the temptation to justify our 'so apparently equitable' inclinations to ourselves and to thereby place our own immediate interest over the demands of reason. Consequently, it would be misleading to think of the practical use of formulas of the categorical imperative in moral reflection as a test or decision-procedure for figuring out what is truly right and wrong. As we already saw in the previous chapter, Kant repeatedly emphasized that common moral agents do not need to engage in any decision-procedure in order to know what is right and wrong, just like they don't need any such thing to know their left hand from their right. Kant's concern is not with the puzzled agent who wants to do what is right but is uncertain about his moral obligation; Kant's concern is with our propensity to rationalize away the requirements of morality in favor of our inclinations.[22]

---

[21] For related discussion of Kant's notion of 'rationalizing' (vernünfteln) see also Sticker 2021, especially pp. 29-30.
[22] The fact that Kant was not concerned with providing a test or decision-procedure to be used in practical deliberation for figuring out what is truly right and wrong also becomes apparent when we keep in mind how Kant's remarkable disinterest in one prominent topic of contemporary Kant scholarship: the problem of maxim description. Kant repeatedly tells us of his concern with 'rationalizing' (vernünfteln). And on the presumption that the formulas of the categorical imperative are supposed to provide us with a test or decision procedure, the paradigmatic example of rationalizing would be to change the description of one's maxim. On this assumption, it would be quite remarkable that Kant does not consider a change of maxim description as a way of rationalizing away the requirement of morality. Of course, it is always possible to think that Kant just failed to consider the obvious. However, in light of what Kant says about common moral knowledge and the natural dialectic, it is much more plausible that Kant simple did not think of moral knowledge as arising through the application of a decision-procedure. (Cf Sticker 2021 pp. 10-11)

But if this is the case, what is happening in the passage of the *Critique of Practical Reason* cited above? Doesn't Kant there describe a situation that sounds much like a kind of self-help exercise, a conscious walking-through a test or decision procedure in order to figure out what is really right and wrong rather than just getting to the core of what we knew all along?

> The rule of judgment under laws of pure practical reason is this: ask yourself whether, if the action you propose were to take place by a law of the nature of which you were yourself a part, you could indeed regard it as possible through your will. Everyone does, in fact, appraise actions as morally good or evil by this rule. Thus one says: if everyone permitted himself to deceive when he believed it to be to his advantage, or considered himself authorized to shorten his life as soon as he was thoroughly weary of it, or looked with complete indifference on the need of others, and if you belonged to such an order of things, would you be in it with the assent of your will? Now everyone knows very well that if he permits himself to deceive secretly it does not follow that everyone else does so, or that if, unobserved, he is hardhearted, everyone would not straightaway be so toward him; accordingly, this comparison of the maxim of his actions with a universal law of nature is also not the determining ground of his will. Such a law is, nevertheless, a type for the appraisal of maxims in accordance with moral principles. (KpV 5:69)

In this part of the *Critique of Practical Reason*, Kant is concerned with the question how we can conceive of the idea of laws of *freedom* in the sensible world. Laws of nature, for their part, correspond to a 'schema,' "that is, a universal procedure of the imagination (by which it presents a priori to the senses the pure concept of the understanding which the law determines)." (KpV 5:69) However, laws of freedom cannot, by definition, apply to intuitions in this way. Consequently, there seems to be no schema through which we can represent a 'law of freedom' in the sensible world. Nevertheless, Kant argues that the idea of a law of nature can serve as a *type* of the moral law, because it captures the form of pure practical reason. Thus, in the passage quoted above, Kant is concerned to demonstrate just this: that the structure or form of common moral judgment takes the form of a universal law, and that the idea of a universal law of nature can serve as a type for

67

representing the idea of laws of freedom for sensible beings. The reflective thought-process that Kant describes here is not supposed to be the source or origin of the moral judgment, but bringing something to the fore that was within us all along.[23] In this way, the typic remains squarely within the original project of moral philosophy set out in the *Groundwork*: to start with common moral cognition, determining its underlying principle, and proceeding to a *Critique*, which would examine if reason is in fact capable of acting in the way our analysis of common moral cognition suggested. And if this project were successful, we would have corroborated Kant's original claim about practical faith in the *Critique of Pure Reason*, that "even the highest philosophy cannot advance further than the guidance that nature has also conferred on the most common understanding" (KrV A830/B858-A831/B859).

Lastly, consider Kant's infamous four examples of 'derivations of imperatives of duty' in section II of the *Groundwork*. These four examples are usually read as paradigmatic examples of how, on a supposedly Kantian view, reflection can help us figure out what is right and wrong. Although I believe that this is a misrepresentation of the purpose of this infamous passage,[24] it is worth asking how Kant portrays the common moral agent in his examples. On the face of it, Kant does not portray the common moral agents as concerned with the question what really is right and wrong in a specific situation. Instead, what stands out about the four agents in Kant's famous examples is that they struggle with their inclinations. "Someone feels sick of life because of a series of troubles that has grown to the point of despair," so much so that the person just has enough strength left to put themselves to reflect about their situation. "Another finds themselves urged by need [durch Not gedrungen] to borrow money." This is a person in despair (Not), not someone unclear about the

---

[23] For a similar reading of the typic, see Westra 2016. Cf Zimmermann 2015.
[24] I discuss section II of Kant's *Groundwork* at length in 'Kant's Derivation of Imperatives of Duty.'

morality of truthfulness. The third is tempted by the comfort of an easy and pleasurable life as opposed to the hassle of developing excellence in their natural gifts. Only the fourth seems to be doing alright for themselves, and is merely tempted by a desire to ignore the hardship of others in order to enjoy their own fortune unimpededly. (GMS 4:421-423) In none of these instances is Kant concerned with the merely puzzled agent who has good intentions but is unclear about which course of action is truly the right one. Rather, Kant's examples are concerned with bringing out the form underlying our common moral cognition in cases where agents are tempted to rationalize away their moral judgment.

Whatever we may think the purpose of Kant's four examples in the *Groundwork* is, Kant's paradigmatic instance of practical questions in moral life concerns agents who are struggling because their moral judgment is clouded by the natural dialectic.[25] What practical reflection (as a conscious thought-process) does for these agents is help them appreciate the very conditions of their immediate moral judgments in the form of practical reason. To come back to the analogy with theoretical cognition, as Robert Pippin has aptly put it: "The examples [of *Groundwork II*] are not meant to reveal what actually goes on in the assessment of a course of action as impermissible, any more than the claim that the form of our perception is judgemental is meant to suggest that perception consists in judgements."[26] The Categorical Imperative as the form of pure practical reason is a constitutive element of our representation of actions as intrinsically good; thus, it need not figure in a conscious process of deliberation in the form of a test or decision-procedure any more than categories of the understanding would need to figure in perception in some process of assertoric judging or conceptual sorting. Rather, for Kant, reflecting on the conditions of our moral

---

[25] See also the discussion of Kant's casuistical questions in the *Metaphysics of Morals* in chapter 7.
[26] Pippin 2013 p. 378

judgments will thus highlight the demands of morality against the temptation of the natural dialectic. To have such a considered (überlegter) resolve to do what is morally right even when it is difficult is virtue.[27]

Making Sense of Virtue

If Kant's account of practical reason contains two kinds of reflection – a constitutive kind of reflection and a further, normative requirement to reflect – why might we find such an account attractive? Independently of the purely exegetical benefit, why should we think about practical reason as involving two different kinds of activities, a cognitive act imposing a form through which represent actions and ends as to be done from duty without any conscious process of deliberation or thinking-through, and a further element of reflecting about the sources of our moral judgments that might help us overcome the temptation to place inclinations above morality? One reason for taking Kant's account of reflection seriously is simply that it helps us make sense of our own ethical life. Both kinds of reflection outlined above are familiar to us. Moral deliberation plays an important role in the life of most of us, and certainly in the lives of all those who aspire to develop virtue. At the same time, there are many instances in which we do not need to deliberate in order to make a confident moral judgment – even more, in some cases it is the immediacy of unclouded moral judgment itself that carries particular moral significance.

---

[27] How much of this further kind of reflection is necessary in order to lead us to a virtue? The short answer is: as much as is needed in order to overcome the natural dialectic that also makes moral philosophy necessary (GMS 4:405 and 4:411n). The long answer – how much *that* might be – I leave aside for now. Note also that if there was only one kind of 'reflection' in moral knowledge – a conscious act of deliberating about the universalizability of one's maxims – then one would have expected Kant's account of virtue to say something about how often we need to apply this procedure. (Indeed, tentative answers recent Kant scholarship to this question seem less like Kant exegesis and more like groping in the dark. See e.g., Sticker 2015 and Geiger 2015.) Virtue as a steadfastness in doing what you ought to do makes sense primarily for agents who already have moral knowledge.

In addition to making sense of our lived ethical experience, Kant's account of reflection in moral knowledge also provides a tempting starting point to think about virtue.[28] For we might think that even we who are not virtuous must have *some* moral cognition that makes it possible for us to act morally and immorally, and to be held responsible for those acts specifically, rather than for our failure or inability to develop our virtue. As mentioned above, three plausible aspects of virtue are that it (i) is exceedingly difficult to fully achieve, that (ii) the non-virtuous still have enough moral knowledge to act morally and immorally and can be held responsible for their actions, and that (iii) moral judgment is non-contingently connected to being virtuous. Only if moral judgment is truly egalitarian can the first two requirements be fulfilled; and if reflection about the conditions of our moral judgments can help us become more virtuous, we are also in a position to see an intrinsic connection between the common moral knowledge of all and the virtue most (if not all) of us are still aspiring to.

Thus, in a final plot-twist, Kant's account of reflection in moral knowledge is, at least in some sense, all about virtue. But there also remains a final caveat: the Kantian conception of virtue outlined here is quite thin. The content of this kind of virtue remains limited by the content of our immediate moral judgment. Kantian virtue is unclouded moral judgment and the reflected resolve to follow it, but the further requirement to reflect on the source of our moral cognition does not, by itself, make this moral cognition more substantive than it already is. In brief: further reflection about the form of practical reason may uncloud our moral judgment, and it may introduce systematicity into our moral cognition,[29] but it is far from obvious that it would add substantive moral knowledge

---

[28] Although I believe that Kant's conception of virtue is indeed tempting, I argue that Kant's account of virtue is, ultimately, too thin to be fully satisfying in 'Kant's Racism as a Philosophical Problem.'

[29] I discuss the way in which Kant's project of a metaphysics of morals introduces systematicity into our moral cognition in more detail in chapter 6.

over and above our immediate moral judgment. The next chapter raises a first worry about this

apparent limitation of Kant's moral philosophy.

Chapter 3

The Formalism of the Categorical Imperative

Introduction

No criticism of the Categorical Imperative continues to shape people's understanding of Kant's ethics like Hegel's empty-formalism charge. Indeed, Hegel's empty-formalism charge is not merely the most famous criticism of Kant's ethics, but also one of the most prominent aspects of Hegel's own ethical thought. Literature on Kant's ethics frequently summarizes Hegel's empty-formalism charge as on a par with J. S. Mill's famous criticism of the Categorical Imperative. On this view, both philosophers criticized Kant for providing an alleged test or decision-procedure for the morality of our actions that turns out to be practically useless because it is overly formal. However, there is a striking difference between their criticisms of Kant: while Mill has little to nothing good to say about Kant's moral philosophy, Hegel's criticisms are consistently accompanied by high praise. Even more, Hegel incorporates Kant's moral philosophy as a necessary step (*Moralität*) into his own philosophical system. This ostensible tension makes Hegel's empty-formalism charge significantly more interesting than most other criticisms of Kant's Categorical Imperative. If Hegel praises Kant for being the first philosopher to adequately capture the form moral judgment in the good will, what was he criticizing? If Hegel was criticizing the uselessness of the Categorical Imperative as a potential criterion for right and wrong action, what was he praising? Given the centrality of Hegel's empty-formalism charge to our understanding of the practicality of Kantian ethics, it is worth parsing through existing interpretations to get a better sense of the potential limits of Kant's practical therapy.

In this chapter, I argue that there is no tension in Hegel's approach. The criticism and praise just mentioned are not just compatible but necessarily go together. Hegel's empty-formalism charge is neither a knockdown argument nor a failed criticism. Instead, Hegel's empty-formalism charge highlights a limitation of Kant's ethics that goes hand in hand with Kant's insights: it is one thing to show that all moral judgments share a particular form in representing actions and ends as intrinsically good, but it is another to determine which actions can and should be represented in accordance with that form. On Hegel's view, Kant was right about the former, but the formalism of the Categorical Imperative alone does not provide the resources to achieve the latter. Whether or not this limitation is a problem for Kant's ethics depends on what we want from it. Historically, it also depends on whether Kant intended to overcome this limitation of the *Groundwork* and the second *Critique* in his later work, and if so, how he intended to overcome it.[1] For the purpose of this essay, I leave this latter question aside and focus on Hegel as an insightful reader of Kant.

Much literature on Hegel's empty-formalism charge has considered Hegel's criticism as a simple yes-or-no criticism. Consequently, literature on Hegel's empty-formalism charge is roughly divided between two ways of interpreting the criticism. According to one interpretation, Hegel's empty-formalism charge is simply a criticism of the universal law formulas as practically useless. And since the alleged practicality of the Categorical Imperative is a major selling point of Kantian ethics today, the empty-formalism charge continues to elicit responses from Kantian moral philosophers. According to a second interpretation, the empty-formalism charge identifies an inconsistency in Kant's critical system: while Kant rejects a formal criterion of truth in the first *Critique*, Kant allegedly adopts a similarly formal criterion for right and wrong action. Against this reading of the

---

[1] I discuss these two questions at length in chapters 5-7.

empty-formalism charge, defenders of Kant have argued that Hegel simply misunderstands Kant's notion of 'metaphysics of morals,' and thus errs in linking the categorical imperative to Kant's 'general logic.'

Both these interpretations of the empty-formalism charge face a considerable difficulty: they inevitably assume that Hegel was either a lousy reader of Kant or significantly misrepresented Kant's views. By contrast to these two interpretations, I argue that it is unnecessary to attribute such grave errors to Hegel's reading of Kant in order to give a compelling interpretation the empty-formalism charge. On the alternative reading I propose, Hegel thought that any formula of the Categorical Imperative would be useless as a criterion for right action because they *correctly* capture the form of moral judgement. On Kant's hylomorphic conception of practical reason, the content of a practical judgment (namely, the specific action) does not, by itself, contradict the Categorical Imperative; on Kant's view, the Categorical Imperative contradicts the form of certain maxims. To anticipate: on this view, mere content does not contradict form; content must first receive some form in order to so much as contradict another formal criterion. But what maxim an action figures in is not determined by the Categorical Imperative. Consequently, the form of moral judgment by itself cannot provide a systematic criterion for right action.

Hegel's empty-formalism charge thus emphasizes a crucial limitation of Kantian ethics.[2] By separating the form of moral judgment from its content, Kant's ethics cannot derive a systematic account of right action from the form of moral judgment alone. If Kant had wanted to provide a systematic account of right action, then the Categorical Imperative would be insufficient.

---

[2] Hegel provides two other important criticisms of Kant. Namely Kant's alleged inability to explain moral motivation and the alleged inexplicability of freedom in the phenomenal realm. In this essay, I focus exclusively on the empty-formalism charge.

Alternatively, if the Categorical Imperative was not meant to be a foundation for a systematic account of right action, then Kant's ethics – and specifically, his *Metaphysics of Morals* – might be less practical than we want it to be.

In order to provide a Kantian response to Hegel's criticism, we may have to reconsider what practical purpose, on Kant's view, moral philosophy was supposed to serve. We would need to pay close attention to Kant's entire system, including his *Metaphysics of Morals* and potentially his work on religion as well as anthropology. Did Kant ever intend to provide a systematic account of right action by means of deriving it from the form of moral judgment (the Categorical Imperative)? And if not, was Hegel right to think that moral philosophy *should* attempt to produce a systematic account of right action, or was Kant right to constrain the ambitions of his moral philosophy? Although these questions go far beyond the scope of this paper, they will emerge as the central issues dividing Kantian and Hegelian outlooks on moral philosophy.[3]

A Criticism of the 'Universalization Test'?

Consistently throughout his major works, Hegel criticizes Kant's approach to moral philosophy as being 'empty,' or too abstract. Perhaps the most often quoted version of this criticism can be found in Hegel's *Elements of the Philosophy of Right*. Hegel also makes similar remarks in his *Phenomenology of Spirit* and his lectures on the history of philosophy, but he gives the most extensive version of the empty-formalism charge in his essay on natural law, written during his early Jena period. There, Hegel portrays Kant's Categorical Imperative as a practical instance of the law of excluded middle. From this, he concludes that the Categorial Imperative would be useless as a

---

[3] I discuss these questions together with the purpose of Kant's *Metaphysics of Morals* at length in chapters 4-8.

criterion for the appraisal of actions. Specifically, Hegel contends that, if it were applied as a test, the

Categorical Imperative allows us to portray any action as permissible – foreshadowing his comments

in the *Philosophy of Right* (§135). In his most famous example, Hegel suggests that taking someone

else's property cannot be ruled out by universalizing the maxim of taking other people's possessions,

but could only be ruled out if we already presuppose the institution of property. In addition, Hegel

also claims that applying the Categorical Imperative as a test would have the absurd result of

rendering some maxims impossible although they involve moral action-types – mirroring his

comments in the *Phenomenology* (§§419-37).[4]

So far, the majority of commentators have understood the purpose of Hegel's empty-

formalism charge as a criticism of the alleged 'universalization test.' On this view, Hegel's main

criticism of the Categorical Imperative would be that it is unsuccessful as a decision-procedure for

moral action. Kant scholars have spent much effort debating how to interpret a possible

universalization test such that its results 'come out right,' while critics continue to look for plausible

counterexamples. Given the centrality of such an alleged universalization test for both Kant

scholarship and popular understanding of Kant's ethics, it is intuitive to interpret Hegel's empty-

formalism charge as a criticism of this universalization test.

Authors who adopt this interpretation of Hegel's empty-formalism charge typically agree that

Hegel's supposed objection to a universalization test is that too much depends on the input of the

test: we only get out what we put in. This interpretation of Hegel's criticism has been spelled out

---

[4] Note that, strictly speaking, the formulations of Hegel's empty-formalism charge do not assert that Kant attempted to provide the categorical imperative as a test or decision-procedure. (In this way, Hegel is a significantly more astute reader of Kant than many contemporary neo-Kantian ethicists.) Rather, Hegel is concerned with the development of a systematic account of our moral knowledge and claims that if the categorical imperative *were* applied as a test for this purpose, then it would be useless. I return to this point in section 3 below.

slightly differently by various authors. For instance, some scholars think that Hegel charges Kant with presupposing the goodness or badness of the consequences of the universalized action. In other words, these authors interpret Hegel's empty-formalism charge as tantamount to J. S. Mill's famous criticism.[5] Similarly, some authors believe that Hegel charges Kant with presupposing the existence and goodness of certain institutions or practices.[6] And some authors understand Hegel's empty-formalism charge as concerning the proper description of maxims; on this view, the Categorical Imperative could only function successfully as a universalization test if we had an independent, prior way of determining the right description of our maxims to be tested.[7]

This interpretation of the empty-formalism charge as directed specifically against a universalization test appears particularly popular with defenders of Kant; numerous Kant scholars have accordingly responded by defending the practicality of the formulas of the Categorical Imperative. For instance, some defenders of Kant have accepted the presumed criticism of a universalization test but argued that Hegel fails to consider the Formula of Humanity (FH), which states "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." (GMS 4:429) Some defenders of Kant also make the more subtle claim that the Formula of Universal Law (FUL) must be interpreted as less abstract because of the equivalence between the different formulas of the Categorical Imperative, or that teleological considerations about human nature can be taken into consideration in the application of both the FUL and FH.[8] These defenders of Kant therefore conclude that

---

[5] Wood 1990 p. 155, Singer 1961 pp. 251-253.
[6] Walsh 1969 pp. 23-24, Acton 1970 pp. 24-25, Wildt 1982 pp. 45-49, and O'Hagan 1987 pp. 140-141.
[7] Hoy 1989 pp. 215-6, Westphal 1991 pp. 158-59, Lottenbach and Tenenbaum 1995 pp. 223-5, and Freyenhagen 2011 pp. 175-180.
[8] Lottenbach and Tenenbaum suggest this as a potential Kantian response to the empty-formalism charge (Lottenbach and Tenenbaum 1995 p. 229). Sedgwick 2011 also proposes a (partial) response to Hegel that focuses on the other formulas of the Categorical Imperative in the *Groundwork* to show that the moral law is not so empty after all.

Hegel's empty-formalism charge is ultimately unsuccessful.[9] Some even have decried Hegel as

"manifestly unfair,"[10] and "careless in understanding his predecessor."[11] Some have also rejected

Hegel's criticism of the FUL as conceptually confused: they claim that Hegel failed to recognize the

difference between a contradiction in willing and the non-existence of an institution or practice. In

other words, Hegel supposedly ignores the fact that the universalizability test applies to maxims of

the will, not to the existence of institutions.[12] Yet other defenders of Kant have focused on the

problem of maxim-description,[13] arguing that once we can demonstrate a correct way of formulating

maxims, then the decision-procedure of the Categorical Imperative might work just fine. And, of

course, some defenders of Kant have adopted a combination of these possible responses.

The Kantian responses just mentioned share two considerable weaknesses. One of these

weaknesses is that they risk portraying Hegel as a surprisingly sloppy reader of Kant.[14] Indeed, some

---

Specifically, Sedgwick argues that the other formulations of the Categorical Imperative provide some of the moral 'content' that is presupposed in an application of the FUL. Similarly, Mark Timmons has argued that the FUL provides for a universalization test, and that the FH fixes the proper description of maxims and thereby overcomes the problem of maxim-description (Timmons 2017 pp. 67ff). (Although Timmons proposes this solution in response to the problem of maxim-description rather than Hegel explicitly, his solution works much like other responses to Hegel's empty-formalism charge.)

[9] Knox 1957 p. 75, Lo 1981, Freier 1992 pp. 317-320, Wood 1990 pp. 156.

[10] Ameriks 2000 p. 313

[11] Lo 1981 p. 198

[12] Singer 1961 pp. 251-253, Wildt 1982 pp. 45-49, O'Hagan 1987 pp. 140-141, Freier 1992 pp. 313-314, Korsgaard 1996 pp. 86-87, and Iwasa 2013 p. 69. Cf. Wood 1999 p. 90 and 1990 p. 157-158, arguing that the 'contradiction' in the commonly debated example concerns not the existence or absence of a practice (i.e., of making of deposits) but the presupposition of mutual trust.

[13] Herman 1993 and Wood 1990 pp. 160-61.

[14] Another Kantian response might augment the FUL not with resources allegedly provided by the FH, but with teleological considerations in general. Although this is not a common move in the literature on Hegel's empty-formalism charge, Lottenbach and Tenenbaum 1995 suggest that this might be a potential route for the Kantian. One difficulty with this response is that it is controversial whether Kant's insistence on the formality of the Categorical Imperative is compatible with such a teleological reading. For a criticism of this kind, see Freyenhagen 2012 p. 46. As Matthias Haase has pointed out, an even bigger problem for the teleological defense of Kant is the following: by the time Hegel considers the concept of morality, the concepts of life and teleology are already on the table in his philosophical system and we should therefore also expect Hegel to make use of them when he has to. Nevertheless, Hegel claims that the Kantian concept of morality is *both* necessary and empty despite having teleology already on the table. Thus, a Kantian response to Hegel which reads teleology into the Categorical Imperative would have the awkward downside of portraying Hegel as a poor reader of Hegel.

of the authors who have taken Hegel's remarks about the Categorical Imperative to imply a confusion between contradictions with the absence of institutions have also expressed rather derogatory views about Hegel, some calling his criticism "almost incredibly simpleminded."[15] However, as I will argue below, drawing such extreme conclusions about Hegel's reading of Kant is just not inevitable.

Another weakness of the Kantian replies outlined above is that they implausibly narrow the scope of Hegel's explicit criticism. By his own light, Hegel is not merely criticizing one of Kant's formulations of the Categorical Imperative – as if Kant had simply failed to identify the correct formal principle that should guide our conduct, or misrepresented this principle through his formulations. By way of criticizing the Categorical Imperative as an empty-formalism, Hegel targets Kant's conception of practical reason in general. In his *Natural Law Essay* (NLE), Hegel explicitly targets the Kantian idea that it is the "essence of the pure will and of pure practical reason to abstract from all content." (NLE p. 124) Similarly, in the *Philosophy of Right*, Hegel criticizes Kant's attempt to ground morality in the 'standpoint of morality.'[16] This discrepancy between Hegel's own description of the empty-formalism charge and narrow interpretations of his criticism provides another reason to be skeptical of reading him as targeting an alleged universalization test.

An Inconsistency Charge?

Another tempting interpretation of the empty-formalism charge is to understand it as a charge of inconsistency. This interpretation tries to understand both Kant's ethics as well as Hegel's

---

[15] Singer 1961 p. 251.

[16] One of the few recent commentators to point this out is Sally Sedgwick 2011 p. 268. See also Wood 1990 pp. 154-173.

criticism as connected to their larger philosophical systems. According to this interpretation, Hegel accuses Kant of ignoring his own insights from the *Critique of Pure Reason*. There, Kant had rejected a formal criterion for truth. As Kant points out, such a formal criterion could not determine the truth of any belief, since a formal criterion would abstract from all objects.

> If truth consists in the agreement of a cognition with its object, then this object must thereby be distinguished from others; for a cognition is false if it does not agree with the object to which it is related even if it contains something that could well be valid of other objects. Now a general criterion of truth would be that which was valid of all cognitions without any distinction among their objects. But it is clear that since with such a criterion one abstracts from all content of cognition (relation to its object), yet truth concerns precisely this content, it would be completely impossible and absurd to ask for a mark of the truth of this content of cognition, and thus it is clear that a sufficient and yet at the same time general sign of truth cannot possibly be provided. Since above we have called the content of a cognition its matter, one must therefore say that no general sign of the truth of the matter of cognition can be demanded, because it is self-contradictory. (KrV A58-59/B82-B83)

Hegel explicitly refers to this passage of Kant's first *Critique* when he formulates the empty-formalism charge in his *Natural Law Essay*. According to Hegel, the Categorical Imperative is formal in the same sense as a formal criterion for truth. And, importantly, the way Hegel formulates his criticism suggests that he thought a true Kantian should therefore reject the Categorical Imperative as criterion for right action:

> […] the pure identity of the understanding, expressed in theoretical terms as the principle of contradiction, remains exactly the same when applied to the practical form. If the question 'What is truth?', when put to logic and answered by it, provides Kant with 'the ridiculous spectacle of one man milking a billy-goat while another holds a sieve beneath it', then the question 'What is right and duty?', put to and answered by pure practical reason, is of exactly the same kind. Kant recognises that a universal criterion of truth would be one which would apply to all knowledge [Erkenntnisse], irrespective of the objects of this knowledge, but it is clear that, since this involves abstracting from the whole content of knowledge (although truth is concerned with precisely this content), it is quite impossible and absurd to ask for a test of the truth of this content' (because this test is supposed to

have nothing to do with the content of knowledge). *With these words, Kant in fact passes judgement on the principle of duty and right which is set up by practical reason.* (NLE p. 124, emphasis added)

Hegel's wording of this passage might suggest that he charges Kant with inconsistency: perhaps Kant had failed to apply this insight from the first *Critique* consistently throughout his critical philosophy.[17] Hegel again draws this parallel between theoretical and practical cognition in his criticism of Kant's ethics in the *Phenomenology*, albeit more briefly. There, Hegel writes: "It would also be very peculiar if tautology, the principle of non-contradiction, which everyone concedes to be only a formal criterion for knowing of theoretical truth, i.e., as something which is supposed to be wholly indifferent to truth and untruth, were for the knowing of practical *truth supposed to be more than that*."[18] (PG §430)

Support for the inconsistency interpretation of Hegel's empty-formalism charge also comes from an unlikely corner: contemporary Neo-Kantian ethics. At least in their rhetoric, many contemporary readers of Kant suggest that the Categorical Imperative can be characterized as a formal principle of the logic of practical reasoning. Thus, the Categorical Imperative has been described as a "principle of the logic of practical deliberation"[19] and a "purely formal constraint."[20] It has also been compared to a law of logic,[21] and to the principle of non-contradiction.[22] At the same time, prominent neo-Kantians like Christine Korsgaard and Onora O'Neill also argue that one can derive substantive implications from this formal principle of practical reasoning.[23] Thus, many contemporary neo-Kantians ascribe, at least implicitly, an asymmetry to Kant's critical philosophy:

---

[17] Knox 1957 p. 74, Sedgwick 1988, Hoy 1989 p. 214
[18] Original emphasis, compare Hegel 1986 PG, Werke 3 p. 319.
[19] Korsgaard 2008 p. 321
[20] Herman 1993 p. 230
[21] Stern 2013 p. 23
[22] Sensen 2013 pp. 64-5
[23] Building on a comparison between Kant scholarship and early 20th century analytic philosophy, Jeremy Schwartz has aptly labelled these Neo-Kantian approaches 'practical logicism' (Schwartz 2017 pp. 259-262).

while Kant rejects a formal criterion of truth in the theoretical realm, he allegedly embraces a parallel, formal criterion for right action in the practical realm. Hegel's comparison between Kant's Categorical Imperative and a purely formal criterion of truth strikingly mirrors these Neo-Kantian characterizations of the Categorical Imperative.

Therefore, a tempting way of reading Hegel's empty-formalism charge is as a charge of inconsistency. Hegel's complaint might be that Kant's moral philosophy stands in tension with his attempt to articulate a system of critical philosophy that treats practical and theoretical reason as two applications of one and the same capacity. Unfortunately, there is a significant problem for this reading of Hegel's empty-formalism charge: Kant explicitly rejected the idea that his ethics would rest on such a formal approach. Already in the introduction to the *Groundwork*, can explains to his readers how this preliminary work on moral philosophy fits within his larger critical system:

> All rational cognition is either material and concerned with some object, or formal and occupied only with the form of the understanding and of reason itself and with the universal rules of thinking in general, without distinction of objects. Formal philosophy is called logic, whereas material philosophy, which has to do with determinate objects and the laws to which they are subject, is in turn divided into two. For these laws are either laws of nature or laws of freedom. […] All philosophy insofar as it is based on grounds of experience can be called empirical; but insofar as it sets forth its teachings simply from a priori principles it can be called pure philosophy. When the latter is merely formal it is called logic; but if it is limited to determinate objects of the understanding it is called metaphysics. In this way there arises the idea of a twofold metaphysics, a metaphysics of nature and a metaphysics of morals. Physics will therefore have its empirical part but it will also have a rational part; so too will ethics, though here the empirical part might be given the special name practical anthropology, while the rational part might properly be called morals. (GMS 4:387-8)

In this passage, Kant explicitly tells us that the investigation of the *Groundwork* is concerned with a kind of cognition that "has to do with" or "is limited to" "determinate objects." This is what makes the book a groundwork for a *metaphysics* of morals. Thus, when Kant analyses the capacity of

finite, rational beings to represent actions and ends as intrinsically good, and then tries to articulate the form of this capacity in different formulas, he is already thinking about that capacity in relation to the determinate representations of moral actions and ends in our ethical lives.

Moreover, Kant repeatedly emphasizes that the content of the representations of practical reason is not determined by the form of practical reason alone. In several passages in the *Metaphysics of Morals*, Kant mentions the importance of knowledge of human nature for morality.

> […] just as there must be principles in a metaphysics of nature for applying those highest universal principles of a nature in general to objects of experience, a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles. But this will in no way detract from the purity of these principles or cast doubt on their a priori source. - This is to say, in effect, that a metaphysics of morals cannot be based upon anthropology but can still be applied to it. (MS 6:216-7)[24]

In light of passages like these in both the *Groundwork* and the *Metaphysics of Morals*, Kant appears to be very consistent. The fundamental principle of morality is not an instance of general logic, which is merely the "formal business of the understanding and reason," and which remains "entirely indifferent with regard to the objects."[25] Instead, Kant's moral philosophy is an exercise in metaphysics. By this, Kant means that his moral philosophy is "material philosophy," which is "limited to determinate objects of the understanding." A metaphysics "has to do with determinate objects and the laws to which they are subject" (GMS 4:387). Hence, charging Kant with inconsistency might not be so tempting after all.[26]

---

[24] See also MS 6:468-9.

[25] Note that the general logic Kant discusses in the *Critique of Pure Reason* is not limited to theoretical reason. Rather, it is supposed to apply to reason in general, whether theoretical or practical.

[26] By rejecting the idea that there is a stark asymmetry between Kant's practical and theoretical philosophy, in the way in which both some Neo-Kantian interpretations and some of Hegel's passages might suggest, I do not mean to endorse the opposite extreme, i.e., the view that there would be a perfect symmetry between Kant's practical and theoretical

However, we might instead conclude that Hegel is the one who is mistaken. Perhaps Hegel really *did* charge Kant with being inconsistent, but simply overlooked (or misunderstood) the passages in which Kant denies taking a formalistic approach to ethics. In her early work, Sally Sedgwick has argued that Hegel makes such a mistake when he portrays Kant's ethics as an instance of general logic. Consequently, Sedgwick has suggested that we can answer Hegel's empty-formalism charge by getting clear about the difference between general and 'transcendental logic.'[27] The difference between general and transcendental logic can roughly be summarized as follows. Kant holds that general logic, or "pure general logic," cannot tell us anything substantive about "the material (objective) truth of knowledge" (A60/B85). Rather, general logic provides us with the "absolutely necessary rules of thought" (A52/B76) and "deals solely with the form of thought in general" (A131/B170). Consequently, the validity of the laws of general logic can be assessed independently of any determinate object of cognition. By contrast, transcendental logic provides the a priori norms that enables us to have representations of determinate objects. Without these a priori norms, our representations would not refer to any determinate content at all. Therefore, the validity of the norms of transcendental logic can only be determined in relation to content. A metaphysics supposedly deals with exactly those norms that make the cognition of certain objects possible. If the norms are 'laws of nature,' Kant talks about a metaphysics of nature; if the norms are 'laws of freedom,' Kant talks about a metaphysics of morals. In this way, Kant's metaphysics is separate from both formal philosophy (i.e., general logic) and empirical philosophy (i.e., empirical physics and practical anthropology).

---

philosophy either. Here, I merely want to point out the difficulties for the views that attribute a strict asymmetry to Kant's practical and theoretical philosophy, while leaving the precise nature of their differences and parallels as open as possible.

[27] Sedgwick 1988 pp. 68ff

Can Hegel's empty-formalism charge be answered simply by pointing out that Kant's ethics was an exercise in metaphysics, and thus concerned with transcendental, not general, logic? In her early work, Sally Sedgwick argued that it can:

> Kant's specific concern in his transcendental philosophy with its transcendental logic is to provide a deduction of the laws without which thought would not relate to content. […] Turning to the moral philosophy, we contradict the ground of the possibility of morality for rational nature as such when our maxims fail the test of the categorical imperative considered within the context of the *Critique of Practical Reason* or the *Foundations*. […] the Hegelian critique is misdirected in so far as it underplays the role that empirical content plays in Kant's moral theory and in so far as it characterizes the command of the categorical imperative as requiring no more of us than consistency in our choice of maxims […][28]

Sedgwick here provides another tempting response to Hegel empty-formalism charge. But, like many of the readings surveyed above, she also faces two significant difficulties. First, Sedgwick's response slightly misses Hegel's criticism. Hegel is not merely saying that the Categorical Imperative must be applied to empirical content, i.e., to concrete situations with knowledge about human nature in order to work as a decision-procedure. Rather, Hegel argues that the formality of the Categorical Imperative makes it unclear how it *could be* applied to content such that it would determine what we should do. In other words, Hegel is not merely saying that we need some knowledge about human nature in order to feed maxims into the test; Hegel is saying that the Categorical Imperative is not providing a useful criterion for the determination of right action no matter what empirical knowledge of human nature we have.

---

[28] Sedgwick 1988 p. 79

Second, and more importantly, Sedgwick's response risks portraying Hegel as a lousy reader of Kant once again. As Sedgwick herself points out, Hegel is of course aware of the distinction between general and transcendental logic in Kant's work.

> What is puzzling is that Hegel elsewhere seems to acknowledge the distinction between general and transcendental logic in Kant, and the special role of synthetic a priori laws of reason. (See, e.g., *Enzyklopddie der philosophischen Wissenschaften I*, §40.) Nevertheless, his reading of Kant's moral philosophy does not reflect an awareness of these distinctions.[29]

This puzzle alone should make us cautious to adopt Sedgwick's interpretation of Hegel's empty-formalism charge. Hegel famously disdained the distinction between general and transcendental logic as Kant drew it, and he argued that Kant was overly formalistic not merely in his practical philosophy, but also in his theoretical philosophy. Finally, it is worth noting that general logic just deals with "the form of thought *in general*" (A131/B170). That is, Kant intended the general logic to apply to both theoretical and practical reason alike – even though his examples in the first *Critique* are taken from theoretical reason, as was (and still is) common practice in discussions of philosophical logic. In this light, it seems unlikely that Hegel criticized Kant's approach to moral philosophy for being overly formalistic simply because he was mistaken about the right parallel between Kant's theoretical and practical philosophy, or about the place Kant's metaphysics of morals occupied in Kant's philosophical system. Therefore, we should reconsider Hegel's texts and see if a compelling alternative interpretation of the empty-formalism charge is available.[30]

---

[29] Sedgwick 1988 p. 69

[30] In her later work, Sedgwick (1996) herself eventually rejected this response. Eight years after her initial response to Hegel discussed above, Sedgwick argued that Hegel's empty-formalism charge is part of his larger criticism of both Kant's theoretical as well as practical philosophy as overly formalistic. See Sedgwick 1996 and 2011.

Against Separating Form and Content

So far, I have argued that we should be skeptical of interpretations on which Hegel criticized Kant because he either overlooked the FH, or misunderstood Kant's notion of contradiction, or because he disregarded the distinction between general and transcendental logic. Instead, it is worth considering if Hegel compared Kant's ethics to general logic because he believed that *all* Kantian metaphysics is overly formalistic.[31]

Kant had argued that the unification of representations in our experience of the world is made possible by certain a priori rules contributed by our rational faculties. And he thereby also argued that our knowledge of the world is fundamentally dependent on these a priori rules contributed by our rational faculties. How things are in themselves, independent of our possible experience of them, could therefore not be known. By contrast, Hegel attempted to draw a very different conclusion from Kant's insight into the necessary contributions of our rational faculties to experience: namely, that the philosophical distinction between 'form' (the a priori rules contributed by our own rational faculties) and 'content' (the independently given content of experience) should somehow be overcome. Kant, however, did not attempt to overcome this ostensible limitation of human knowledge, focusing instead on deepening our understanding of the formal contributions

---

[31] In her recent work on Hegel's criticism of Kant, also Sally Sedgwick has defended this interpretive option, and it provides a particularly helpful starting point for understanding the empty-formalism charge. Although Sedgwick's recent work provides an excellent exposition of the empty-formalism charge as part of Hegel's overall criticism of Kant, I believe that her interpretation still falls short in a crucial respect: Sedgwick argues that Kant can be saved from the charge by acknowledging that the "categorical imperative test" was not meant to be formal, and that Kant's discussion of the FH in the *Groundwork* provides the content needed to make the Categorical Imperative less formal (Sedgwick 2012 pp. 273-77). However, this response to the empty-formalism charge faces the same difficulty as the interpretations discussed in section 1 above. Moreover, this response is difficult to square with Kant's repeated insistence in the *Groundwork* that the Categorical Imperative is indeed formal (GMS 4:400, 403-4, 416, 421, 426, 428, 436-7, 444, 458, 461-2).

made by our rational faculties to experience.[32] In this way, according to Hegel, Kant's theoretical

philosophy remained overly 'formalistic.' In what way, then, might Kant's *practical* philosophy be

overly formalistic?

Kant's conception of metaphysics in general is an attempt to discover the 'form' of reason.

That is, Kantian metaphysics attempts to investigate the conceptual norms without which our

cognition would not be about any determinate content at all. These conceptual norms are what Kant

called the 'formal' element of cognition. In the theoretical philosophy, space, time and the categories

are 'form' whereas the manifold is matter which are accessible to us because of that form. (As Kant

tells us early in the first *Critique*, "I call that in the appearance which corresponds to sensation its

matter, but that which allows the manifold of appearances to be intuited as ordered in certain

relations I call *the form* of appearance" (KrV A20/B34, emphasis added).) Similarly, practical agents

represent actions as to be done through subjective principles of action (maxims). And these

principles for representing actions as to be done all have a specific form or structure. Significantly,

according to Kant, the particular principles making it possible to represent actions and ends as

intrinsically good have a very specific form, and we can attempt to analyze and understand this form

in more detail: the form of these principles making the representation of actions and ends as

intrinsically good possible is the Categorical Imperative. Or, in other words, the Categorical

Imperative is the form of moral judgment.[33]

---

[32] See especially part one of Hegel's *Encyclopedia* §§40-46, and §60. For a detailed account of Hegel's criticism of Kant's Idealism see the first two chapters of Robert Pippin's *Hegel's Idealism* 1989 and chapter three of Sally Sedgwick's *Hegel's Critique of Kant* 2012. Compare also Sedgwick's 1996 pp. 569-578. Whether or not this interpretation and criticism of Kant's theoretical philosophy is ultimately convincing is certainly up for debate – however, even if we are unconvinced by Hegel's criticism of Kant's theoretical philosophy, we would still thereby attribute a significantly smaller misreading of Kant to Hegel than the currently standard interpretations of Hegel's empty-formalism charge.

[33] See GMS 4:400, 403-4, 416, 421, 426, 428, 436-7, 444, 458, 461-2. Note that in Kant's terminology, insofar as reason abstracts from the influences of sensibility, reason is 'pure:' "I call all representations pure (in the transcendental sense) in which nothing is to be encountered that belongs to sensation" (KrV A 20/B 34). We encounter the same sense of

The Kantian project of analyzing the forms that enable us to have concrete representations also comes with a methodologically necessary limitation: on this picture, form and matter cannot be separated in anything more than abstract speech. Regarding theoretical cognition for instance, it is impossible to look at the manifold over here and then turn to our conceptual capacities (i.e., the forms of reason) over there. You just cannot take off your spectacles and look through them as well. Therefore, the validity of the norms of transcendental logic can only be determined in relation to specific content. Conversely, it is impossible to represent a determinate object without the formal principles of our rational capacities. Regarding practical cognition, a similar inseparability obtains between the form and content of practical judgment. It is impossible to represent an action as to be done without a formal principle, i.e., through a maxim. That is, you just cannot represent an action *as to be done* without the formal principle, i.e., the maxim that makes such representations so much as possible. Conversely, you cannot find the specific form of such representations except by analyzing what it means to have such representations, because the form does not exist independently by itself but only as the form of a particular kind of representation.

Hegel calls the Kantian attempt at abstracting the form from the matter of practical cognition the "non-identity of ideal and real." In his Natural Law Essay, Hegel characterizes this attempted abstracting of form and matter in the practical philosophy of his predecessors as follows:

> [D]espite the fact that the ideal and the real are identical in this practical reason, the real remains purely and simply opposed [to the ideal]. This real is essentially posited [as] outside reason, and practical reason is [to be found] only in its difference [Differenz] from it. […] This science of the ethical, which talks of the absolute identity of the ideal and the real, accordingly does not do what it

---

purity of form in Kant's practical philosophy. In the *Groundwork*, Kant says that "[p]ractical principles are formal if they abstract from all subjective ends, whereas they are material if they have put these, and consequently certain incentives, at their basis" (GMS 4:427-8).

90

claims to do: its ethical reason is, in truth and in its essence, a non-identity of the ideal and the real. (NLE 120)

In this passage, Hegel makes a couple of striking – although impossibly worded – claims. First, Hegel contends that genuinely practical reason is reason not merely thinking but acting in the world (the 'identity of real and ideal'). Second, he claims that on the Kantian picture, reason appears as those formal principles which we contribute to our practical cognition. Consequently, Hegel maintains that on this Kantian picture, practical reason appears as defined by its opposition to the actual content, or matter, of acting in the world. Finally, Hegel plausibly assumes that a 'science of the ethical' is about practical reason, i.e., about reason acting in the world. Thus, Hegel concludes that if Kantian ethics were to be a 'science of the ethical,' it would fail to properly address its own topic; for Kantian ethics merely investigates practical reason in its 'purity' or 'formality,' abstracted from its concrete content; and thus, Kantian ethics stops short of considering practical reason as what it truly is, namely embodied in action in the world.

These claims will become clearer after considering Hegel's later formulations of the charge, in a moment. But in order to get the criticism in view, it is helpful to put the overall point in more familiar language: Hegel's empty-formalism charge tries to show that no critical theory of right action –no 'immanent doctrine of duties' – can be developed by separating the form of practical reason from its determinate content. Kant's account of practical reason is thoroughly hylomorphist; and from conceptually analyzing the form alone a hylomorphic account cannot derive the resources to provide a critical theory of right action. Put programmatically: if no determinate content can be represented without representing it according to some form, then form alone cannot add a criterion that is not already contained in the representation of the content. Conversely, if form is conceptually separated from content, then it remains unclear what could make *this* rather than *that* content

91

incompatible with one form or another. In other words, any concrete action could be represented as good depending on which maxim it figures in, because neither the concrete content (the action) nor the form of representing it by themselves determine which content can figure in which form.

Thus, in the *Philosophy of Right*, Hegel claimed that no 'immanent doctrine of duties' could be developed on the basis (or from the point of view) of the Categorical Imperative:

> However essential it may be to emphasize the pure and unconditional self-determination of the will as the root of duty […] to cling on to a merely moral point of view without making the transition to the concept of ethics reduces this gain to an empty formalism, and moral science to an empty rhetoric of duty for duty's sake. From this point of view, no immanent theory of duties is possible. (R §135)

Hegel continues to explain this claim in the following way: "One may indeed bring in material from outside and thereby arrive at particular duties…" In other words, one may think about examples of actions one already accepts as good, and then see whether the representations of such actions we already presume to be good pass the 'test' of the Categorical Imperative, and thereby confirm them as moral duties.

> …but it is impossible to make the transition to the determination of particular duties from the above determination of duty as absence of contradiction, as formal correspondence with itself, which is no different from the specification of abstract indeterminacy (R §135)

That is, applying the Categorical Imperative as a test to actions we already presuppose to be good would not help us make a transition from the form of practical reason to a critical theory of right action. For the application of this 'test' to an action we presuppose as good does not add anything that is not already contained in the representation of this action as good. After all, the form that serves as the 'test' is already contained in the very representation of an action as morally good. Consequently, such a 'test' would be entirely superfluous. Hegel adds:

92

…and even if such a particular content for action is taken into consideration, there is no criterion within that principle for deciding whether or not this content is a duty.

In less enigmatic wording: whatever content we bring in 'from outside' to the test of the Categorical Imperative, the form with which we try to determine our duties would already be contained in the representation of the content that we bring in. Once we 'test' any representation of an action as intrinsically good, our representation of that action already implies that it is represented in a particular way, through a particular form, through a maxim with a particular form. Therefore, we have 'no criterion' for deciding whether or not the proposed content *really* is an appropriate duty or not; the only criterion we have – the form of representing actions as intrinsically good – necessarily comes too late. To summarize: Hegel suggests that any formal criterion of the Categorical Imperative can only be applied to a particular content once this content is represented as a potential moral act. But once an action is so represented, the representation of this action would be tested against something that it itself already contains (since the Categorical Imperative just is the form through which we represent actions as intrinsically good). As Hegel puts it in the *Phenomenology*, "The standard of the law which reason has in its own self therefore fits every case equally well and is thus in fact no criterion at all." (PG §430)[34] Of course, we could also take the converse example of an action we presuppose as bad and represent thus; again, the test would come too late, merely confirming that the form through which we represented the action in question was not the representation of an action as morally good.

This is a genuine problem for any criterion or standard that one attempts to apply as a test. For in order to function as a test, the standard must be separable from the thing it is supposed to

---

[34] We are now also in a better position to see why it is misguided to assume that Hegel just misunderstands how the contradiction is supposed to work in Kant's universalization test. The 'absence of contradiction' in Hegel's empty-formalism charge in the *Philosophy of Right* refers to the correspondence between the form and the content of practical reason.

evaluate. The standard must enable a potential discovery: it must (at least in theory) be able to tell us something that we did not by definition know before. Like a scale must be separable from what it measures. And this kind of separability is just not available on Kant's hylomorphism. Form cannot be a test for mere matter; form can be incompatible with a different form, but not merely with matter. Therefore, Kant also emphasizes that, in the case of immoral action, the Categorical Imperative is not contradicting a specific action, specific content, but *maxims*. If the form of pure practical reason is incompatible with something, it is incompatible with another form, namely the form of a maxim representing an action as merely subjectively good. Articulating the form of the good will can make explicit its difference to a bad will; but articulating the form of the good will does not, by itself, tell us *what* can be willed by a good will unless the concrete matter of human action also had a form that can correspond or contradict the form of the will. However, that final option is not part of Kant's explicit moral philosophy; for Kant, form is the contribution of reason, which explicitly appears in opposition to the actual content, or matter, of acting in the world. In Kant's practical philosophy, the form of practical reason is the form of willing, and the form of practical reason is imposed by the spontaneity, or autonomy, of reason. It is not something 'found' in the world, in the form of particular actions that in themselves already contain a certain form.

But things may be even worse. According to Hegel, not only is Kant's hylomorphism and his specific focus on the form of practical reason unable to provide us with the resources for developing a critical theory of right action; it may even appear to justify bad action. Hegel's explanation in the *Philosophy of Right* continues thus:

> On the contrary, it is possible to justify any wrong or immoral mode of action by this means. – Kant's further form – the capacity of an action to be envisaged as a universal maxim – does yield a more concrete representation of the situation in question, but it does not in itself contain any principle apart from formal identity and that absence of contradiction already referred to.

94

The Categorical Imperative, as the form of pure practical reason, is the particular way in which we represent acts as intrinsically good. Consequently, no content that we so represent can be ruled out as an immoral act by the very form through which it is represented. This is what Hegel means by describing the formal criterion of the Categorical Imperative as both 'absence of contradiction' and 'formal correspondence with itself' or 'formal identity:' form and content correspond, i.e., do not contradict each other because they are inseparable in our representation. As noted above, whatever content we bring to the 'test' we already necessarily represent through a particular form, making the 'test' superfluous.[35] Hegel then attempts to drive home this point with two famous examples: private property and human life.

> The fact that *no property* is present is in itself no more contradictory than is the non-existence of this or that individual people, family, etc., or the complete *absence of human life*. But if it is already established and presupposed that property and human life should exist and be respected, then it is a contradiction to commit theft or murder; a contradiction must be a contradiction with something, that is, with a content which is already fundamentally present as an established principle. (R §135, original emphasis)

On Kant's conception of form, content like specific actions does not itself have a form and therefore cannot contradict the form of pure practical reason either – only representations have form, and consequently maxims can contradict the form of pure practical reason. For instance, the bare content of taking things according to one's needs and pleasure is not itself contradictory with anything yet: a "contradiction must be a contradiction with something." But this content (taking things according to one's needs and pleasure) can only contradict the form of practical reason if this content is represented through a specific form ("is already fundamentally present as an established

---

[35] To draw an analogy with theoretical cognition: one might similarly think that only conceptual content can rationally limit the concepts one brings to bear on this content.

principle"), like the form it received when described through a maxim about taking other people's property.

Contemporary readers of Kant might well find Hegel's charge preposterous on first sight – 'Of course,' one might think, 'some matter is incompatible with a particular form! Didn't we show how some maxims contain e.g., practical contradictions when universalized?' And we might be tempted to agree – does it not seem like some maxims cannot properly be universalized? But such an imagined response would miss the point of Hegel's criticism. The problem is not that universalization cannot rule out some maxims – it of course *can* rule out certain maxims. (Hegel, in fact, explicitly agrees.[36]) Rather, the problem is that no type of action by itself determines what kind of maxim it should figure in, and thus it seems like any kind of action can potentially be represented in a universalizable maxim. Hegel specifically complains that any immoral *action* can be made to pass the test by putting it in the right kind of maxim, while some maxims that contain actions with supposedly moral content (like 'help the poor') are supposedly difficult to universalize. Hegel elaborates this point in the *Philosophy of Right* thus:

> [I]t is possible to justify any wrong or immoral *mode of action* by this means. – Kant's further form – the capacity *of an action* to be envisaged as a universal maxim – does yield a more concrete representation of the situation in question, but it does not in itself contain any principle apart from formal identity and that absence of contradiction already referred to. (R §135, emphasis altered)

For Hegel, the crux of the issue is that *any* formulation of a maxim represents a particular action according to a certain form. But whether we put some action into a maxim that represents this act as moral is, at least on the Kantian picture, neither determined by the specific action nor by the form through which we represent actions as moral. Put programmatically: we can distinguish a good will

---

[36] NLE pp. 127-128

from a bad will through their form, but that does not, without more, determine the content of a good will. Of course, people don't always have good intentions; but you can have 'good intentions' doing horrific things. And from the fact that we know what makes some practical judgments a good will and others a bad will it does not follow that we thereby already know which actions and ends should figure in the practical judgment of a good will.

On the face of it, this limitation seems closely related to the alleged problem of maxim description in the literature on Kant's ethics: what is the right way to formulate one's maxim? And what considerations can legitimately be taken into consideration when formulating one's maxim, without either incorporating normative content into the description or leaving out anything significant form the description?[37] However, Hegel's empty-formalism charge is not simply the same as the alleged problem of maxim-description in Kantian ethics; rather, the problem of maxim-description is a result – or better: a symptom – of the formalism of Kant's ethics. That is, Hegel's empty-formalism charge identifies the systematic limitation of Kant's standpoint, of which the contemporary, academic topic of 'maxim description' is merely the surfacing consequence. As we saw, neither the specific actions nor the forms of our maxims by themselves determine which action should be represented through which maxim. Therefore, something else must provide a way of determining through which form a specific action should be represented. Accounts of the right way to describe one's maxim promise to do exactly that: they promise to tell us which actions can figure in a good will, and which ones cannot. Could Hegel's empty-formalism charge then be answered by providing an account of maxim-description? Unfortunately, not quite. Such an account of maxim-description could not answer the empty-formalism charge without also going beyond the Kantian

---

[37] Unsurprisingly, many recent responses and endorsements of Hegel's empty-formalism charge discuss the problem of maxim description at length. See for instance Hoy 1989 pp. 215-6, Westphal 1991 pp. 158-59, and Freyenhagen 2011 pp. 175-180.

moral standpoint. Of course, one may either give up Kant's hylomorphism and instead take form to be contained in the social and historical structures of our lifeworld. But by doing so, one would also give up one (if not *the*) central element of Kant's ethical thought. Consequently, trying to address the empty-formalism charge by focusing on the difficulty of maxim description would be to overlook the central point of Hegel's criticism: the problem is that once we formulate a maxim, we have *thereby already* given it a particular form. And once we have given it a particular form, the further 'test' whether it conforms to this or that form is moot. A solution to the problem of maxim-description that also responds to the empty-formalism charge would have to be one that overcomes the form-matter division, and shows how there is some kind of unity, or 'fit', between the possible actions – and the institutions in which they are embedded – and the forms of practical reason that represents such actions as to be done. Thus, such an account would not so much vindicate the viability of a 'CI-procedure' as it would make such a test superfluous.

The Significance of Hegel's Empty-Formalism Charge

Hegel's empty-formalism charge is neither a misreading nor misrepresentation of Kant's Categorical Imperative. Quite to the contrary, Hegel remains an unusually insightful reader of Kant. As I have argued above, Hegel's empty-formalism charge is just one facet of his larger criticism of Kant's critical philosophy as overly 'formalistic:' in both the theoretical and the practical philosophy, Kant's alleged formalism lies in his attempt to build a philosophical system around the idea of conceptually separating the form of cognition from its content: to see the formal as contributed by reason, and to locate the material outside of it. Therefore, Hegel thinks that any formula of the Categorical Imperative would be insufficient for the development of a critical theory of right action

– that is, a critical engagement with our common moral knowledge, determining which actions are truly right and truly wrong. If no determinate content can be represented without form, then the form of the good will alone cannot provide a criterion that is not already contained in the respective representation; any 'CI-procedure' would come too late.

In this way, Hegel's empty-formalism charge highlights both an insight and a limitation of Kant's moral philosophy. It is one thing – and a big thing, for that matter – to know that all moral judgments represent actions in accordance with some particular form. This is Kant's insight into practical reason, for which he receives ample praise from Hegel. Importantly too, this insight must remain part of any post-Kantian story about what is right and wrong – as it is in Hegel's. But it is quite another thing to know *which* actions should be represented in accordance with exactly that form, and to develop a systematic theory of right action. According to Hegel's empty-formalism charge, Kant was not wrong about the 'standpoint of morality' (i.e., the form of pure practical reason); but from the perspective of Kant's moral philosophy, it also seems that we lack the resources to provide a critical and substantive theory of right action.[38]

The significance of Hegel's empty-formalism charge, however, is not simply that Kant's moral philosophy contains both insights and limitations. The significance of Hegel's charge is that the original insight and limitation of Kant's moral philosophy are one and the same. And this is an infinitely deeper point that e.g., G. A. Tittel or J. S. Mill's criticisms of Kant's Categorical Imperative

---

[38] I discuss Kant's view on the purpose of moral philosophy in more detail in chapters 1-2. Despite their seeming agreement on the purpose of moral philosophy, there is one significant difference between Hegel and contemporary neo-Kantian ethicists in this context: most Neo-Kantian approaches take the Categorical Imperative to be a test or decision-procedure and try to argue that such a test can bridge the gap between the form of practical reason and a substantive system of duties. By contrast, Hegel's empty-formalism charge does not, strictly speaking, claim that Kant intended the Categorical Imperative to function as a test or decision-procedure; rather, Hegel believes that there is a gap between the form of practical reason and a substantive theory of right action, and that the Categorical Imperative cannot help us bridge that gap by functioning as a test. Thus, Hegel's criticism applies whether or not the Categorical Imperative was ever intended to be used as a test or decision-procedure.

as practically useless. On Hegel's view, it is simply not possible to do moral philosophy adequately without taking heed of Kant's original insight, and without struggling with the limits of Kant's ethics. At its most extreme, the failure to take either one of them seriously would be to either change the topic away from ethics (by failing to appreciate the moral standpoint), or to change the subject away from philosophy (by failing to appreciate the limits of what can be sensibly said on the topic). Thus, Hegel's empty-formalism charge leaves us with two questions. How much, exactly, can be said about ethical life from a Kantian perspective? And how important, exactly, is it to transcend the Kantian perspective?

A related upshot of Hegel's empty-formalism charge is that it highlights the importance of Kant's and Hegel's respective philosophical systems. Both Kant and Hegel agree that there is a formal standpoint which Kant's analysis of the Categorical Imperative captures; and both agree that the story continues beyond this analysis of the Categorical Imperative. Kant continued to write a metaphysics of morals in two volumes, while Hegel tried to go beyond morality in his treatment of ethical life (Sittlichkeit). Hegel's criticism was that Kant's analysis of the form of pure practical reason does not give us the resources to proceed – not, perhaps, to Kant's own metaphysics of morals, but to the kind of substantive moral philosophy that Hegel thought we need. Conversely, if neither Kant nor Hegel had been interested in going anywhere beyond an analysis of the form of the good will, Hegel's attitude toward Kant's formalism would make significantly less sense. (Therefore, contemporary readers who treat Kant's ethics as if it was essentially finished by the end of section II of the *Groundwork* won't avoid the difficulty in explaining how Hegel can simultaneously praise and criticize Kant's discussion of morality.[39])

---

[39] For a prominent example, see Korsgaard's characterization of the *Groundwork* (Korsgaard 1998) and her rejection of Hegel's empty-formalism charge in her discussion of the universal law formula (Korsgaard 1996 pp. 86-87).

It would therefore be ill-advised to attempt to defend Kant against the empty-formalism charge by e.g., focusing on the distinction between general- and transcendental logic, or the specific notion of contradiction involved in any alleged universalization test, or the right ways of describing maxims. Answering the empty-formalism charge first requires determining to what extent, if any, Kant needs to be defended against the empty-formalism charge at all: does Kant really attempt to provide a test or decision-procedure with his formulations of the Categorical Imperative? And does his *Metaphysics of Morals* attempt to provide a theory of right action? And if Kant attempted neither, what reasons might we have for being more ambitious than Kant, such that his formalism were to become a problem?

PART II

FORM AND THERAPY IN KANT'S METAPHYSICS OF MORALS

Chapter 4

Kant's Derivation of Imperatives of Duty

Introduction

Contemporary readers of Kant are almost unanimously united in the assumption that Kant's famous derivation of 'imperatives of duty' in Section II of the *Groundwork* introduces the Formula of Universal Law (FUL) and the Formula of Humanity (FH) as 'tests' or 'decision procedures' for the derivation of concrete duties.[1] In this way, Kant's *Groundwork* holds the promise of being both objective and practicable: because the Categorical Imperative abstracts from all contingent, empirical facts about human nature, it is applicable everywhere and at all times; and because the Categorical Imperative allegedly provides a decision procedure for the derivation of concrete duties, we expect that it will be useful in guiding our moral deliberation. Precisely, this promise of an objective and practicable foundation for normative ethics has made Kant's ethics attractive to many readers.

However, Kant's own engagement with his contemporary critics complicates this standard assumption about the *Groundwork's* famous derivation of imperatives of duty. An early reviewer had criticized the *Groundwork* for providing a principle that is 'unfit for practical use' in determining which actions are right and thus concluded that the *Groundwork* did not provide any new principle for the 'reform' of ethics. In his response, Kant claimed that his reviewer had hit the mark better than he might have realized because the *Groundwork* simply never intended to provide such a 'new

---

[1] At least since Sidgwick's *Methods of Ethics*, the terms 'test' and 'decision procedure' have been used roughly interchangeably to refer to some (ideally) useful and reliable method for moral deliberation by which agents can figure out what they ought to do. In the current literature, the terms are often used to talk either of decision procedures in general or 'universalizability tests' and 'ends-in-themselves tests' specifically. However, current usage is loose enough to preserve the terms' interchangeability, and so I too will use them interchangeably.

principle' for the reform of moral philosophy, but merely to capture the formula of our common

moral knowledge. When we read Kant's reply together with the original criticism, it becomes clear

that Kant himself was remarkably unconcerned with arguing for the practicality of the Categorical

Imperative as a decision procedure for the derivation of concrete duties.[2]

Kant's response should make us open to reassessing the standard assumption about the

Groundwork's examples of derivations of imperatives of duty and to reconsider what their purpose

might be. I argue that Kant attempted to show, in his derivation of imperatives of duty, how his

analysis of moral judgement and its fundamental principle underlies the four types of moral

imperatives that philosophers of his time recognized and that outlining a decision procedure for the

derivation of concrete duties was not his main concern in section II of the *Groundwork*. In other

words, Kant's infamous 'derivation of imperatives of duty' is best understood literally as a derivation

of the traditionally accepted types of imperatives, rather than a derivation of duties.

Unfortunately, the most important translation of Kant's work into English still today

misidentifies the author of the commentary mentioned above.[3] Moreover, the actual commentary,

---

[2] By contrast to 'test' and 'decision procedure', moral 'criterion' is usually used to refer to a principle specifying the ultimate right-making feature of actions from a third-person perspective; both functions have been ascribed to either or both the FUL and FH in Kant's 'derivation of imperatives of duty'. (For an account that ascribes merely the former function to the FUL and merely the latter function to the FH, see Timmons 2017, chapter 2.) In this article, my aim is to challenge the (still) widespread assumption that either formula in Kant's 'derivation of imperatives of duty' constitutes a test or decision procedure for guiding moral deliberation and choice. Therefore, I leave aside the question which formula(s) might serve as a moral criterion.
[3] The Mary Gregor translation of the second *Critique* in the Cambridge Edition of the Works of Immanuel Kant (in Kant 1996), as well as in the Cambridge Texts in the History of Philosophy series (in Kant 2015), wrongly identify this reviewer as Johann Friedrich Flatt (and mis-cite Flatt's actual review). Flatt had indeed written an equally critical review of Kant's *Groundwork* in the same year as Tittel, and there is some overlap between their respective criticisms to the extent that Tittel acknowledges Flatt's (1786) review in his preface. However, the criticism that Kant responds to at CPrR 5: 8 does not appear in Flatt's review and is made by Tittel. It is also worth noting that Paul Natorp, in his annotations to volume 5 of the Academy Edition of Kant's writings at 5:506-7), Karl Vorländer in his introduction to the Philosophische Bibliothek edition of the second Critique (Kant 1959 pp. xvi–xvii), and Werner Pluhar in the preface to his translation of the second *Critique* (Kant 2002a p. 13) all correctly identify Tittel as the reviewer in question. The mistaken attribution of authorship in the CUP editions has also been noted by Pluhar and Walschots (2020). For a plausible explanation of the misattribution see Pluhar (Kant 2002a pp. 12-13).

authored by Gottlob August Tittel, has until recently not been translated into English.[4] At first sight, both the misidentification and the lack of an English translation might appear to be of historical interest only. But if we take a close look at Tittel's actual criticism and keep in mind that Kant arguably meant to demonstrate how the Categorical Imperative underlies the four standardly recognized types of moral imperatives, we can also better appreciate Kant's response.

Revising our assumptions about the purpose of Kant's derivation of imperatives of duty in the Groundwork also implies a break with a long tradition of Kant scholarship. The most influential examples of this tradition are no doubt Onora O'Neill's *Acting on Principle*,[5] Christine Korsgaard's well-known 'Kant's Formula of Universal Law',[6] and John Rawls' *Lectures on the History of Moral Philosophy*, in which he characterizes the Categorical Imperative as providing a 'procedure by which that imperative is applied to us as human beings'.[7] Today, the standard assumption of this tradition still permeates introductory texts and popular philosophy.[8] So far, only a handful of scholars have deviated from this standard assumption about the *Groundwork*, and most deviations have been far less radical than the textual evidence would allow. Moreover, the relatively few deviations so far rely on philosophical resources outside the *Groundwork* for a full motivation. By contrast, I argue that the *Groundwork* taken together with Kant's rejoinder allows for the deviation.[9] To my knowledge, until

---

[4] An English translation of selections of Tittel's book by Michael Walschots is forthcoming in Kant's *Critique of Practical Reason: Background Source Materials* by Cambridge University Press.
[5] O'Neill 2013, first published 1975 under the name Onora Nell.
[6] First published 1985, reprinted in Korsgaard 1996 pp. 77-105.
[7] Rawls 2000 p. 162. For a brief discussion of the influence of Rawls in the study of Kant's ethics see Schneewind 2010 pp. 239-47 and Ameriks 1996 p. 38. For additional examples of this interpretation see (among others) Kemp (1958), Gregor (1963), Dietrichson (1964 and 1969), O'Neill (2013[1975], 1991 and 1998), Höffe (1977), Schnoor (1989), Sullivan (1989), Reath (1989 and 1994), Allison (1990 and 2011), Schneewind (1991), Joerden (1993), Korsgaard (1996), Wood (1997 and 1999), Kerstein (2002), Guyer (2005), Hill (2007), Timmermann (2007), Galvin (2009), Velleman (2012), Timmons (2017) and Kleingeld (2017 and 2019).
[8] See for instance the BBC Radio 4's (2017) 'Kant's Categorical Imperative'.
[9] So far, almost all deviations remain far less radical than, I believe, they should be. Recently, Manfred Kuehn has expressed some doubt about the Categorical imperative's use as a test, although based on his reading of the importance of the *Metaphysics of Morals* (2010 pp. 23-27) rather than on a reading of the *Groundwork*. Allen Wood (2007) has suggested that the categorical imperative is not providing a strict 'algorithm' for testing maxims, but continues to take the FUL and

now only Ido Geiger has argued in print that Kant's four examples of duties in the *Groundwork's*

discussion of the FUL are not meant to provide examples of a universalization test.[10] But since this

pathbreaking departure from the standard assumption also still omits a full explanation of Kant's

likely reason for going through the traditional four kinds of imperatives with both the FUL and FH,

my discussion below intends to fill this gap.

In this way, my argument also joins recent commentators who have emphasized the

importance of understanding Kant's *Groundwork* as containing a defense of 'immediate' common

moral knowledge.[11] Since this recent literature on the method and structure of the *Groundwork* has so

far omitted a detailed discussion of Kant's aim of the 'derivation of imperatives of duty' in light of

his response to G. A. Tittel, my argument provides additional support to the reading of the

Groundwork as providing a defense of 'immediate' or 'unreflected' common moral knowledge.

Indeed, I believe that despite its invaluable contributions, recent scholarship pointing to Kant's

FLN as negative tests for maxims and the FH as providing a more concrete test for the derivation of duties. However, on this view the overall purpose of the 'derivation of imperatives of duty' in the *Groundwork* remains an exemplification of a decision procedure, albeit not a strict algorithm. Similarly, Barbara Herman (1993) has pointed out that the 'CI procedure' was not meant for a derivation of a set of strict rules, but argued that it should instead be understood as a 'rule for deliberation' that generates 'deliberative principles' or 'presumptions' that provide the baseline for subsequent moral judgement in difficult situations. Again, although Herman's early view moved away from the interpretative paradigm somewhat, the *Groundwork's* 'derivation of imperatives of duty' still remains a test that would generate 'deliberative principles'. Only in her more recent work (2011), has Herman suggested that 'we make a mistake if we regard the examples [of the FUL in section II of the *Groundwork*] as models of duty-derivation; their role in the text is clarificatory, to confirm at each stage of a very abstract argument about an a priori rational principle that we are in fact tracking morality' (2011 p. 51, emphasis added). While this comes close to a sound rejection of the long- standing assumption of taking Kant's examples in the *Groundwork* to exemplify a decision procedure for right and wrong action, her rejection remains far less radical than it appears on first sight. For Herman's reason for rejecting the standard reading of the *Groundwork* is that the alleged test only works when we presuppose a prior notion of moral literacy. As she argues, 'its [the *Groundwork's*] principles are not intended to be sufficient for moral judgment' (p. 51, emphasis added), and she consequently denies that 'the candidate supreme principle of morality [as exemplified in the *Groundwork*] is intended to be, on its own, generative of duties and obligations' (p. 52, emphasis added). However, as I hope to show in this article, no such view on prior moral literacy or the 'sufficiency' of either formula of the CI is necessary in order to doubt that the *Groundwork's* 'derivation of imperatives of duty' was supposed to exemplify a decision procedure.

[10] Instead, Geiger has argued that: 'the discussion of the FUL speaks of agents who generally know what their duties are and how these duties are fulfilled or violated in most all everyday situations, and they know these things without putting their intentions to any universalization test. The *Groundwork*, and the discussion of the FUL in particular, make explicit the implicit common moral knowledge of ordinary rational agents'. (Geiger 2010 p. 272)

[11] Callanan 2019; Grenberg 2013; cf. Sticker 2016.

defense of 'immediate' common moral knowledge retains a lacuna as long as it does not provide an

explanation of the purpose of Kant's 'derivation of imperatives of duty' in section II of the

Groundwork. For if the standard view of Kant's derivation as exemplifying a decision procedure was

correct, then it would imply that either Kant's moral philosophy is much more reformist than he

claims,[12] or that ordinary ethical life typically consists in applying universalization tests, or possibly

both.[13]

Revising our assumptions about Kant's derivation of imperatives of duty in section II of the

Groundwork also carries consequences for Kant scholarship in general. This passage is commonly

viewed as the most explicit passage in which Kant derives concrete duties from the Categorical

Imperative by means of a decision procedure; by contrast, his other references to universality and

humanity are not usually taken to be straightforward examples of the application of such a test.

Consequently, if Kant is not concerned with the introduction of a decision procedure in section II

of the *Groundwork*, Kant scholarship would have to look to other parts of Kant's opus for

paradigmatic examples of the role of the Categorical Imperative in moral deliberation, and this may

ultimately require us to rethink the practical purpose of Kant's moral philosophy more generally.

In section 2, I present Tittel's criticism of the *Groundwork* and Kant's response in the preface

to the *Critique of Practical Reason*. In section 3, I argue that both Kant's *Groundwork* as well as its

historical-philosophical background provide some pushback against reading Kant's infamous

'derivation of imperatives of duty' as examples of a decision procedure; instead, I argue that Kant's

infamous four examples in the *Groundwork* attempt to show that his analysis of the structure of

---

[12] For a defense of the view that Kant's ethics is indeed more revisionary than he admits, see Sticker 2016.
[13] For the view that Kant's ethics is both more revisionary than he admits and that his examples in the *Groundwork* – specifically in section I – describe the paradigmatic reasoning of common moral agents, see Sticker 2016 and 2021 pp. 11-12.

moral judgement underlies the general types of imperatives already recognized by his contemporaries. In section 4, I argue that the same is true of Kant's discussion of the FH in section II of the *Groundwork*. I conclude with some promissory remarks about the upshots of my reading.

No 'New Principle' but a 'Determinate Formula'

Less than a year after the 1785 publication of the *Groundwork*, a book-length commentary by G. A. Tittel reproached Kant for his failed attempts to 'reform' moral philosophy by providing an a priori foundation for a theory of right action. In the course of his criticism, Tittel covered a whole range of topics that are still echoed in contemporary criticisms of Kant: for instance, Kant's insistence that a goodwill is the only thing of unlimited moral value, prior to happiness,[14] the talk of 'pure' practical reason,[15] Kant's conception of obligation,[16] as well as his obscure terminology and 'mysticism'.[17] Importantly, Tittel also argued that Kant's 'highest principle of morality' would be practically useless. For, he claimed, the Categorical Imperative is entirely formal, and thus empty, and since it is empty, it would not generate any concrete duties unless it was supplemented with a calculation of consequences and with theoretical knowledge of which consequences would further people's happiness. Tittel continued to charge Kant with having smuggled just such empirical considerations into his discussion of the Categorical Imperative in the derivation of imperatives of duty.[18] Thus, Kant allegedly failed to give moral philosophy a practically useful a priori foundation.

---

[14] Tittel 1786 pp. 7-13, 37-38
[15] Ibid pp. 17-29
[16] Ibid pp. 29-32
[17] Ibid p. 4
[18] Ibid pp. 32-37

Elaborating on his criticism, Tittel asks: 'Is the entire Kantian reform of ethics indeed to confine itself to just a new formula?'[19] And he continues:

> For what help should an empty formula be that one has made to look purely rational, but that is unfit for practical use unless one supplements it with material of experience? It cannot help that Kant now says 'all duties depend upon this one principle of morality, as far as the kind of obligation, not the object of their action, is concerned'.[20] For one clearly saw in these examples [the four derivations of imperatives of duty] that the point of obligation, or rightfulness, must be determined only through the nature of the object and its relation to the human being's nature, through its consequences. The categorical imperative (so that I say it once and for all), together with its imprinted stamp of pure reason, appears to me to be a veritable plaything. 'Act so that you can will your maxim to become universal' may sound pretty and new. But if one searches for meaning, not words, then it can mean nothing else but: 'Act so as you think it good and beneficial, if all other beings (that can act so) acted this way'. Thereby a prior, and separately determined, concept of goodness and benefit is necessarily presupposed; or the entire formula does not make sense. I do not learn what is good from the formula; instead, the formula merely tells me: 'Do that which is good (generally beneficial)'. Not a single step am I taken further by the amended, seemingly new formula: 'Act so that you can will that your maxim become universal' than by one of the oldest known principles: Act so that perfection be furthered by your actions and maxims (taken in their entirety and universality). I still must first abstract the concept of the good and perception from the stuff of experience (empirically) through collection and comparison of cases.[21]

In summary, Tittel complains that Kant's 'new principle' does not provide us with a practically useful theory of right action – of determining which actions are right and which actions are wrong. Allegedly, the Categorical Imperative does not allow us to derive duties without presupposing empirical content and thinking about consequences, and once we have seen that Kant's 'new principle' smuggles empirical content into its test, we must conclude that his 'new principle' has

---

[19] Ibid p. 35. All translations of Tittel in this article are my own.
[20] Here Tittel slightly misquotes GMS 4:424, but still conveys the correct meaning of Kant's passage. There, Kant actually says: 'so all duties, as far as the kind of obligation (not the object of their action) is concerned, have by these examples been set out completely in their dependence upon the one principle'.
[21] Tittel 1786 pp. 35-37

failed to reform the moral discipline by giving it an a priori basis. For by itself, it is entirely useless as a decision procedure. If we apply Kant's alleged test, we either do not get any answer at all, or we must presuppose empirical content, which we can only get by comparing different actions and their results. But in the latter case, Kant's formulations of the Categorical Imperative do not do the promised work; instead, we must rely on consequentialist considerations.

In his preface to the *Critique of Practical Reason*, Kant responds to Tittel's criticism in a striking footnote. In that preface, Kant tells the reader that his second *Critique* 'presupposes, indeed, the *Groundwork of the Metaphysics of Morals*, but only insofar as this constitutes preliminary acquaintance with the principle of duty and provides and justifies a determinate formula of it;* otherwise, it stands on its own' (KpV 5:8, emphasis added). In the footnote concerning this 'determinate formula', Kant remarks as follows:

> A reviewer who wanted to say something censuring [the *Groundwork*] hit the mark better than he himself may have intended when he said that no new principle of morality is set forth in it but only a new formula. But who would even want to introduce a new principle of all morality and, as it were, first invent it? Just as if, before him, the world had been ignorant of what duty is or in thoroughgoing error about it. But whoever knows what a formula means to a mathematician, which determines quite precisely what is to be done to solve a problem and does not let him miss it, will not take a formula that does this with respect to all duty in general as something that is insignificant and can be dispensed with. (KpV 5:8)

On first sight – and especially to a reader unfamiliar with Tittel's commentary – this footnote might seem innocuous. After all, few present-day readers of Kant's ethics would follow Tittel's description of Kant's ethics as attempting to introduce a new principle into ethics. For Kant presents the Categorical Imperative as the form of pure practical reason. That is, Kant presents the Categorical Imperative not as a new principle but as the underlying principle that was there all along in our

practical judgements, and that his philosophy supposedly, for the first time, unearths in its precise form. Therefore, this footnote might not appear to demand much attention.

But on a second look, Kant's response becomes more interesting. As we saw, Tittel's reason for denying that the Categorical Imperative would be a 'new principle' was that he found it practically useless for the derivation of concrete duties. So why would Kant be happy to go along with this criticism if the purpose of the derivation of imperatives of duty in the *Groundwork* had been to exemplify a useful decision procedure for the derivation of concrete duties? If we read Kant's derivation of imperatives of duty as the introduction of a decision procedure, his response to Tittel should puzzle us.

Unfortunately, there are still few discussions of Kant's response to Tittel.[22] But his response is quite significant in the context of the present-day reception of Kant's ethics. For Kant here explicitly addressed someone who shared what I have called the standard assumption about Kant's *Groundwork* – namely that its examples of 'derivations of imperatives of duty' were supposed to demonstrate the application of a decision procedure – and then criticized Kant's book for not providing a principle that could be put to use as a test for the determination of concrete duties. And in this light, Kant's response to Tittel's criticism is striking: in his footnote, Kant decidedly did not insist on the practicality of the formulas of the Categorical Imperative presented in the *Groundwork* as providing a decision procedure for the derivation of duties, nor did he provide new examples of such a derivation or point to another passage elsewhere that might provide such examples. Instead, Kant's response was to question the reasonableness of the very idea that moral philosophers should attempt to introduce any 'new principle' of morality, as if people before had not known what duty is.

---

[22] See for instance Beck 1963 p. 57, Sala 2004 p. 65, Timmermann 2007 pp. xii-xiii, Schneewind 2009 p. 143, and Walschots 2020.

One may object to this reading that even if Kant was apparently not interested in insisting on the practical usefulness of the categorical imperative in his response to Tittel, he might have responded to this worry indirectly elsewhere. While it is important to distinguish Tittel's criticism of uselessness from the conceptually separate claim that Kant's ethics must be given a specifically consequentialist reading in order to be practical, Tittel joins these two charges together. Thus, if Kant had responded to the charge of implicit consequentialism elsewhere, might such a response not provide an indirect response to Tittel's charge of uselessness[23] and thus vindicate the standard reading of the *Groundwork*? I believe that this would be a premature conclusion. For even if Kant had such a response to the charge of implicit consequentialism elsewhere, he explicitly decided to respond to Tittel's separately. And what is at issue in Tittel's 'no new principle' objection and Kant's response is whether or not the formulas of the Categorical Imperative as presented in the *Groundwork* were meant to exemplify a decision procedure to be actively applied by agents to guide them in their moral deliberation and choice. Thus, even if Kant wanted to eventually provide a practically useful decision procedure elsewhere by way of answering the charge of implicit consequentialism, this would not change his response regarding the *Groundwork*.

Apart from providing a theoretical insight into the form of common moral judgement, what else, if anything, might his formula then be able to do? After all, Kant apparently believed that anyone who 'knows what a formula means to a mathematician' will understand its significance: it purportedly 'determines quite precisely what is to be done to solve a problem'. And is this not just another way of saying that the formulas of the Categorical Imperative should be used as a decision

---

[23] The charge of implicit consequentialism was also made by H.A. Pistorius, and Kant arguably responded to this charge of implicit consequentialism in the second *Critique's* 'The concept of an object of practical reason' and the 'Typic of pure practical judgment'. The fact that Kant arguably responds to both these charges separately indicates that even if he was concerned with the practicality of the Categorical Imperative elsewhere, he was not so concerned with the *Groundwork*. For discussion of Kant's response to Pistorius see Guyer (2021) and Walschots (2021).

procedure? If one were to read this remark with an emphasis on the 'what is to be done' without regard to the final formulation 'to solve a problem', one might indeed get the impression that Kant must be talking about a test or decision procedure that tells us what to do. However, this would be to read Kant's passage too selectively. Instead, the crucial question is what kinds of problems Kant might have had in mind here. Mathematical formulas help solve problems, so what problems does moral philosophy help solve? If Kant had in mind theoretical problems, which the formulas of the Categorical Imperative help us solve, they are presumably the philosophical questions he addresses in the work he was writing, the *Critique of Practical Reason*, and the later *Metaphysics of Morals*. By contrast, if Kant was thinking about practical problems, what might he have had in mind? Strikingly, Kant did not believe that figuring out what to do would be a practical problem requiring philosophical knowledge of the formulas of the Categorical Imperative. As Kant remarks repeatedly, he believed that anyone, even children, could judge 'without hesitation' what is right to do without engaging in any technical decision procedure or having been taught the formulas of the Categorical Imperative.[24] By contrast, the central problem Kant identifies as providing a practical reason for doing moral philosophy is the difficulty of overcoming rationalizations against prior (common) moral knowledge in favor of inclinations.[25] As Kant points out in section I of the *Groundwork*, '[t]he human being feels within himself a powerful counterweight to all the commands of duty, which

[24] See, e.g., MS 6:480-1 and TP 8:286.
[25] For a detailed discussion of this mechanism of rationalizing against morality in favor of inclinations and its place in Kant's argumentative strategy in the *Groundwork*, see especially Callanan (2019). For recent discussion of rationalizing in Kant more generally see Papish (2018) and Sticker (2021). The corresponding therapeutic practical purpose of moral philosophy in general has received comparatively less attention so far. For a brief discussion of the potential heuristic value of the FUL in the *Groundwork* see Geiger (2010 pp. 286-8). Since the aim of the present discussion is a careful exegesis of Kant's 'derivation of imperatives of duty' in light of his response to G. A. Tittel, it would be immodest to venture further into these surrounding topics. For the present purpose, it is only important to emphasize that Kant's remark about the 'problems' which a 'new formula' is said to help solve cannot simply be presupposed to refer to the determination of which actions are really right and wrong, but must be read in the context of his account of that very problem: the natural dialectic.

reason represents to him as so deserving of the highest respect…'. Because of this aspect of human nature, Kant adds that 'from this there arises a natural dialectic, that is, a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations…'. And according to Kant's remarks in the *Groundwork*, it is this natural dialectic that gives a practical purpose to moral philosophy:

> In this way common human reason is impelled … on practical grounds themselves, to go out of its sphere and to take a step into the field of practical philosophy, in order to obtain there information and distinct instruction regarding the source of its principle and the correct determination of this principle in comparison with maxims based on need and inclination, so that it may escape from its predicament … (GMS 4:405)

Importantly, this account of the practical difficulties that require moral philosophy for their solution also matches Kant's formulation of the four examples of the derivation of imperatives of duty. For the latter conspicuously take the form of agents already aware of what they ought to do, but who are in danger of rationalizing away their moral knowledge in favor of one selfish inclination or another. 'Someone feels sick of life because of a series of troubles that has grown to the point of despair', so much so that the person just has enough strength left to put themselves to reflect about their situation. 'Another finds themselves urged by need [durch Not gedrungen] to borrow money'. This is a person in despair (Not), not someone unclear about the morality of truthfulness. The third is tempted by the comfort of an easy and pleasurable life as opposed to the hassle of developing excellence in their natural gifts. Only the fourth seems to be doing alright for themselves and is merely tempted by a desire to ignore the hardship of others in order to enjoy their own fortune unimpeded (GMS 4:421-23). In none of these examples did Kant choose a merely puzzled agent; rather, they all struggle against the temptation to rationalize away their common moral knowledge.

114

Still, one might object that, surely, the analogy with a mathematical formula solving a

problem must entail a test or decision procedure. Taking one of the (if not the) best-known

mathematical formulas, the Pythagorean theorem, as an example, readers are frequently tempted to

think of Kant's analogy between a formula of duty and a mathematical formula in the following way.

The formula of duty might stand to duty as $a^2 + b^2$ stand to c^2. On this view, $a^2 + b^2$ would be

analogous to the formulas of the Categorical Imperative and c analogous to duty. On this

interpretation of the mathematical analogy, the Categorical Imperative would provide a decision

procedure for the determination of a duty, as the root of $a^2 + b^2$ determines c. However, this way of

interpreting the analogy would be misleading. For $a^2 + b^2$ is only part of the formula, the formula

being $a^2 + b^2 = c^2$. In the analogy with the Pythagorean theorem, duty is not analogous to c; rather,

duty is analogous to a right triangle. Thus, in Kant's analogy comparing a formula of duty to a

mathematical formula, $a^2 + b^2 = c^2$ would be analogous to the Categorical Imperative, and duty

analogous to a right triangle. Consequently, the mathematical analogy does not support interpreting

the Categorical Imperative as an algorithm for determining duty as $a^2 + b^2$ determines c^2. How

exactly a formula can be practical depends on the respective formula and the problems in question.

But here, the mathematical analogy ends. Which problems the formulas of the Categorical

Imperative can help us solve depends on what Kant believed made moral judgement difficult, not

on mathematics. Consequently, Kant's remark that a formula 'determines quite precisely what is to

be done to solve a problem' does not entail that he would have been concerned with exemplifying a

decision procedure in the *Groundwork*.

If Kant's own response to Tittel shows him as remarkably uninterested in defending the practical usefulness of the formulas of the categorical imperative as a test or decision procedure, how has the standard reading first become the uncontroversial assumption among readers of the *Groundwork*? Due to its popularity, it has so far received little explicit defense. But two things conjointly make the standard reading so intuitively plausible on first sight: Kant's metaphorical remark about the fundamental principle of moral judgement being a 'compass' in section I and Kant's terminology of 'derivation' in section II. I briefly consider the former, before discussing Kant's terminology of 'derivation' and offering an alternative reading.

After arriving at a first formulation of the Categorical Imperative in the Universal Law formulation in the first section of the *Groundwork*, Kant tells us that:

> Common human reason also agrees completely with this in its practical appraisals and always has this principle before its eyes. … [T]o this problem, whether a lying promise is in conformity with duty, I ask myself: would I indeed be content that my maxim (to get myself out of difficulties by a false promise) should hold as a universal law (for myself as well as for others)? and could I indeed say to myself that everyone may make a false promise when he finds himself in a difficulty he can get out of in no other way? Then I soon become aware that I could indeed will the lie, but by no means a universal law to lie … Thus, then, we have arrived, within the moral cognition of common human reason, at its principle, which it admittedly does not think so abstractly in a universal form but which it actually has always before its eyes and uses as the norm for its appraisals. Here it would be easy to show how common human reason, with this compass in hand, knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty or contrary to duty, if, without in the least teaching it anything new, we only, as did Socrates, make it attentive to its own principle; and that there is, accordingly, no need of science and philosophy to know what one has to do in order to be honest and good, and even wise and virtuous. (GMS 4:403–4, emphasis added)

Here, Kant speaks metaphorically of an innate principle as a 'compass' that we always have before our eyes that makes moral judgement possible; and yet, importantly, we do not make use of this principle by consciously thinking about it in its abstract form.[26] On the standard reading of the *Groundwork*, Kant is here alluding to an innate formal principle that also serves as a decision procedure in moral deliberation, which he would then exemplify in his derivations of imperatives of duty. However, it is important to keep in mind that such an innate principle, the form of moral judgement, need not necessarily figure in agents' conscious deliberation – as Kant makes clear, common moral judgement does not usually 'think so abstractly' through some universalization test.[27] And correspondingly, examples of common moral judgement conforming to this principle may not be meant to exemplify a decision procedure for practical deliberation but 'to confirm at each stage of a very abstract argument about an a priori rational principle that we are in fact tracking morality.'[28] Importantly too, we should not presuppose a reading of Kant's compass-metaphor that is already influenced by previous assumptions about the 'derivation of imperatives of duty' in section II; in light of Kant's response to Tittel, it is worth reconsidering his infamous derivations themselves.

According to the standard assumption about the *Groundwork*, section II reiterates the highest principle of morality that Kant had just excavated from common moral knowledge in the form of the Formula of Universal Law: 'act only in accordance with that maxim through which you can at the same time will that it become a universal law' (GMS 4:421) and then continues to demonstrate how this formula of his highest principle of morality helps us figure out what is truly right and

---

[26] Kant here suggests that the highest principle of moral knowledge makes a certain kind of cognition possible without (necessarily) consciously featuring in our thoughts. In this regard, we might think about the form of practical reason as at least partly analogous to that which is formal in theoretical reason. For an insightful exploration of this parallel see Pippin (2013).

[27] See also KpV 5:36 and TP 8:287.

[28] Herman 2011 p. 51

wrong. This demonstration allegedly takes the form of a derivation of four duties from the Categorical Imperative: the duty not to take one's own life, the duty not to lie to others, the duty to cultivate one's talents, and finally, the duty to help others in need. And at first sight, the wording of section II of the *Groundwork* indeed suggests this ambitious reading. Immediately after introducing the Formula of Universal Law, Kant says:

> Now, if all imperatives of duty can be derived from this single imperative as from their principle, then, even though we leave it undecided whether what is called duty is not as such an empty concept, we shall at least be able to show what we think by it and what the concept wants to say. (GMS 4:421)

The first half of the sentence clearly talks about deriving 'imperatives of duty'. And after a quick digression into the right ways of formulating the Categorical Imperative, Kant then announces: 'We shall now enumerate 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.' (GMS 4:421) – followed by his famous discussion of the four canonical duties of his discussion of morality. Hence, on first sight, it might seem that Kant's main concern here is to demonstrate how the Categorical Imperative might enable us to derive concrete duties from it.

But before jumping to old conclusions, it is helpful to keep some distinctions in Kant's terminology in mind: Kant here talks about deriving imperatives of duty, and then says that he will demonstrate this by enumerating (herzählen) 'a few duties.' In Kant's terminology, imperatives are distinguished from both duties and obligation. The clearest statement of this distinction can be found in Kant's introduction to the *Metaphysics of Morals*: 'An imperative is a practical rule by which an action in itself contingent is made necessary' (MS 6:222). By contrast, '[o]bligation is the necessity of a free action under a categorical imperative of reason' (MS 6:222). And '[d]uty is that action to which someone is bound. It is therefore the matter of obligation …' (MS 6:223). In other words, a

118

duty is the specific action type that an imperative commands (something that it is right to do),[29] and

an 'imperative' is the representation of the action as something to be brought about.[30] But we need

not look ahead to the *Metaphysics of Morals*. Already in his discussion leading up to the 'derivation of

imperatives of duty', Kant makes clear that imperatives are specific ways of representing action.

There, he repeatedly points out that only through representations can there be practical judgement,

and that these representations are properly called 'imperatives.' More precisely, 'imperatives' are the

linguistic formulations that attempt to express the representation of a necessary (free) action. This is

perhaps most clearly expressed in the following passages:

> The representation of an objective principle, insofar as it is necessitating for a will, is called a
> command (of reason), and the formula of the command is called an imperative. (GMS 4:413,
> emphasis added)

> Now, all imperatives command either hypothetically or categorically. The former represents the
> practical necessity of a possible action as a means to achieving something else that one wills (or that it
> is at least possible for one to will). The categorical imperative would be that which represented an
> action as objectively necessary of itself, without reference to another end. Since every practical law
> represents a possible action as good and thus as necessary for a subject practically determinable by
> reason, all imperatives are formulae for the determination of action that is necessary in accordance
> with the principle of a will which is good in some way. (GMS 4:414, emphases added)

With these terminological distinctions in mind, it is possible to think about another purpose than the

provision of a test or decision procedure for Kant's derivation of imperatives of duty: Kant's aim in

section II of the *Groundwork* is to argue that all representations of actions as intrinsically good share

---

[29] Note, however, that Kant is not entirely consistent in his usage of the terms *Verbindlichkeit* (obligation) and *Pflicht* (duty). At GMS 4:439, for instance, he appears to use *Verbindlichkeit* (obligation) and *Pflicht* (duty) in exactly reverse ways. At MS 6:379 Kant adds that the 'concept of duty is already the concept of a necessitation' and the imperative 'makes this constraint known'. Although his formulation here is vaguer, it is still possible to read this sentence (in a way that makes Kant's view overall coherent) as holding that the concept of an action that one has to perform entails the concept of a constraint of free choice, which is represented through an imperative.

[30] Kant reserves the term 'imperative' for the practical laws guiding the actions of imperfect beings, i.e., beings that can fail morally. In a perfect being, actions would still be represented as necessary by some practical laws, but these practical laws would not be accompanied by a subjective force that might run against the being's inclinations.

in the form of one single highest principle of morality, the Categorical Imperative. Thus, Kant first claims that 'all imperatives of duty can be derived from this single imperative [the categorical imperative] as from their principle'.[31] Then, he enumerates a handful of examples of duties (action types) that he believes his readers will readily agree to be morally obligatory in order to substantiate his claim. And finally, he attempts to demonstrate that despite their many differences, the respective imperatives through which practical reason represents the necessity of actions always share the same form – whether people describe them as perfect, or imperfect, towards others or oneself.[32]

Before going through the four examples in his argument, Kant also tells us about his motivation for this argument. He thinks it important to 'enumerate 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' (GMS 4:421, emphasis added). The infamous distinction between perfect and imperfect duties does not originate with Kant but with the writers of the Natural Law tradition.[33] Its most likely origin is the work of Hugo Grotius, and it was taken up in the works on moral and political philosophy by such writers as Pufendorf and Wolff, with whose work on natural law Kant was familiar.[34] Traditionally, this distinction was supposed to signify a distinction between duties that should or should not be enforced. And at the time of the publication of the *Groundwork*, it appears to have been an open philosophical question how this traditional distinction between perfect and

---

[31] The qualification 'of duty' singles out categorical from hypothetical and prudential imperatives and thus narrows the scope of the remark to moral imperatives. At the beginning of the *Groundwork*, Kant started off his investigation into the nature of morality by analyzing the idea of moral goodness in action and argued that the idea of duty is fundamental to the idea of morally good actions; hence, the qualification 'of duty' clearly indicates a qualification like the term 'moral', albeit in more Kantian terminology.

[32] This special emphasis on the types of imperatives has not gone unnoticed by Kant scholars. See for instance Allison 2011 p. 179.

[33] For a comparison of their respective conceptions of perfect and imperfect duties in Grotius, Pufendorf, Wolff, and Kant, see Salam 2014 pp. 19–44.

[34] Although, to my knowledge, it is not clear if Kant owned copies of their work, his personal library as indexed for the purpose of an auction in 1922 contained books in which Grotius and Pufendorf figure prominently, such as Burlamaqui (1747), von Vattel (1760) and Huber (1735). For a catalogue of Kant's personal library see Warda (1922).

imperfect duties could be explained and justified.[35] This distinction might easily lend itself to the thought that the representation of the necessity of an action by practical reason differs correspondingly. Given the philosophical context, Kant's remarks in the second part of the *Groundwork* on his 'derivation of imperatives of duty' must be read as acknowledging this debate about different kinds of duties and rejecting the thought that the will's representation of actions as to be done would differ correspondingly.[36]

But in what sense is this argument that all practical representations of actions as intrinsically good share a common form a derivation of imperatives of duty? I take it that Kant uses the term 'derive' (ableiten) quite consistently to mean simply the drawing of a conclusion from a principle: 'every syllogism is a form of derivation [Ableitung] of a cognition from a principle' (A300/B357). Thus, when Kant uses the term 'derive' here at GMS 4:421,[37] he likely means that there can be arguments in which the formula of the Categorical Imperative serves as a major premise, and from which we can draw conclusions about the form of other commonly accepted imperatives of morality. A helpful analogy might be to compare Kant's argument in the 'derivation of imperatives of duty' with some hypothetical argument about formal logic. Kant claims that the form of the commonly accepted moral imperatives just is the form of the Categorical Imperative. His argument proceeds by taking four representations of actions (maxims) commonly thought contrary to duty and shows how these four representations also contradict the form of the Categorical Imperative. By analogy, we might claim that the forms of certain commonly accepted logical functions in fact share

---

[35] For a historical overview of the different ways of explaining this traditional distinction from Grotius onward, focusing especially on debates in the second half of the eighteenth century, see Kersting (1982).

[36] Note, however, that Kant here is interested in the 'usual division' for the reasons described above; he is not claiming that this 'usual' division is also the only, the best, or the final way of dividing duties. Consequently, Kant explicitly reserves, in a footnote, his final word on the matter to his future *Metaphysics of Morals* (GMS 4:422n).

[37] And in a similar context later at GMS 4:429, which I discuss in more detail below.

in the form of another logical function – take, for instance, A ∨ B, ¬B → A and (P | P) | (Q | Q).

In an argument resembling Kant's 'derivation of imperatives of duty', we might try to show how

contradicting one logical function also means contradicting the other logical functions. And the

conclusion we might draw is that the different functions in fact share a specific form. In a similar

way, I believe, Kant wants to show that the form of the Categorical Imperative is indeed the form

underlying all representations of actions as intrinsically good.[38]

At this point, one might object that surely Kant's central concern must have been to provide

a decision procedure in section II of the *Groundwork* because he talks about a 'derivation' after the

conclusion of his notorious four examples. In most contemporary English translations, Kant is there

recorded as saying: 'These are a few of the many actual duties, or at least of what we take to be such,

whose derivation [Ableitung] from the one principle cited above is clear' (GMS 4:423-4). But as

these translations usually point out, this is not what Kant says in the original. In the first edition of

the *Groundwork*, Kant says '[t]hese are a few of the many actual duties, or at least of what we take to

be such, whose division [Abtheilung] from the one principle cited above is clear.'[39] Starting with

Gustav Hartenstein in 1838, some editors – and most English translators[40] – have decided to replace

Kant's Abtheilung (division, or classification) with the term 'derivation'.[41] Unfortunately, the

---

[38] I am not concerned here with the question whether, ultimately, Kant's argument is either stringent or even plausible. I am only concerned with showing the argument's purpose.

[39] Kant 1785 p. 57

[40] With some notable exceptions: Allen Wood has translated the term as 'partitioning' in his translation of the *Groundwork* (Kant 2002b p. 40), Jens Timmermann has translated the term as 'division' in his 2011 revision of Mary Gregor's translation (Kant 2011 p. 75), and Thomas Abbott translated the passage as reading '…which obviously fall into two classes on the one principle that we have laid down' (Kant 1895 p. 49).

[41] The likely reason Hartenstein decided to exchange 'division' for 'derivation' is the odd style of the sentence. A derivation from a principle is an inconspicuous expression; by contrast, a division from sounds significantly less elegant. And while Kant uses the equivalent expression 'division from' ('Eintheilung aus' – see, e.g., Opus Postumum, 22:501), I am personally not aware of another passage in Kant's work containing the precise expression 'Abtheilung aus'. Either way, it is not obvious to me that speaking of a 'division from' is philosophically as odd as it is stylistically wanting. After all, Kant seemed concerned to show both that (i) the traditionally accepted types of imperatives share the form of the highest principle of morality while also (ii) not picking a fight with the traditional division. He did the latter by drawing a

contemporary German terms for division (Abteilung) and derivation (Ableitung) are anagrams, which might predispose contemporary readers to assume a mistake in the original printing – however, they were not during Kant's time. In Kant's time, the spelling contained an additional h (Abtheilung) (which is the spelling of the original editions of the *Groundwork* as well as in volume IV of the Akademie Ausgabe, first published 1903). This alone makes it rather unlikely that the original wording would have been a mistake. Moreover, the context of the passage does not obviously support reading 'derivation' instead of the original 'division'. In this passage, Kant argues that maxims that violate perfect duties cannot even be thought as universal laws of nature, whereas those maxims that violate only imperfect duties could be thought as universal laws of nature but cannot be willed as such universal laws. As Allen Wood has pointed out, this makes it sound like Kant is more concerned with the classification of distinct types of duties in this passage than with a potential derivation of duties.[42] Moreover, even on the assumption that Kant really meant to write 'derivation', it would still be far from obvious that Kant's reference to a derivation of duties could not be understood as shorthand for the derivation of types of duties as they are commonly recognized as indicated by the foregoing discussion of the *Groundwork*. In short, the sentence following the infamous four examples is not good evidence in favor of the standard assumption.

The reading of the *Groundwork* I have defended so far is further supported by Kant's last sentence in the very same passage following the 'derivation of imperatives of duty'. Kant there concludes his discussion of the four famous examples by pointing out that 'so all duties, as far as the

---

distinction between the ways in which these traditional imperatives share the form of the highest principle of morality, namely by distinguishing between contradictions in conception and in the will. Thus, to speak somewhat awkwardly of a division 'from' or 'out of' a principle does not seem entirely unreasonable. I therefore agree with Allen Wood and Jens Timmermann in keeping close to the original expression with 'partitioning' and 'division' in their respective translations (see above, n. 31). I am grateful to Ido Geiger for pressing me to clarify my view on this point.

[42] Wood 2017 pp. 26-27. For discussion of this replacement of 'division' with 'derivation' see also Wood in Kant 2002b pp. 40-42, Kuehn 2010 p. 23, and Timmermann 2007 pp. 86-87.

kind of obligation (not the object of their action) is concerned, have by these examples been set out completely in their dependence upon the one principle' (GMS 4:424, emphasis added). Recall that in Kant's terminology, an imperative is the representation of the necessity of an action, and the obligation is the subjective necessitation that results from the representation of the necessity of an action. Thus, by demonstrating that the traditionally distinguished imperatives of morality are all based on the same underlying principle, Kant also takes himself to demonstrate that the 'kind of obligation' is equally based on the same underlying principle: the categorical imperative. In other words, by showing that the representation of the necessity of an action has the same form in all instances of duties traditionally distinguished, Kant also believes himself to show that the kind of obligation an agent takes herself to be under depends in each case on the same form of representation. This indicates once more that the purpose of the 'derivation of imperatives of duty' was to show how moral imperatives of all classes share a particular form, rather than providing examples of a decision procedure.

Kant's Introduction of the Formula of Humanity

Kant's introduction of the 'Formula of Humanity' (FH) in the *Groundwork* is the second passage apparently aimed at providing a test or decision procedure for the derivation of concrete duties from the Categorical Imperative. In the forgoing sections of the *Groundwork*, Kant had first argued that moral cognition concerns the representation of actions as good independently of inclination (a good will) and subsequently argued that the foundational principle of the capacity to represent actions in this way (the Categorical Imperative) could be captured adequately by the Formula of Universal Law. In the second section of the *Groundwork*, Kant proceeds with his analysis

of this capacity to represent actions as morally good and to determine oneself to action through such representations. He then eventually concludes that such a capacity supposes that every rational being must represent her own existence, and the existence of every other rational being, as an end in itself (i.e., as capable of setting oneself ends autonomously).[43] This fact, that rational beings must represent their own and others' existence as an end in itself, must serve as 'an objective principle from which, as a supreme practical ground, it must be possible to derive all laws of the will' (GMS 4:429, emphasis added). He continues: 'The practical imperative will therefore be the following: 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. We shall see whether this can be carried out' (GMS 4:429). Kant then turns back to the examples of duties discussed earlier in his 'derivation of imperatives of duty' to see if 'all laws of the will' can indeed be derived from the new formulation of the Categorical Imperative offered here.

Once again, on first sight this might appear as an attempt to derive duties from the Categorical Imperative by means of another decision procedure – this time, from the FH. However, I believe that we should again be cautious about such an interpretation of the passage. For one, the fact that Kant moves through the same examples suggests that the point of his discussion of the FH is serving the same purpose as his discussion of the FUL in the derivation of imperatives of duty discussed above. Earlier, Kant attempted to demonstrate that the form of all four types of imperatives commonly recognized is that of a universal law; now, he attempts to show that the form of practical reason also entails a special status of those beings who represent actions through it. In each case, Kant can only demonstrate this by going through examples of each type of imperative

---

[43] For the purpose of this article, I leave out Kant's argument for this inference.

commonly recognized by his contemporaries, in order to give some plausibility to the claim that he is really unearthing the form of all moral cognition.

Importantly too, note that Kant explicitly speaks about 'laws of the will'. While it can be frustrating to have an author use the term 'law' in so many ways, we should be careful when conflating different kinds of laws with each other in Kant's text. In this context, 'laws of the will' are the laws through which rational beings represent actions and ends as good: imperatives. Recall that earlier in the *Groundwork*, Kant had described the capacity to represent moral actions in this way:

> The categorical imperative would be that which represented an action as objectively necessary of itself, without reference to another end. Since every practical law represents a possible action as good and thus as necessary for a subject practically determinable by reason, all imperatives are formulae for the determination of action that is necessary in accordance with the principle of a will which is good in some way. (GMS 4:414, emphasis added)

For the purpose of the present discussion, what matters is the connection between what Kant calls a practical law that represents a possible action as good and what he calls the imperative: practical laws of the will and imperatives are just one and the same thing. It is through certain laws that reason represents actions as necessary. The different kinds of practical laws are different kinds of imperatives.

Thus, when Kant goes through the four examples of duties again here, he is doing two things. First, he strengthens his previous claim that the different kinds of imperatives commonly recognized at his time share the same form. Second, Kant's previous argument regarding the FUL in the derivation of imperatives of duty did not imply that it perfectly, or even sufficiently, captured the form of representation of moral actions. In his discussion of the FH, Kant further specifies the form that all moral representation of ends supposedly share: not only are all moral imperatives universal, but according to Kant they also concern the status of rational beings as ends in themselves.

126

Concluding Remarks

In light of Kant's reluctance to insist on the practical usefulness of the categorical imperative as a decision procedure, and in light of the historical and philosophical context of the *Groundwork*, we have good reasons to doubt the standard interpretation of Kant's infamous derivation of imperatives of duty as providing examples of a decision procedure for the derivation of concrete duties in moral deliberation. On a closer look, Kant's argument is concerned with showing that the Categorical Imperative, as the form of practical reason, underlies all commonly recognized types of imperatives. And in his response to Tittel in the preface to the *Critique of Practical Reason*, Kant explicitly replied to someone who adopted the standard reading of the *Groundwork* and criticized it for failing to provide a practically useful test for the derivation of concrete duties. In this light, his remark that Tittel understood his *Groundwork* better than Tittel himself seemed to have realized can help us better appreciate the purpose of Kant's derivation of imperatives of duty. For if we read Kant's derivation of imperatives of duty as demonstrating that his highest principle of morality really underlies all imperatives of duty commonly recognized, then Kant's response makes perfect sense: there was no need to argue about the practicality of the Categorical Imperative as a decision procedure because the *Groundwork* just did not intend to provide a new principle 'fit for practical use' in this way.

If my exegesis of Kant's derivation of imperatives of duty was correct, then we cannot focus entirely, or even primarily, on the *Groundwork* in order to understand if, and how, the formulas of the Categorical Imperative were supposed to figure in practical deliberation. A full discussion of the practical purpose of the formulas of the Categorical Imperative would go far beyond the scope of

127

this article. Here, I merely want to gesture toward some larger consequences for how we may understand the practical purpose of Kant's moral philosophy more generally.

If Kant believed that either (i) pre-philosophical, common moral knowledge knows right from wrong by applying a test or decision procedure along the lines of the formulas presented in the *Groundwork*, or (ii) philosophically informed moral reflection would proceed through such a decision procedure, then we should expect Kant to exemplify or describe this decision procedure somewhere. So far, Kant's derivation of imperatives of duty in the *Groundwork* has commonly been viewed as the most explicit such example – but if this standard assumption is ill-founded, where else should we look?

The most important passage outside the *Groundwork* in which Kant might appear to describe a decision procedure for moral deliberation is the 'typic' of pure practical judgement in the *Critique of Practical Reason* (KpV 5:69). As some authors have suggested, Kant there arguably responds to Pistorius' charge of implicit consequentialism. And since Tittel's objection was that Kant's ethics could not be practical unless it availed itself of consequentialist considerations, one might arguably take the typic to be Kant's implicit insistence on the practical usefulness of the Categorical Imperative.[44] As mentioned above, this still leaves intact Kant's response to Tittel as indicating that the *Groundwork* itself was not concerned with exemplifying a decision procedure – rather, this would indicate that Kant might have introduced such a decision procedure in the typic. However, it is controversial if the typic serves chiefly as exemplifying a decision procedure to be used by agents in

---

[44] As mentioned above, Kant's disparaging response to Tittel's charge of uselessness does not explicitly direct us to any other part of his work where he might provide either a direct or indirect response like this. Given that Kant did not respond to many criticisms, one may not think much of this fact. But given that the typic (KpV 5:69) is contained in the very same book as his response to Tittel (KpV 5:8), it is at least odd that Kant would choose to separate his responses like this if the typic had also been meant to provide an indirect response to Tittel's charge of uselessness.

practical deliberation, or an unconscious cognitive mechanism.[45] Although I am skeptical of reading

the typic as the introduction of a decision procedure, even this reading would benefit from my

present argument. Taking the typic to describe a decision procedure suggests a break with the

Groundwork, since in the second *Critique* neither the Categorical Imperative itself nor its formulas, but

a type of the moral law, would provide the decision procedure.[46] However, if my reading of the

Groundwork is correct, then proponents of this reading of the typic would have no need to posit any

break in Kant's ethics with respect to the derivation of imperatives of duty, for the former was

simply not in the business of exemplifying a test or decision procedure.

Apart from this brief section of the second *Critique*, we might also look to the *Metaphysics of*

Morals, Kant's work on pedagogy, and the 'Common saying' essay. However, the reason Kant

scholarship has focused on the derivation of imperatives of duty in the *Groundwork* for examples of

how the Categorical Imperative ought to figure in moral deliberation is that few passages elsewhere

provide straightforward examples of an agent applying a decision procedure. Despite Kant's brief

remark about the Categorical Imperative as allowing us to 'test' maxims for their universality in his

introduction to the *Metaphysics of Morals* (MS 6:225), commentators have frequently noted the almost

complete absence of the Formula of Universal Law in the work;[47] and it has also been questioned

whether Kant's references to 'humanity' throughout the *Doctrine of Virtue* can plausibly be read as

constituting derivations from the Formula of Humanity.[48] Similarly, Kant's 'casuistical questions' do

not contain the application of any test or decision procedure. And Kant's remarks about common

---

[45] For recent discussion see Westra (2016), suggesting that the typic's described cognitive process operates unconsciously and a typic procedure may just additionally 'amplify one's conscience'. Cf Zimmermann (2015), who takes the typic to describe a decision procedure.
[46] Zimmermann 2015 pp. 439-40
[47] See for instance Herman 1993 p. 133, Lottenbach and Tenenbaum 1995 p. 228, Korsgaard 1996 p. 124, Wood 1997 pp. 4, 11-12, Wood 1999 pp. 139-40, Hill 2007 p. 486, Geiger 2010 p. 274, and O'Neill 2013 p. 33.
[48] Kuehn 2010 pp. 23-27

moral cognition in his lectures on pedagogy, his 'fragments of a moral catechism', as well as his 'Common saying' essay do not portray agents as walking through any kind of decision procedure either.[49] Finally, the same applies to Kant's notorious essay 'On a supposed right to lie from philanthropy'. There, Kant provides a brief paragraph that might appear like an explanation of why truthfulness is really a universal, unconditional duty; but even there, Kant's actual aim is to argue that being truthful is not a duty to a specific person regarding a particular object but a duty to humanity in general (VRML 8:426). Consequently, if we discount the *Groundwork* as providing examples of a decision procedure for the derivation of concrete duties, then the passages where Kant seems to describe anything like the application of such a test are both remarkably rare and remarkably less straightforward than we might have expected. And given how rarely Kant describes anything resembling a decision procedure on the basis of the formulas of the Categorical Imperative outside the *Groundwork*, we may ultimately have to conclude that Kant was simply not interested in general in providing a new principle that would be 'fit for practical use' for the derivation of concrete duties – in other words, that for Kant the practical purpose of moral philosophy was not primarily in giving us a test or decision procedure.

Consequently, this revision of our background assumptions about Kant's moral philosophy could open up a newly unprejudiced engagement with his remarks about the natural dialectic and the practical necessity of moral philosophy (see section 2, above). For if the practical purpose of moral philosophy is to help us overcome the natural dialectic, then we should not expect – and certainly must not presuppose – this therapeutic benefit of moral philosophy to function merely, or even primarily, in terms of a test or decision procedure.

---

[49] See MS 6:480-81; TP 8:286; KpV 5:36, 79-80, 161-62; SF 7:63; TP 8:284-87.

Chapter 5

A Sudden Change of Mind? Kant's Project of a 'metaphysics of morals' in the *Groundwork*

A Puzzle About the 'Purity' of a Metaphysics of Morals

Kant's plan to write a 'metaphysics of morals' long predates the final *Metaphysics of Morals* in two volumes that was published in 1797. Already in 1768, he mentioned a project of that name in a letter to Herder. Since Kant never produced such a work during his pre-critical period, we can only speculate as to what it might have contained. But by the time he wrote the *Critique of Pure Reason*, the project of a 'metaphysics of morals' appears to have taken its mature shape. Kant's description of the project in the prefaces to both editions of the first *Critique* and in his chapter on 'the architectonic of pure reason' correspond neatly with his comments in the introductions to the *Metaphysics of Morals* and the *Doctrine of Virtue*, as well as with comments made during his lectures on moral philosophy in the 1780s. Nevertheless, some readers have questioned the consistency of Kant's project of a 'metaphysics of morals' throughout his critical period. Specifically, they have found Kant's description of a 'metaphysics of morals' in the *Groundwork* to be incompatible with his description of the project in the final *Metaphysics of Morals*: while both Kant's comments in the *Critique of Pure Reason* as well as in the *Metaphysics of Morals* suggest that a 'metaphysics of morals' will consist of a union of a priori principles and knowledge of human nature, and thus cannot be a system of 'pure' a priori cognition alone, Kant's remarks in the *Groundwork* have frequently been taken to suggest the opposite. This has led some prominent readers to doubt whether the *Metaphysics*

of Morals is really a part of Kant's critical system at all.[1] In this vein, Jens Timmermann has devoted the majority of his entry on 'Metaphysik der Sitten' in the De Gruyter Kant-Dictionary to the *Groundwork*, only to briefly claim at the end of his entry that it is not clear if the *Metaphysics of Morals* is actually the work which was envisaged in the *Groundwork* as a metaphysics of morals.[2] Perhaps, then, Kant suddenly changed his mind about what a metaphysics of morals would contain when he came to write the *Groundwork*?

This skepticism vis-à-vis the *Metaphysics of Morals* is motivated not only by Kant's brief description of the project of a 'metaphysics of morals' in the preface to the *Groundwork*, but perhaps equally so by a common reading of section II of the same work. Kant entitles this section the "Transition from popular moral philosophy to the metaphysics of morals." Therefore, we can expect that whatever Kant did in section II of the *Groundwork* was meant to provide a blueprint for what a later, more elaborate metaphysics of morals would contain. Section II of the *Groundwork* has traditionally been understood as aiming to establish a test or decision-procedure from which we can then derive duties for all rational beings – the underlying assumption being that only once we would have such a test could we come to know what morality really requires. Consequently, such a test would have to be established independently of any substantive moral knowledge, which could only succeed an application of this test to our specific situation. Such a 'pure' metaphysics of morals

---

[1] See for instance Anderson 1923 p. 41, Siep 2009 pp. 77-85, Timmermann 2007a p. 168-172, and Timmermann 2015 pp. 1542-1543. Similarly, Volker Gerhardt has argued that Kant's remarks in the preface to the *Groundwork* is far more restrictive than his moral philosophy would allow (Gerhardt 2009). According to Nandi Theunissen, "The Metaphysics of Morals indeed contains metaphysics* so understood, but it is a strangely mixed kind of book. It is a mix of metaphysics as critique, [to investigate the capacity of reason to have substantive a priori knowledge] of metaphysics* [laying out the substantive truths arising out of pure reason], and of the various divisions of practical philosophy which lie outside of the system of metaphysics strictly speaking, and to which any mention of specifically human duties belongs." (Theunissen 2013 pp. 110-11) Ralf Ludwig (1991) has also argued that Kant's *Doctrine of Virtue* introduces a "new model" of moral philosophy (p. 297) which is supposedly incompatible with his critical philosophy. Compare also Kim 2009 p. 337.

[2] Timmermann 2015 pp. 1542-1543.

could thus provide the test that would determine unconditionally for rational beings what is right and wrong for their circumstances. If that was indeed the aim of Kant's transition to a metaphysics of morals in the *Groundwork,* then the *Metaphysics of Morals* of 1797 could not be part of the same project. Kant's introductions to the *Metaphysics of Morals* and the *Doctrine of Virtue* (as well as the actual content of that book[3]) make that clear.

However, we have good reason to be skeptical of this received reading of Kant's infamous 'derivation of imperatives of duty' in section II of the *Groundwork*. As I have argued elsewhere,[4] it seems unlikely that Kant intended his discussion of the formulas of the Categorical Imperative in the *Groundwork* to supply a test or decision-procedure form which a system of concrete duties would eventually be derived. In other words, when Kant was trying to articulate the a priori form of practical reason he did not intend this form to function as a test that could be separated from our knowledge of human nature and moral ends. And if we do not gratuitously presuppose such a problematic standard reading of section II, then we can read both Kant's discussion in section II of the *Groundwork* as well as his descriptions of the project of a metaphysics of morals in the preface to the *Groundwork* as compatible with both his earlier and later remarks about a metaphysics of morals. As I argue in this chapter, Kant's project of a metaphysics of morals remains consistent throughout his critical period.

In the first section of this chapter, I bring together Kant's descriptions of the project of a metaphysics of morals in the *Critique of Pure Reason*, the *Groundwork* and the *Metaphysics of Morals*, which have frequently appeared incompatible to readers. In section two, I argue that Kant's

---

[3] I further discuss the alleged tension between the content of the *Metaphysics of Morals* and Kant's earlier work on moral philosophy at length in 'A System of Moral Knowledge'.
[4] See chapter 4, adapted from Ramsauer 2024

description of a metaphysics of morals as concerning both systematic a priori cognition as well as relying on knowledge of human nature remains consistent throughout his critical period, and is not only illustrated in his *Metaphysics of Morals* but also schematically in section II of the *Groundwork*. In the previous chapter, I argued that we should read the four notorious examples of Kant's "derivation of imperatives of duty" in section II as showing how the structure of practical reason underlies all four types of commonly recognized moral imperatives; drawing on this reading of Kant's derivation of imperatives of duty as well as Kant's remarks about the purity of a 'metaphysics of morals,' section three outlines how this reading allows us to understand the progression of Kant's argument in second section of the *Groundwork*. Specifically, I suggest that this allows us to read Kant's project of a metaphysics of morals as an attempt to analyze the rational principles that underlie our moral judgment and their systematic interconnection. Finally, I end with some remarks about what it means to think about the purity of rational principles that are inseparable from contentful moral cognition.

Kant's Scattered Descriptions of the Project of a Metaphysics of Morals

In both the *Critique of Pure Reason* and the *Metaphysics of Morals*, Kant suggests that a metaphysics of morals cannot completely dispense with concrete knowledge of human nature. There, his comments indicate that a metaphysics of morals may involve a priori principles, but cannot constitute a system of 'pure' knowledge involving *only* a priori cognition. By contrast, some of Kant's remarks in the *Groundwork* appear to paint the project of a metaphysics of morals in a different light: as an exercise of pure philosophy, cleansed entirely of everything empirical. Consequently, some readers have taken Kant to change his mind, by the time he wrote the

Groundwork, about what a metaphysics of morals should contain.[5] And if the project of a

metaphysics of morals was supposed to be cleansed entirely of everything empirical, then, perhaps,

the *Metaphysics of Morals* was not in fact the kind of work Kant envisaged in the *Groundwork*.[6] (Thus,

the *Metaphysics of Morals* has even been belittled as a 'senile's quirks'[7] and a relapse into pre-critical

thinking.[8])

In the introduction to the first *Critique*, Kant outlines his project of a transcendental

philosophy and explicitly excludes moral philosophy, because it cannot dispense with empirical

concepts.

> The chief target in the division of such a science [transcendental philosophy] is that absolutely no
> concept must enter into it that contains anything empirical, or that the *a priori* cognition be entirely
> pure. Hence, although the supreme principles of morality and the fundamental concepts of it are *a
> priori* cognitions, they still do not belong in transcendental philosophy, for, while they do not, to be
> sure, take the concepts of pleasure and displeasure of desires and inclinations, etc., which are all of
> empirical origin, as the ground of their precepts, they still must necessarily include them in the
> composition of the system of pure morality in the concept of duty, as the hindrance that must be
> overcome or the attraction that ought not to be made into a motive. (A14/B28-A15/B29, original
> emphasis)

In this passage, Kant points out that the very concept of duty already places moral philosophy

outside of the transcendental part of his critical system, because it concerns the necessity of an

action for finite beings, who perceive practical laws as imperatives that can go against their

inclinations. Hence, although cognition of the 'supreme principles of morality and the fundamental

concepts of it' are supposed to be a priori, such a priori cognition cannot be separated from our

---

[5] See e.g., Siep 2009.
[6] See e.g., Timmermann 2015.
[7] Paulsen 1899 p. 360
[8] See e.g., Ritter 1971 and Ilting 1983.

knowledge of human nature as finite, dependent beings. While one can supposedly find and analyze the a priori concepts of morality, Kant here suggests that one can only cognize them as components of knowledge that already involves a union of practical a priori principles and knowledge of human nature. The highest principle of morality (the Categorical Imperative) must be a priori, but our cognition of it is only possible by analyzing the idea of duty, and thus an idea that can only apply to finite beings like us.

Similarly, in the introduction to the *Metaphysics of Morals* published with the *Doctrine of Right* in 1797, Kant repeats his claim about the necessity of knowledge about human nature to a metaphysics of morals.

> If, therefore, a system of a priori cognition from concepts alone is called *metaphysics*, a practical philosophy, which has not nature but freedom of choice for its object, will presuppose and require a metaphysics of morals […] But just as there must be principles in a metaphysics of nature for applying those highest universal principles of a nature in general to objects of experience, a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles. But this will in no way detract from the purity of these principles or cast doubt on their a priori source. - This is to say, in effect, that a metaphysics of morals cannot be based upon anthropology but can still be applied to it. (6:216-217)

Here, Kant makes clear that a metaphysics of morals will often have to take human nature as its objects. Importantly, he points out that a metaphysics of morals will 'show' what is a priori not independently but *through* our knowledge of human nature ("in order to show *in it* what can be inferred from universal moral principles").[9]

---

[9] "[…] um *an ihr* die Folgerungen aus den allgemeinen moralischen Principien zu zeigen."

In the second volume of the *Metaphysics of Morals*, Kant repeats his claim from the introduction to the first *Critique*, about the connection between the concept of duty and our nature as finite, dependent beings.

> The very *concept of duty* is already the concept of a *necessitation* (constraint) of free choice through the law. [...] The moral *imperative* makes this constraint known through the categorical nature of its pronouncement (the unconditional ought). Such constraint, therefore, does not apply to rational beings as such (there could also be *holy* ones) but rather to *human beings*, rational *natural* beings, who are unholy enough that pleasure can induce them to break the moral law, even though they recognize its authority; [...]. (MS 6:379, original emphasis)

The concept of duty, through which we begin our investigation into the form of practical reason, shows itself through our first-personal experience of moral cognition. Thus, the very starting point of doing moral philosophy, from which we will eventually move on to a metaphysics of morals, is only initially accessible to us through the moral cognition of finite and dependent rational beings.

As these passages quoted above show, Kant's description of the project of a metaphysics of morals remains consistent throughout the *Critique of Pure Reason* and both volumes of the *Metaphysics of Morals*. However, his remarks about a metaphysics of morals in the *Groundwork* have been taken to depart from both his earlier and later descriptions. There, after a brief outline of the architectonic of his system, Kant writes:

> Here, however, I ask only whether the nature of science does not require that the empirical part always be carefully separated from the rational part, and that a metaphysics of nature be put before physics proper (empirical physics) and a metaphysics of morals before practical anthropology, with metaphysics *carefully cleansed of everything empirical so that we may know how much pure reason can accomplish* in both cases [...] (GMS 4:388, emphasis added)

Just two pages later, Kant adds that: "[...] the metaphysics of morals has to examine the idea and the principles of a possible pure will and not the actions and conditions of human volition generally,

137

which for the most part are drawn from psychology." (GMS 4:390) And Kant emphasizes the importance of grounding a metaphysics of morals on a priori principles rather than human nature a third time in section II of the *Groundwork*:

> [...] it is of the greatest practical importance [...] because moral laws are to hold for every rational being as such, to derive them from the universal concept of a rational being as such, and in this way to set forth completely the whole of morals, which needs anthropology for its *application* to human beings, at first independently of this as pure philosophy, that is, as metaphysics (as can well be done in this kind of quite separated cognitions); (GMS 4:412)

To many readers, these passages have appeared to create a tension between Kant's description of the project of a metaphysics of morals in the *Groundwork* and his descriptions of the same project in the early *Critique* and the late *Metaphysics of Morals*. According to Jens Timmermann, these passages show that "the Metaphysics of Morals published in 1797 is not the book that Kant promises his readers in the Preface of the Groundwork."[10] Similarly, Ludwig Siep has argued that Kant's differentiation between a metaphysics of morals and anthropology in the *Groundwork* rules out that the *Metaphysics of Morals* of 1797 is really the execution of the project envisaged in the *Groundwork*.[11] And according to Georg Anderson, Kant's numerous descriptions of a metaphysics of morals indicate that his understanding of this project are 'wavering and versatile', and that Kant even revised his entire concept of metaphysics.[12]

While these authors all agree on the alleged incompatibility of Kant's description of a metaphysics of morals in the *Groundwork* and in the *Metaphysics of Morals* of 1797, they take different views on Kant's remarks in the first *Critique*. Georg Anderson has argued that Kant's late conception

---

[10] Timmermann 2007a p. 171
[11] Siep 2009 p. 83
[12] Anderson 1923 p. 47-48

of a metaphysics of morals in 1797 is closer to the first *Critique* than to the *Groundwork*.[13] On this view, Kant's conception of metaphysics of morals in the *Groundwork* presents a momentary change of mind. Similarly, Ludwig Siep has interpreted Kant's separation of a metaphysics of morals from anthropology in the *Groundwork* as stricter than their separation in the first *Critique* and the *Metaphysics of Morals*.[14] By contrast, Jens Timmermann has interpreted Kant's remarks in the first *Critique* as suggesting simply that Kant just always felt uneasy about the place of moral philosophy in his critical system.[15]

This common attribution of a tension in Kant's description of a metaphysics of morals is unappealing in two ways. First, it is unfortunate to credit Kant with a major change of mind about the notion of a metaphysics of morals between the *Critique of Pure Reason* and the *Groundwork*, and back again between the *Groundwork* and the *Metaphysics of Morals*. While many authors have been content to disregard Kant's late *Metaphysics of Morals* as a 'senile's quirks,' ascribing an inconsistency to Kant that includes also the *Critique of Pure Reason* sounds even less appealing.

In addition, to read Kant's claims in the *Groundwork* about the 'purity' of a metaphysics of morals as incompatible with his description of moral philosophy in the first *Critique* and the *Metaphysics of Morals* casts doubt on Kant's project not merely in the late *Metaphysics of Morals*, but already on his 'metaphysics of morals' in section II of the *Groundwork*. As both Jens Timmermann and Ludwig Siep have correctly pointed out, Kant's discussion of the common distinction of duties into perfect vs imperfect, to self and to others in section II of the *Groundwork* is difficult to square

---

[13] Anderson 1923 p. 47
[14] Siep 2009 p. 78
[15] Timmermann 2007a p. 170

with an alleged attempt to cleanse moral philosophy from everything empirical. As Siep has aptly put it:

> This distinction – or more precisely, this grounding of a traditional distinction – seems to be the only example of purely metaphysically grounded systematic moral philosophy on Kant's part. But we may well harbour doubts even here. For it follows neither from the difference between thinking and willing, nor from that between laws of thought and laws of action, that specific principles of action can be thought as law without contradiction (as a 'universal law of nature'), but not willed as such. This difference only emerges in relation to specific natural conditions of willing.[16]

Similarly, Jens Timmermann has claimed that Kant's discussion of concrete duties *as duties* in section II of the *Groundwork* is incompatible with the claim of doing 'pure' moral philosophy.

> A pure will is located in the world of understanding, in which there are no duties because the moral law applies descriptively. Only when we consider ourselves as members of both the world of understanding and the world of sensibility do we experience morality as an unconditional ought (IV 454.9–15). Kant presumably thought that the metaphysics of morals would not regard the kinds of obligation to be classified duties, in the thick sense of the word, with all the moral psychological details they entail – that must be left to 'anthropology' – but rather as the commands of an autonomous human will. […] Yet a certain amount of 'impurity' would inevitably affect such a project. It is – comparatively – easy to see how a pure will could contain the moral law in itself; it is rather more difficult to see how it might contain particular token duties, such as the duty not to lie or to care for one's friends, which draw on the specifics of human nature;

Thus, there are two difficulties for the reading, shared by authors like Anderson, Siep and Timmermann, of Kant's description of a metaphysics of morals in the *Groundwork*. For one, this reading makes it difficult to square the *Groundwork* with both the *Critique of Pure Reason* as well as the *Metaphysics of Morals* of 1797. For another, this reading also casts doubt on the feasibility of Kant's discussion of the *Groundwork* itself.[17] Fortunately, it is not inevitable to arrive at this particular

---

[16] Siep 2009 p. 85

[17] Compare also Ludwig 1991, who calls Kant's notion of the good will in the *Groundwork* unfortunately "empirisch und psychologisch stark vorbelastet." (Ludwig 1991 p. 69)

reading of Kant's remarks about the purity of moral philosophy in the *Groundwork*. Another

appealing reading is available. In the remainder of this chapter, I argue that Kant asks us to carve up

the conceptual space in a way that makes room for a conception of purity of rational principles as

grounding, without thereby presupposing their separability from content.

How Could Kant's Project be Consistent?

If Kant's project of a metaphysics of morals were actually inconsistent between the passages

surveyed so far, then the same inconsistency would arguably also exist within both the *Critique of Pure*

Reason and the *Metaphysics of Morals*. In his section on 'the architectonic of pure reason,' Kant outlines

his envisaged division between a metaphysics of nature and a metaphysics of morals:

> Metaphysics is divided into the metaphysics of the **speculative** and the **practical** use of pure reason,
> and is therefore either **metaphysics of nature** or **metaphysics of morals**. The former contains all
> rational principles form mere concepts (hence with the exclusion of mathematics) for the **theoretical**
> cognition of all things; the latter, the principles which determine **action and omission** *a priori* and
> make them necessary. Now morality is the only lawfulness of actions which can be derived entirely *a*
> *priori* from principles. Hence the metaphysics of morals is really the pure morality, which is not
> grounded on any anthropology (no empirical condition). (A840/B868-A842/B870, original
> emphasis)

Thus, while Kant's introduction to the first *Critique* pointed out cognition of the 'supreme principles

of morality and the fundamental concepts of it' cannot be separated from our knowledge of human

nature as finite, dependent beings, he continues to call a metaphysics of morals the "pure morality,

which is not grounded on any anthropology." Could this side-by-side of such seemingly different

remarks in the same book be a result of a change of mind, too? Perhaps, Kant changed his mind

about the possibility of a 'pure' metaphysics of morals just after he had written the *Groundwork*, and

incorporated this change of mind in his revised introduction to the first *Critique* of 1787 and the later *Metaphysics of Morals*?[18] I believe it cannot. For the same side-by-side of such seemingly different remarks can also be found in the introduction to the *Metaphysics of Morals.* There, Kant puts a description of the project of the metaphysics of morals as 'pure' and its inseparability from our knowledge of human nature as finite and dependent side by side. In the passage already quoted above at 6:216-217, Kant claims that "we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles. *But this will in no way detract from the purity of these principles* or cast doubt on their a priori source […]" (6:216-7, emphasis added) Already in this passage, we have an emphasis on both the connection between a metaphysics of morals and anthropological knowledge as well as its purity. And immediately afterwards, Kant adds:

> The counterpart of a metaphysics of morals, the other member of the division of practical philosophy as a whole, would be moral anthropology, which, however, would deal only with the subjective conditions in human nature that hinder people or help them in fulfilling the laws of a metaphysics of morals. […] It cannot be dispensed with, but it must not precede a metaphysics of morals or be mixed with it; for one would then run the risk of bringing forth false or at least indulgent moral laws, which would misrepresent as unattainable what has only not been attained just because the law has not been seen and presented in its purity […] (6:216-217)

More likely than any change of mind – whether before or after writing the *Groundwork* – Kant continuously held both following beliefs: (i) that cognition of the highest principle of morality, and the fundamental concepts connected with it, cannot be separated from our knowledge of humans as rational yet finite beings; and (ii) that a metaphysics of morals would not be grounded on any empirical knowledge of human nature and would therefore be 'pure.'

---

[18] Ludwig Siep suggests as much in his discussion of Kant's project of a metaphysics of morals, Siep 2009 p. 78

Kant emphasizes this second requirement – that a metaphysics of morals be free of empirical knowledge – in the context of its ground or foundation. For instance, in 'the architectonic of pure reason,' Kant claims that a metaphysics of morals is "not *grounded* on any anthropology (no empirical condition)." (A840/B868-A842/B870) In the *Groundwork*, Kant repeats his claim that "all moral philosophy is *based* entirely on its pure part; and when it is applied to the human being it does not borrow the least thing from acquaintance with him (from anthropology." (GMS 4:389) Kant emphasizes this 'purity' requirement for a metaphysics of morals again in section II, claiming that "it is of the greatest practical importance not to make its [the metaphysics of morals's] principles *dependent* upon the special nature of human reason" (GMS 4:412, emphasis added). This way of talking about the a priori basis of a metaphysics of morals corresponds well with Kant's introduction to the first volume of the *Metaphysics of Morals*: "They [moral laws] hold as laws only insofar as they can be seen to have an a priori *basis* and to be necessary." (MS 6:214-215) Similarly, a little later Kant holds that a "moral anthropology […] cannot be dispensed with, but it *must not precede* a metaphysics of morals or be mixed with it." (MS 6:216-217)

Thus, Kant speaks about the 'purity' requirement for a metaphysics of morals in the context of discussing the ground or basis for moral cognition. By contrast, when he talks about the inseparability of moral cognition from knowledge of human nature, he seems concerned with the overall content of a metaphysics of morals. A discussion of the a priori foundation of moral knowledge will necessarily be accompanied by content relating to human life. Thus, in the introduction to the *Critique of Pure Reason*, Kant claims that "although the supreme principles of morality and the fundamental concepts of it […] do not, to be sure, take the concepts of pleasure and displeasure of desires and inclinations, etc., which are all of empirical origin, as the ground of their precepts, *they still must necessarily include them in the composition of the system of pure morality* in the

143

concept of duty […]" (A14/B28-A15/B29, emphasis added) And similarly, in the introduction to the *Metaphysics of Morals*, Kant says that "we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles. But this will in no way detract from the purity of these principles or cast doubt on their a priori source." (MS 6:216-217)

How, then, could Kant be committed to both (i) that cognition of the highest principle of morality, and the fundamental concepts connected with it, cannot be separated from our knowledge of humans as rational yet finite beings; and (ii) that a metaphysics of morals would be an exercise in 'pure' moral philosophy, not grounded on any empirical knowledge of human nature? The possibility of an answer will depend on how we interpret Kant's claim that a metaphysics of morals must be *grounded* a priori.

For Kant, practical knowledge is grounded in a priori principles. That is, the representation of actions and ends as intrinsically good is only so much as possible because reason provides a priori principles through which we can represent specific objects as such moral actions and ends. Kant also believed that it followed from this view about practical knowledge that an analysis of the principles that make moral judgments possible (a metaphysics of morals) can therefore not be derived from anthropology. For one, empirical study alone could not observe such a thing as unconditionally applicable norms, for its object would always remain empirically conditioned. But equally important: without a preceding metaphysics of morals, an anthropological study of moral judgment could not even have the alleged object of such an that investigation (moral judgment) properly in view. Unless we already knew about the a priori principles that make moral judgment possible independently of empirical observation, we would not be in a position to identify any object out there as a potential

instance of moral judgment. A metaphysics of morals, i.e., an analysis of the principles that make moral judgment possible, must then be grounded a priori.

Kant's project of metaphysics of morals in section II of the *Groundwork* is an analysis of the rational principles that underlie our common moral judgments and their systematic interconnection. (Such a preliminary metaphysics of morals as is provided in the *Groundwork* does, of course, not yet suppose that these principles are really applicable – to make this judgement we would need a 'critique of practical reason,' which Kant attempts to provide, in rough outlines, in the third section of the *Groundwork* and the second *Critique.*) And on this reading, Kant's project of a metaphysics of morals is compatible with both (i) and (ii). As an analysis of the rational principles that underlie our moral cognition and their systematic interconnection, a metaphysics of morals would be an investigation of an aspect of our cognition that cannot simply be distilled from empirical observations, e.g., by generalizing from our experience. In this way, a metaphysics of morals would not be *grounded* in an anthropology – no anthropology can precede it. Both the *Groundwork* and the *Metaphysics of Morals* are centrally concerned with concepts that, according to Kant, are not derived from empirical observation but are conditions of the possibility of practical reason: the Categorical Imperative and freedom. In this vein, the *Groundwork* analyzes the highest principle of practical reason as underlying all our commonly known duties, and the principle's possible formulations and their relations; the *Metaphysics of Morals* begins with a discussion of the a priori concepts that are at the foundation of a metaphysics of morals, like freedom and the Categorical Imperative, and continues to show how a body of common moral knowledge can be unified by recourse to these a priori concepts. In this way, neither the *Groundwork* nor the *Metaphysics of Morals* can be said to be 'grounded' in an anthropology, as if an anthropology would precede them. At the same time, the project of a metaphysics of morals as an analysis of the rational principles that underlie our moral

145

cognition and their systematic interconnection will also contain reference to our nature as finite, dependent beings. For the very starting point of doing moral philosophy, from which we will eventually move on to considering what could make moral judgment possible, is only initially accessible to us through the moral cognition of finite and dependent rational beings. And in order to show how a priori principles underlie concrete ethical obligations, such an analysis will, ultimately, also have to discuss these concrete ethical obligations of human beings.

Thus, the *Groundwork* relies on a set of commonly accepted moral duties in order to demonstrate that the representation of each of them indeed contains the form of a universal law; and the *Metaphysics of Morals* considers our concrete duties as finite, dependent beings even more systematically, connecting them into a coherent body of moral knowledge. In the following section, I outline how this reading allows us to understand the progression of Kant's argument in section II of the *Groundwork*.

A Proposal for Reading the Groundwork's 'Transition to a Metaphysics of Morals'

Kant's first section of the *Groundwork* is entitled 'Transition from common to philosophical moral rational cognition.' By starting the *Groundwork* with reference to 'common moral cognition' (*gemeine sittliche Vernunfterkenntnis*) Kant did not mean to simply start the *Groundwork* with the prevailing mores of the day. Although Kant was indeed confident about the adequacy of people's moral judgments in general, the primary concern here is not the average moral opinions of Kant's contemporaries. Kant's starting point in the *Groundwork* is the intuitive, first-personal awareness of morality in our everyday practical judgments: that sometimes it is right to do things simply because they are the right things to do. *Philosophical* moral cognition would be to find, and give expression to,

146

the principle (or principles) making this common moral cognition possible. According to Kant, our common moral cognition already tells us that the good will is the only thing of unlimited moral worth. (4:393-396) Thus, the start of the *Groundwork's* project is to unravel the concept of a good will, leading Kant to three 'propositions': (i) that the concept of the good will is the concept of acting from duty (4:397-399); (ii) that the moral worth of acting from duty lies in the maxim of an action (4:399-400); and (iii) that duty is the necessity of an action from respect for the moral law (4:400-401). From these three propositions that attempt to unravel the concept of a good will, Kant then concludes with a preliminary formulation of the underlying principle of common moral cognition: the universal law formulation of the Categorical Imperative (4:402-403).

Once we have an initial idea of what the underlying, rational principle of moral judgment might be, we can then proceed to analyze this principle further. At the beginning of section II of the *Groundwork*, Kant is adamant to point out that taking the concept of duty from the 'common use of practical reason' does not mean that we therefore derive the concept from experience. In fact, Kant claims that no certain examples of an action from duty could be found, even in our own self-examination (4:406-407). Importantly, Kant claims that one could not derive morality from examples, for in order to count as an example it must be appraised by principles of morality which determine whether it should count as an example.[19] Consequently, there is an important distinction for Kant between deriving a concept from experience by generalization on the one hand, and analyzing cognition in order to find its rational (a priori) condition of possibility on the other hand. Although the *Groundwork* starts with examples of our common moral cognition, Kant emphasizes that his analysis of the concept of duty is not taken from experience; rather, Kant asks what kind of

---

[19] Kant had already made this point in the first *Critique*, while commenting on Plato's insight into the conative nature of reason (A314-A315/B371).

rational principle *could* make the kind of judgment possible that we find in our own, common moral awareness.

Thus, we have already progressed from common to philosophical moral cognition. However, Kant believes that the philosophical discussion so far has remained 'popular philosophy,' i.e., too much constrained by examples. Therefore, we need to analyze this ability for moral judgment more closely in order to arrive at a true 'metaphysics of morals' (4:412). This is the project of section II of the *Groundwork*, and it will be the 'transition from popular moral philosophy to metaphysics of morals.'

Kant begins his further analysis of practical reason (i.e., the will) by looking closer at the notion of an imperative, which is the way in which finite beings cognize practical laws (4:413-414). At this point, Kant famously distinguishes between different kinds of imperative – hypothetical imperatives of skill[20], hypothetical imperatives of welfare/prudence[21], and categorical imperatives[22] (4:414-416). The necessity represented by a hypothetical imperative is not so difficult to recognize: it depends on a previously given, and sensibly determined, desire for some end. But a categorical imperative does not depend on any prior desire and is supposed to command unconditionally. Therefore, Kant asks how such an imperative would be possible, if it cannot rely on an antecedently given desire. Kant's famous answer is that a categorical imperative must 'furnish' its own formula (4:417-420). Kant then asks what this formula could be, and comes to the conclusion that an a priori principle underlying universal, practical demands must have the form of universality. Therefore, one way of formulating the principle is by capturing this form, i.e., the form of universality. In this way,

---

[20] Or 'technical imperatives'
[21] Or 'pragmatic imperatives'
[22] Or 'moral imperatives'

Kant arrives at two formulations of the Categorical Imperative in the form of a universal law, and of a universal law of nature.

If this Categorical Imperative is supposed to underlie all our representations of moral duties, then Kant should be able to demonstrate this by means of the formula of universality. Among the commonly recognized duties of his time are duties to others vs. duties to the self, and perfect vs. imperfect duties – particularly regarding duties to the self, the thought that our representation of such duties takes the form of universality is not exactly intuitive. Therefore, Kant moves through four examples of all types of duties according to their 'common division' and attempts to show how the form of universality is really presupposed in our representation of each of these duties. (4:422-423) In this way, Kant's infamous 'derivation of imperatives of duty' in section II demonstrates the unity of commonly recognized moral imperatives by making explicit the form of practical reason in the moral knowledge of ordinary agents.

Having demonstrated this shared a priori basis of all moral judgments (or, at least, having made his analysis of the will more plausible through four examples), Kant then exalts the a priori nature of his investigation.

> For the purpose of achieving this it is of the utmost importance to take warning that we must not let ourselves think of wanting to derive the reality of this principle from the special property of human nature. For, duty is to be practical unconditional necessity of action and it must therefore hold for all rational beings (to which alone an imperative can apply at all) and only because of this be also a law for all human wills. (4:425)

> The question is therefore this: is it a necessary law for all rational beings always to appraise their actions in accordance with such maxims as they themselves could will to serve as universal laws? If there is such a law, then it must already be connected (completely a priori) with the concept of the will of a rational being as such. But in order to discover this connection we must, however reluctantly, step forth, namely into metaphysics […] namely a metaphysics of morals. (4:426)

149

Kant then proceeds to 'step further into a metaphysics of morals', and analyzes the above-mentioned idea of a connection between the universality of moral representations and its (a priori) connection to the idea of a will of a rational being. There, Kant argues that if this form of practical cognition is a priori, and thus belongs to every practically rational being, then it also indicates the capacity of practically rational beings to represent ends through practical, universal laws. This leads Kant to his second formulation of the Categorical Imperative, the Formula of Humanity. Here again, Kant then moves through the four examples of duties he already discussed before, in order to show how also the status of practically rational nature is implicit in our representation of moral action and ends. (4:427-430) And since the 'self-legislation' of practical reason must not be dependent on sensible desires in order to command unconditionally, Kant argues that practical reason must be free, i.e., autonomous. This consideration leads Kant to eventually to another formulation of the Categorical Imperative, the formulation of the kingdom of ends, as an ideal of self-legislating members of an intellectual realm (4:432-434).[23]

At this point, the discussion of section II of the *Groundwork* has come to its promised goal: we are now familiar with the underlying principle of the good will not just by popular philosophical insight (through examples), but have found the more specific formulas of this one principle and their relation through philosophical analysis. Importantly, our analysis of the formulations of the highest principle of practical reason and their relation was not, Kant would insist, derived from empirical investigation. The content of the different formulations of the Categorical Imperative, as well as their connections, was not taken from experience by generalization but from the concept of

---

[23] In this context, Kant also provides his most eloquent formulations of human dignity (4:434-435).

duty and the idea of a will of a rational being as that which makes moral judgment so much as possible. Thus, we have reached the limit of a 'metaphysics of morals:'

> So far, we've only analyzed the concept of morality, and from the concept of the good will have derived the notion of autonomy, vs heteronomy, and out of the supreme principle of autonomous legislation have derived three formulas. But the existence, or possibility even, of such an a priori practical synthetic proposition is not within the bounds of metaphysics of morals. In order to show the possibility and necessity of such a proposition we first need a full critique of practical reason, and the next section will outline its main features. (4:444-445)

Kant's transition from popular moral philosophy to a metaphysics of morals in section II of the *Groundwork* is an analysis of the rational principles that underlie our moral cognition and their systematic interconnection. Neither the rational principles underlying our moral cognition themselves nor their connections are justified on empirical grounds: our entire investigation into the a priori principles of practical reason and their connection proceeded without grounding any of our insights on empirical research. However, Kant's transition to a metaphysics of morals in the *Groundwork* does not thereby banish all knowledge of human nature: indeed, Kant's demonstrations of how the Categorical Imperative underlies all (commonly recognized) types of imperatives relies on examples of concrete moral duties, and thus of course involves more than merely a priori practical cognition. In this way, the conclusions Kant wants to derive in his argument of section II – and, specifically, the connections between the different formulas of the Categorical Imperative – are not themselves derived from empirical knowledge of human nature. Rather, they are the rational principles that would make our human experience of moral obligation possible. And demonstrating this cannot be done completely in the abstract – quite to the contrary, it requires showing how these rational principles already feature in the actual, contentful moral cognition of finite and dependent rational beings.

151

In light of this reading of Kant's project of a metaphysics of morals, we now also know what the aim of the *Metaphysics of Morals* will be: like section II of the *Groundwork* but on a much larger scale, Kant's *Metaphysics of Morals* will be an attempt to analyze the rational principles that underlie our moral cognition and their systematic interconnection. Specifically, the *Metaphysics of Morals* will try to make explicit the rational basis of our systematic moral knowledge, starting with specific duties and trying to highlight the a priori structure that underlies our knowledge of these duties.

At the same time, the consistency of Kant's project which is available through this reading allows for considerable variation in how 'thick' or 'thin' the knowledge of human nature may be that accompanies the respective accounts of a metaphysics of morals. Both the *Groundwork* and the *Metaphysics of Morals* presuppose a minimal knowledge of human nature, namely the kind of knowledge that necessarily accompanies the conception of an imperative and which implies our finitude and dependence. It is this relatively thin conception of the finite, rational being that predominates Kant's discussion of the *Groundwork*. The same finite, rational being is also presupposed in the *Metaphysics of Morals*, but in making explicit the rational basis of our systematic moral knowledge, the later book now engages an even thicker knowledge of human nature. While it is possible to differentiate how thick our knowledge of human nature is in Kant's respective accounts of a metaphysics of morals, they each embody the same union of a priori principles and knowledge of our humanity as finite and dependent.

How Formal can Hylomorphic Practical Philosophy Be?

It is one thing to make conceptual space for the idea of the 'purity' of the form of practical reason, where this form can only be shown by pointing to its being contained in our already existing

common moral cognition; but it is quite another to demonstrate that this pure form of practical reason itself can actually be the content of any determinate thought – and, specifically, of our philosophizing in a metaphysics of morals. After all, doesn't Kant's moral philosophy suppose just that? In other words, doesn't Kant's moral philosophy pretend to be both hylomorphic and strictly formal? The potentially unsatisfactory answer to this question is, I believe: yes and no. The answer is no, if we expect that this pure form of practical reason should be apprehensible *as it is* independently of any content. An analogy to make this thought more intuitive is the following: we can make drawings that represent two crossing vectors on a plain on a blackboard in order to demarcate a specific point on a plain. But that drawing cannot possibly be more than a makeshift help in order to represent that point. Because as soon as a point is represented on any drawing, it necessarily becomes a (small, perhaps even miniscule) area, and an area is not a point. Similarly, any attempt to represent the form of practical reason will be either showing that form through a determinate content, or it will be an abstract description of the form, but not the form itself.[24] The answer is yes, if we expect the form of practical reason to be given an abstract description and to be demonstrated through the content in which it is contained. Although one cannot represent e.g., 'universality' any more concretely than one can represent a point on a plain on a blackboard, one can give both an abstract description. (Such as: 'Only act on that maxim that you can will to become a universal law…' and '$(a_x, a_y) \times (b_x, b_y)$'.) And just like one can provide a demonstration of the point through the drawing on the blackboard, one can also provide a demonstration of the ethical content in which this 'universality' is contained. One obvious place to look for this in Kant's *Groundwork* is the four

---

[24] We would thus expect that formulating the form of practical reason is both difficult and controversial. Indeed, Kant's struggle to formulate different 'formulas' of the Categorical Imperative demonstrate just this difficulty. If it was easy to represent the form of practical reason, Kant would hardly have felt the need to provide us with three (or five, depending on how we count them) different formulas thereof.

examples of Kant's infamous derivation of imperatives of duty. Take for instance Kant's discussion of the duty to develop one's talents. There, Kant says:

> A third finds in himself a talent that by means of some cultivation could make him a human being useful for all sorts of purposes. However, he finds himself in comfortable circumstances and prefers to give himself up to pleasure than to trouble himself with enlarging and improving his fortunate natural predispositions. But he still asks himself whether his maxim of neglecting his natural gifts, besides being consistent with his propensity to amusement, is also consistent with what one calls duty. He now sees that a nature could indeed always subsist with such a universal law […] only he cannot possibly will that this become a universal law or be put in us as such by means of natural instinct. For, as a rational being he necessarily wills that all the capacities in him be developed, since they serve him and are given to him for all sorts of possible purposes. (GMS 4:423)

Kant here takes as his starting point a person who already knows, or presumes, plenty of things about human nature and society. And in following the fictional agent's thoughts, Kant also believes that we find something that cannot be derived from our knowledge of human nature or society: the universality of the form of practical reason. In other words, by thinking through this situation with our fictional agent, we are supposed to see that they cannot represent their actions as morally worthy ('consistent with what one calls duty') unless they represent them through the form of universality. But according to Kant, this form is not itself based on any anthropological knowledge. Thus, Kant's brief discussions of duties in section II of the *Groundwork* provide some examples of what it is, for Kant, to think about something 'pure' (i.e., not derived from anthropology) that also cannot be thought independently of determinate content. And of course, Kant's four examples of imperatives of duty in the *Groundwork* are not the only place to look for an example: Kant's most extensive treatment of this topic is provided by the *Metaphysics of Morals*. In the following chapter, I will discuss how Kant's last work on moral philosophy similarly attempts to analyze the rational principles that underlie our moral judgments and their systematic interconnection.

Ultimately, we might find Kant's project of a metaphysics of morals wrongheaded. Perhaps, to some readers the idea of a moral philosophy that is both grounded in a priori principles and inseparably connected to cognition of our nature as finite, dependent beings is just implausible or unconvincing. And perhaps, some readers may just find Kant's examples of what it would be to think about such a form that is inseparable from content unimpressive. But irrespective of how successful we might judge Kant's hylomorphism in his project of a metaphysics of morals, we should acknowledge that both of Kant's commitments are, at least in theory, compatible: cognition of the highest principle of morality, and the fundamental concepts connected with it, cannot be separated from our knowledge of humans as rational yet finite beings; at the same time, a metaphysics of morals would be an exercise in 'pure' moral philosophy, not grounded on any empirical knowledge of human nature. Consequently, we also need not assume that Kant had a sudden change of mind about what a metaphysics of morals should contain, or that many of his major works are full with contradictory thoughts. If Kant's project of a metaphysics of morals still seems unconvincing, it can at least do so consistently.

Chapter 6

A System of Moral Knowledge: Kant's *Metaphysics of Morals*

A Book with Few Fans

To many readers, the content of Kant's *Metaphysics of Morals* has seemed to be in tension with Kant's earlier works on moral philosophy. For instance, one might have expected Kant to develop a doctrine of virtue by picking up where he left off in the *Critique of Practical Reason*, namely with the highest good as the union of virtue and happiness. Or one might have expected Kant to determine the laws that are necessary for obtaining this highest good by deriving them from one or some of the formulas of the Categorical Imperative as outlined in the *Groundwork*. But Kant does neither of these things. Instead, Kant first provides a concise restatement of central terms of his practical philosophy, and then plunges into a discussion of right, which he gives its own underlying principle (the 'universal principle of right'). And in the second part of his *Metaphysics of Morals*, Kant introduces the idea of an objective end, from which he then builds his discussion of concrete ethical duties. This seeming tension has left numerous readers with the impression that we must either background the *Metaphysics of Morals* in favor of Kant's earlier work, or vice versa.

Because of this impression, some authors have downplayed the importance of the *Metaphysics of Morals*. Schopenhauer famously derided the work as senile nonsense.[1] Hannah Arendt's judgment about the first part of the *Metaphysics of Morals* was equally derogatory. Finding it "difficult not to agree with Schopenhauer,"[2] she described the *Doctrine of Right* as "rather boring and pedantic." In her

---

[1] Schopenhauer 2010 pp. 554-560.
[2] Arendt 1992 pp- 7-8

lectures on Kant's political philosophy, Arendt made clear that she did not find Kant's late works to

be of any interest, for "the decrease of his mental faculties, which finally led into senile imbecility, is

a matter of fact."[3] Similarly, Friedrich Paulsen has called Kant's philosophy of right deplorable

'Schrullen des Greises' (senile's quirks), and generously absolved the fragile old man from

responsibility for his allegedly cranky ideas.[4] Some readers have even found the *Metaphysics of Morals*

to be a relapse into precritical dogmatism. Claims like this have been made about both the *Doctrine of*

Right[5] as well as the *Doctrine of Virtue*.[6] And even a sympathetic reader like Manfred Kuehn has

recently claimed that "[t]he *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."[7]

Unfortunately, this rather derogatory interpretation of the *Metaphysics of Morals* has persisted for

almost two centuries. After an initial row of critical reviews of the first two editions of 1797 and

1803, literature dealing specifically with the *Metaphysics of Morals* ebbs away. Only with a renewed

interest in political philosophy in the second half of the 20th century have philosophers again

devoted more attention to the book, but interpretive efforts have so far focused mostly on the first

half, Kant's *Doctrine of Right*.[8]

---

[3] Arendt 1992 p. 9. As a judgment about the *Metaphysics of Morals*, this is unfair. It appears that Kant was indeed weakened by the time he published the *Metaphysics of Morals*, such that he could neither polish his manuscript as much as other works nor could he oversee the printing process, leading to some corruptions in the text. (For a more detailed discussion see e.g., Kuehn 2001a p. 396 and Ludwig 2015b.) Kant also stopped teaching in 1797 due to his worsening health. Note however that the deterioration of Kant's mental strength is usually said to have begun noticeably ca. 1799, 4-5 years before his death. (Compare Kuehn 2001a pp. 413-422.).

[4] Paulsen 1899 p. 360

[5] See for instance Ritter 1971. Cf. Busch 1979.

[6] See Ilting 1983.

[7] Kuehn 2010 p. 10. For the view that Kant's *Groundwork* and the *Metaphysics of Morals* cannot be reconciled as belonging to the same philosophical system, see also Ralf Ludwig 1991.

[8] As Bernd Ludwig has pointed out, even the renewed interest in Kantian political philosophy after Rawls initially focused on the *Groundwork* (Ludwig 2015b p. 1561). For a concise overview of the history of the reception of the *Doctrine of Virtue*, see Rühl 2015 p. 1567.

On the other hand, some readers have suggested that the *Metaphysics of Morals* has the potential to revolutionize our understanding of Kant's ethics. Rather than rejecting the *Metaphysics of Morals* as a senile relapse into precritical thinking, or a badly edited selection of lecture notes, we might have to revise our reading of Kant's earlier work, and in particular of his alleged test for right action in the *Groundwork*.[9] Indeed, two of the most senior Kant scholars of the last decades have taken precisely this view. For instance, in the introduction to her translation of the *Metaphysics of Morals*, Mary Gregor hinted that "the neglect of the *Metaphysik der Sitten* has left us with a distorted view of Kant's moral philosophy as a whole."[10] Similarly, Allen Wood has argued that Kant's late work on moral philosophy can serve as a corrective to popular, but hasty, interpretations of the *Groundwork*.[11]

Despite the opposing nature of these two ways of reading Kant's *Metaphysics of Morals*, most readers still agree that there is a genuine tension between the presumed methodology of Kant's earlier works on ethics and the *Metaphysics of Morals*. The reason for this agreement is that they appear to share what I will call a 'standard assumption' about the *Metaphysics of Morals*. Kant describes his work as the application of the highest principle of morality (the Categorical Imperative) to the entire system of duties; and according to the standard assumption, this 'application' is to derive the concrete duties for human beings from the method of ethics allegedly provided in the *Groundwork*. On this view, the *Groundwork* provided the initial test or decision-procedure for ethics, and the *Metaphysics of Morals* is supposed to derive a complete system of duties from this test. But

---

[9] A somewhat middle position has been defended by Manfred Kuehn, who argues that the *Metaphysics of Morals* is precritical in its content, but critical in how it fits the substantive doctrine into the critical system. (Kuehn 2010 p. 9)
[10] Gregor 1963 p. xi
[11] Wood 1997 p. 4

since the actual content of the *Metaphysics of Morals* does not fit this description, there appears to many readers to be a tension between the Kant's earlier and later works on moral philosophy.

In this chapter, I want to suggest that an alternative view is available which does not force us to posit such a stark break in Kant's moral philosophy. As I have previously argued, Kant's 'transition to a metaphysics of morals' in the *Groundwork* is concerned with showing that the Categorical Imperative, as the form of practical reason, underlies all commonly recognized types of imperatives.[12] In this way, we can understand Kant's project of a 'metaphysics of morals' in the *Groundwork* as an attempt to analyze the rational principles that underlie our moral judgment and their systematic interconnection. On the reading I suggest here, Kant's *Metaphysics of Morals* provides the long-anticipated elaboration of this very project: it takes a body of moral knowledge as its starting point, and shows how a priori principles unify this moral knowledge into a coherent system. Rather than systematically deriving duties from the highest principle of morality as from a test, Kant's *Metaphysics of Morals* takes up and moves through the content of commonly read treaties on ethics from Kant's time, and attempts to make explicit the rational basis of political institutions and ethical virtues. Consequently, we also need not emphasize one work and downplay the other; on the reading of Kant's *Metaphysics of Morals* I propose here, there just is no big tension in Kant's work on moral philosophy throughout his critical period.

In the first two sections, I briefly outline the protracted origins of the *Metaphysics of Morals* and the requirements that a successful interpretation must meet. In section three, I present a series of difficulties for the standard reading, and argue that we should reject it. Finally, in section four, I

---

[12] Ramsauer 2024 and chapter 4.

suggest an alternative reading of the *Metaphysics of Morals* which allows us to see Kant's moral philosophy as coherent throughout his critical period.

The Origins of the Metaphysics of Morals

When Kant's *Metaphysics of Morals* was finally published in 1797, he had already planned to publish a metaphysics of morals for at least three decades. In 1768, during his pre-critical period, Kant wrote to Herder that he was now investigating the nature (Bestimmung) and limits of human capacities and inclinations, and that he thought he had succeeded in this matter regarding morals. Thus, he wrote that he was now working on a 'Metaphysics of Morals,' and that this work would contain the "evident and fruitful principles and the method" of "this type of cognition." (Br 10:74, translation mine) In his letter to Herder, Kant also mentioned that he hoped to finish this work the same year. During the autumn of 1770, at the beginning of his 'silent decade,' Kant wrote to the mathematician and philosopher Johann Heinrich Lambert that he would spend the winter writing a 'pure moral philosophy,' which would be a 'metaphysics of morals' (Br 10:97). In 1773, Kant wrote to the physician and philosopher Marcus Herz that he was looking forward to finishing his critique of pure reason, which would allow him to finally "turn to metaphysics." This metaphysics, Kant added, "has only two parts, the metaphysics of nature and the metaphysics of morals," of which he hoped to publish the metaphysics of morals first (Br 10:145). Almost ten year later, in the first edition of his *Critique of Pure Reason* of 1781, Kant still had to content himself with envisaging a metaphysics of morals. There, he explains that "[m]etaphysics is divided into the metaphysics of the *speculative* and the *practical* use of pure reason, and is therefore either metaphysics of nature or

160

metaphysics of morals." (KrV A841/B869, original emphasis)[13] And in 1792, five years before its publication, Kant mentions in a letter to Johann Erhard the "Metaphysics of Morals on which I am at work" (Br 11:399).[14] Contrary to his stated intentions, Kant did not proceed from his first *Critique* directly to a metaphysics of morals. Instead, Kant went on to write the *Groundwork* for a future metaphysics of morals and eventually the *Critique of Practical Reason* (which would appear almost a decade before the eventual publication of his *Metaphysics of Morals* in 1797). And while Kant also appeared confident to publish a metaphysics of morals before a metaphysics of nature, he even published the *Metaphysical Foundations of Natural Science* eleven years before his *Metaphysics of Morals*.

As these protracted origins of the *Metaphysics of Morals* show, for at least three decades of his life Kant consistently believed that writing a metaphysics of morals would be a worthwhile project. To some readers of Kant, this might initially come as a surprise. On a popular Neo-Kantian thought, morality consists essentially of a practical-logical criterion for testing maxims (i.e., the Categorical Imperative) and the systematic derivation of duties applicable to human being from this 'method of ethics.' As Mary Gregor describes it in the introduction to her translation of the *Metaphysics of Morals*, "[the *Groundwork's*] purpose is to lay the foundation for the projected *Metaphysik der Sitten*, to investigate the supreme principle of morality, from which, in the subsequent work, 'the whole system' of duties will be derived."[15] But that was the correct way to characterize the purpose of the *Metaphysics of Morals*, why would it be such a worthwhile project? In fact, on such a view, it would seem difficult to see why we would need a metaphysics of morals at all. All such a book might do is go through a range of (necessarily non-exhaustive) aspects of human life and derive concrete duties

---

[13] Kant further envisages a metaphysics of morals at KrV A850/B878 and Bxliii.
[14] For a longer and detailed overview of the history of Kant's project of a metaphysics of morals, see especially Kuehn 2010 pp. 9-16.
[15] Gregor 1963 p. xi. See also Gregor's translator preface 1991 p. 1.

for them – perhaps with the intent to prevent people from misapplying a decision-procedure? But given that no book could reasonably foresee all possible individual situations, it could hardly provide more than a blueprint for how to make such applications in concrete situations, and that is precisely what (at least on this standard assumption) the *Groundwork* already gave us. Indeed, as Mary Gregor herself has pointed out, "as the great majority of his students and critics are concerned, he might as well not have written the *Metaphysik der Sitten* at all."[16]

The standard assumption about the *Metaphysics of Morals* gains much of its plausibility instead from Kant's insistence that his metaphysics must attain the status of a Wissenschaft, i.e., a systematic body of knowledge (Wissen) on the basis of a priori principles.[17] If the *Metaphysics of Morals* really was the systematic application of a test or decision-procedure to situations of human life, then we would have an easy way of explaining why the *Metaphysics of Morals* can claim to present a system of moral knowledge. But while the standard reading has an easy explanation of how a metaphysics of morals as a systematic body of knowledge (Wissenschaft) is possible, it therefore struggles to explain why a metaphysics of morals would be worthwhile.

In the sections below, I suggest an alternative reading. On my view, the *Metaphysics of Morals* is an attempt to make explicit the rational element of our moral knowledge as we find it in our shared institutions and conventional virtues. Rather than attempting to deriving a representative set of duties from a previously established test, Kant's *Metaphysics of Morals* takes up and moves through the content of commonly read treaties on ethics from Kant's time, and attempts to make explicit the

---

[16] Gregor 1963 p. xi. Note that regardless of how plausible we find this view as an interpretation of the *Metaphysics of Morals*, it is already questionable as an interpretation of the *Groundwork*. Because in the *Groundwork*, Kant already envisages a future metaphysics of morals; but if morality consists essentially of a practical-logical criterion right action, why would the book which allegedly lays out this criterion not be enough, and why would Kant at that point already think it so important to eventually write a metaphysics of morals?

[17] For Kant's claim that he will lead metaphysics to the status of a Wissenschaft see KrV Bxxiii.

rational basis of the institutions and virtues contained therein. In this way, we can explain both how a metaphysics of morals is possible and why it is worthwhile. On Kant's view, a metaphysics of morals is possible because a body of moral cognition can be systematically connected, or unified, by an a priori principle. Explaining how such a body of moral cognition can be systematically connected in detail is also worthwhile because it shows how our existing practices of ethical life have a genuinely rational basis.

What Would Make Moral Philosophy Metaphysics?

In the introduction to his first *Critique*, Kant famously tells us that he wants to bring metaphysics onto the "secure course of a science [Wissenschaft]." (KrV Bxxiii) Thus, any subject pretending to be a metaphysics in Kant's sense must meet a handful of requirements – first, to count as a science, and second, to count as metaphysics. On Kant's view, a doctrine (Lehre) must meet two requirements in order to count as a Wissenschaft. First, the cognitions which constitute the doctrine must be necessary and carry conviction (4:468).[18] That is, they must qualify as knowledge. Second, it must be a system, rather than merely an aggregate of cognitions (KrV A832/B860).[19] In order to have such systematic unity, a science needs a unifying idea and scheme (KrV A831/B861). In other words, a mere assemblage of information cannot be a Wissenschaft. In addition, metaphysics is cognition beyond experience (4:265) and thus involves cognition a priori (4:266). Consequently, moral philosophy as Kantian metaphysics is possible if there is (i) moral cognition

---

[18] However, note that this second requirement only applies to what Kant calls the rational sciences. The exception to this second requirement are thus the historical sciences (9:72).
[19] See also 9:72 and 9:139

that qualifies as knowledge, (ii) this cognition is unified into a coherent system by an idea or principle, and (ii) the unifying principle is a priori. Let me briefly outline these requirements in turn.

If the metaphysics of morals is to count as a Wissenschaft, it must be a systematic body of moral knowledge. However, not all practical cognition qualifies as knowledge. According to Kant, theoretical knowledge recognizes what is, practical knowledge what should be (KrV A634/B662). Thus, practical cognition differs from theoretical cognition not in the regard to the kind of objects that are represented, but in the *way* in which it refers to objects (KrV Bx). More specifically, practical and theoretical reason differ regards the actuality of their object of representation.[20] Put briefly, practical cognition is that which concerns our representation of objects as particular (and potential) ends to be brought about. And importantly, for the purpose of this discussion, Kant appears to apply the same taxonomy to the levels or hierarchy of practical cognition as to theoretical cognition. Kant draws a threefold distinction between kinds of cognition (Erkenntnis), or things that are cognizable: knowledge (Wissen), belief (Glaube),[21] and opinion (Meinung).[22] According to Kant's explanation of this distinction, we know something when taking it to be true on grounds that are both subjectively and objectively justified (Log 9:66ff).[23] This kind of cognition consists in objectively valid judgments which are discursive and can be communicated via concepts. Thus, knowledge is certain and to know something is to have conviction (Überzeugung). By contrast,

---

[20] For a helpful overview of Kant's distinction between practical and theoretical knowledge see Engstrom 2002 pp. 55-61. There Engstrom shows how, according to Kant, all knowledge (both practical and theoretical) involves the idea of representations determining the form of their object a priori, and on how the respective kinds knowledge differ (namely as regarding the actuality of their object).

[21] When talking about belief in God and the immortality of the soul, it can be tempting to translate 'Glauble' as 'faith.' However, there are no distinct words for faith and belief in German, and I believe that the context in which Kant uses the term 'Glauble' support translating it consistently as 'belief,' which also allows one to translate the corresponding verb 'glauben' consistently as believing/to believe.

[22] Sometimes, Kant also uses the word 'cognition' (Erkenntnis) interchangeably with 'knowledge' (Wissen), as in A50/B74. However, he usually adheres to the distinction in his usage and explicitly discusses the distinction in both the first and the third *Critique* at KrV A820/B848ff and KU 5:467ff, and the Jäsche Logik (Log 9:66ff).

[23] See also KrV A822/B850

belief (Glaube) is taking something to be true on the basis of a subjectively sufficient but objectively insufficient ground (Log 9:67). Although belief (Glaube) does not entail conviction, it is already sufficient to determine action.[24] Finally, opinion (Meinung) is taking something to be true on grounds that are both subjectively and objectively insufficient (Log 9:66-7). Consequently, an interpretation of the *Metaphysics of Morals* must be able to explain how there can be moral knowledge, i.e., moral cognition which consists in objectively valid judgments, which are discursive and can be communicated via concepts.

In addition, metaphysics must not only be a body of knowledge, but a systematic, or unified, body of knowledge. In the introduction to the *Metaphysical Foundations of Natural Science* of 1786, Kant explains that "Every doctrine that is supposed to be a system, that is, a whole of cognition *ordered according* to principles, is called a science." (MAN 4:467, emphasis added) Later on, in his preface to the *Metaphysics of Morals*, Kant makes a similar claim regarding a system of ethics. In this passage, Kant defends the view that there can only be one true philosophical system, since there cannot be two incompatible- but equally true philosophical views. In this context, he says:

> So the moralist rightly says that there is only one virtue and one doctrine of virtue, that is, a single system that *connects all duties of virtue by one principle*; the chemist, that there is only one chemistry (Lavoisier's); the teacher of medicine, that there is only one principle for systematically classifying diseases (Brown's). (MS 6:207, emphasis added)

The significance of this passage lies in the way Kant puts his claim about morality: that there is one "single system that *connects* all duties of virtue *by one principle*." This one (a priori) principle is the form of pure practical reason: the Categorical Imperative. Thus, a body of knowledge constitutes a system if it is 'ordered according' to or 'connected by' principles. An interpretation of the *Metaphysics of*

---

[24] See also Neiman 1994 pp. 157-158 and Wood 1970 pp. 15-17.

Morals must thus explain how the body of moral knowledge discussed in the book is supposed to be unified into a coherent system.

Finally, the a priori requirement: all Kantian metaphysics concerns a priori knowledge. However, this does not mean that Kantian metaphysics *only* contains a priori knowledge. Kant's critical philosophy contains an important distinction between 'pure' and 'a priori' knowledge. Pure knowledge would be knowledge we have entirely independently of all sense experience. Importantly, pure knowledge concerns not merely the connection between concepts but also the content of the connection: to have *pure* knowledge would be to know both the content of two subject- and predicate-concepts as well as their connection independently of all empirical knowledge. Thus, pure knowledge not only asserts an a priori connection between two empirical concepts, but asserts an a priori connection between two concepts that are not taken from experience. That is, pure knowledge would assert an a priori connection between two concepts that we derive from a reflection upon our activity of reason. 'Pure' knowledge must therefore be distinguished from knowledge in which an a priori connection is made between empirical concepts. In order to count as metaphysics in Kant's sense, the *Metaphysics of Morals* need not be a system of 'pure' knowledge alone; it is enough if the *Metaphysics of Morals* concerns a priori principles. Thus, an interpretation of the *Metaphysics of Morals* must thus give attention to how Kant unifies a body of moral knowledge into a system by an a priori principle.

The Standard Assumption and its Problems

So far, a common reading of Kant's *Metaphysics of Morals* has been to understand its application of a priori principles to a system of duties as the attempt to derive duties from a test: the

166

Groundwork gave us the decision-procedure, and the *Metaphysics of Morals* ought to derive a system of

duties from it. The plausibility of the standard reading comes from the fact that it offers an easy way

to meet the three requirements outlined above for a metaphysics of morals: a system of duties

derived from the Categorical Imperative would seem to be knowledge, since it would be the

application of secure a priori cognition to objective empirical knowledge about human nature;

because the system would be derived from one principle, it would constitute a unified whole; and

because the Categorical Imperative is the form of pure practical reason, such a system of derived

moral knowledge would be based in cognition a priori. As Mary Gregor describes it in the

introduction to her translation, "[the *Groundwork's*] purpose is to lay the foundation for the projected

Metaphysik der Sitten, to investigate the supreme principle of morality, from which, in the subsequent

work, 'the whole system' of duties will be derived."[25] Such claims about the purpose of the

Metaphysics of Morals are (unfortunately) more often stated in casual prefaces to work on other topics

than defended on their own terms. Recent examples include Andrea Esser's recent discussion of

moral conflict in the *Doctrine of Virtue*, where she claims that "they [the *Groundwork* and *Critique of*

Practical Reason] only serve for the preparation of a so-called metaphysics of morals, the task of which

is to devise the contents of the general moral principle and therefore show consideration for the

special nature of man and the particular circumstances given in each individual case."[26] Similarly,

David James prefaces his discussion of Kant's casuistical questions by claiming that the *Doctrine of*

---

[25] Gregor 1963 p. xi. See also Gregor's translator preface 1991 p. 1. Note, however, that Kant does not literally say that the 'whole system of duties' will be derived from the categorical imperative. That is, Kant does not literally say that the categorical imperative will be applied *to human nature* in order to derive 'the whole system' of duties from this application. Technically speaking, Kant only says that a *Metaphysics of Morals* will apply the highest principle of morality *to the whole system*.

[26] Esser 2008 pp. 279-280

Virtue "was intended to display the system of duties which follow from applying the categorical imperative to circumstances in human life."[27]

On this reading, the *Metaphysics of Morals* does not add much beyond the demonstration of the alleged test's application. On a less extreme version of this reading, such a derivation of a system of duties for human beings from the Categorical Imperative might add details to the alleged test developed in the *Groundwork* and correct potential misunderstandings. According to Allen Wood, for instance, the central purpose of the *Metaphysics of Morals* is to derive duties for human beings by means of the Categorical Imperative (and in particular the Formula of Humanity);[28] and according to Wood, this application in the *Metaphysics of Morals* is still crucial for understanding Kant's ethics in addition to the *Groundwork*, for it could potentially correct some common misunderstandings about the content of the duties that one can supposedly derive from the Categorical Imperative.[29]

However, in both its more and less extreme version, this standard assumption of what a metaphysics of morals should contain fits only awkwardly with the actual content of the *Metaphysics of Morals*. Two textual misfits are particularly perspicuous: there are few, if any, passages in the *Metaphysics of Morals* that can seriously be read as applications of a test or decision-procedure for the derivation of duties. In addition, Kant's own description of the project simply does not provide much support the standard assumption either.

---

[27] James 1992 p. 67
[28] Wood 1997 pp. 12-13
[29] Wood 1997 p. 4

The Absence of a 'Test Function' of the Categorical Imperative

By far one of the most striking features of Kant's *Metaphysics of Morals* is that neither part of the work contains the kind of derivation of duties from the Categorical Imperative that readers often ascribe to the *Groundwork*. In this earlier work, Kant is usually read as providing four derivations of duties from the Categorical Imperative – first from the Formula of Universal Law, and then from the Formula of Humanity. But Kant's discussion of duties in the *Metaphysics of Morals* does not contain any derivations in the style of the *Groundwork*. This might be less surprising in the case of the first part, the *Doctrine of Right*, since Kant does not base his discussion of right directly on the Categorical Imperative but, instead, on the Universal Principle of Right. However, this absence of a derivation in the style of the *Groundwork* is particularly striking in the case of the *Doctrine of Virtue*. As Mary Gregor has pointed out, Kant "plunges directly into discussions of specific duties, leaving the reader to abstract from these discussions the subordinate principles operative in them."[30]

Much of the main body of Kant's discussion in the *Doctrine of Virtue* focuses on the role our human nature plays for ethical life. Kant goes through a discussion of our impulses, aspects of our animal nature, which lead him into a discussion of the specific duties that we owe to ourselves as animal beings. Importantly, for the purpose of this chapter, Kant then focuses on explaining which concrete ends humans have a duty to adopt. In all of this, the Categorical Imperative and the formulations provided in the *Groundwork* are strikingly absent in the *Doctrine of Virtue*. Kant refers to the 'dignity of humanity' at several instances, but never in the context of deriving a concrete duty. For instance, Kant refers to the dignity of humanity when explaining the notion of a duty to oneself

---

[30] Gregor 1963 p. xii

as a moral being (MS 6:420) and when discussing the appropriate object for respect (MS 6:435).[31]

But it would be bold hermeneutics to claim that such references should be understood as an

application of the Formula of Humanity. After all, if it really is an application of the Formula of

Humanity, why not say so? And if it was such an application, why would Kant not mention the

formula even once?

The closest Kant gets to deriving a duty from the Formula of Humanity in the *Doctrine of

Virtue* is when, in section II of part II, he says: "Humanity itself is a dignity; for a human being

cannot be used merely as a means by any human being (either by others or even by himself) but

must always be used at the same time as an end." Thus, one "is under obligation to acknowledge, in

a practical way, the dignity of humanity in every other human being." (MS 6:462) On the one hand,

this gets close to a derivation insofar as Kant here makes an argument, with a premise sounding

much like the Formula of Humanity, and derives a particular conclusion from this premise.

However, even this single passage does not provide a derivation of a duty from the Formula of

Humanity in the style we might expect from reading the *Groundwork*, and Kant does not mention any

maxims which could be 'tested' in light of the formula.

Even more striking than the way in which Kant refers to the 'dignity of humanity' is the

rarity of references to the universalizability of maxims in the *Doctrine of Virtue*.[32] The closest Kant

comes to deriving a concrete duty from the Formula of Universal Law is in the passages discussing

---

[31] For other references to 'humanity' of this kind in the *Doctrine of Virtue* see also 6:423, 6:424, 6:429, 6:436, 6:446, 6:447, 6:449, and 6:459, 6:463, 6:466, 6:467.

[32] The almost complete absence of the formula of universal law in the *Metaphysics of Morals* has been frequently noted. See for instance, Geiger 2010 p. 274, Herman 1993 p. 133, Hill 2007 p. 486, Lottenbach and Tenenbaum 1995 p. 228, Korsgaard 1996 p. 124, O'Neill 1975 p. 33; Wood 1997 pp. 4, 11-12 and 1999 pp. 139-40. However, all these authors also claim that Kant derives duties in the *Doctrine of Virtue* from the formula of humanity. Merely O'Neill remarks that, in that light, "Kant's arguments are often extremely sketchy and appeal to the *Formula of the End in Itself* (especially in the second part of the *Metaphysics of Morals*) or to various 'natural purposes'" (O'Neill 1975 p. 33). But how Kant's remarks can about humanity in the *Doctrine of Virtue* be understood as a genuine 'derivation' is unfortunately never discussed.

the duty of benevolence. In a section entitled 'Exposition of duties of virtue as wide duties,' Kant writes the following about the duty to render assistance to others:

> The reason that it is a duty to be beneficent is this: since our self-love cannot be separated from our need to be loved (helped in case of need) by others as well, we therefore make ourselves an end for others; and the only way this maxim can be binding is through its qualification as a universal law, hence through our will to make others our ends as well. The happiness of others is therefore an end that is also a duty. But I ought to sacrifice a part of my welfare to others without hope of return, because this is a duty, and it is impossible to assign determinate limits to the extent of this sacrifice. How far it should extend depends, in large part, on what each person's true needs are in view of his sensibilities, and it must be left to each to decide this for himself. For, a maxim of promoting others' happiness at the sacrifice of one's own happiness, one's true needs, would conflict with itself if it were made a universal law. Hence this duty is only a wide one; the duty has in it a latitude for doing more or less, and no specific limits can be assigned to what should be done. - The law holds only for maxims, not for determinate actions. (MS 6:393)

Later on, in a passage discussion 'the duty of love in particular,' Kant is arguing that the morally obligatory end of benevolence also includes being benevolent to oneself. For, Kant argues, if it excluded oneself, it would lack the universality of a law.

> I want everyone else to be benevolent toward me (benevolentiam); hence I ought also to be benevolent toward everyone else. But since all others with the exception of myself would not be all, so that the maxim would not have within it the universality of a law, which is still necessary for imposing obligation, the law making benevolence a duty will include myself, as an object of benevolence, in the command of practical reason. (MS 6:451, see also MS 6:453)

These passages are the closest Kant ever comes to deriving duties from the Categorical Imperative in the form of the Formula of Universal Law in the *Metaphysics of Morals*, a book of roughly 300 pages. However, the duties discussed in these three passages are also exceedingly broad. Rather than telling us what to do in specific circumstances, as in the application of a decision-procedure, Kant appears to tell us how the idea of universality underlies the general moral end of benevolence. If the

Metaphysics of Morals is supposed to contain the application of a test or decision-procedure, it is an exceedingly abstract and rare application.

As this (highly) limited use of the formulas in the *Metaphysics of Morals* shows, to interpret Kant's project as a derivation of duties from the Categorical Imperative would be to mischaracterize the content *Metaphysics of Morals*. Consequently, we should also disagree with authors like Allen Wood who has claimed that Kant "overwhelmingly prefers the Formula of Humanity as the formula in terms of which the moral law is to be applied." And that the Formula of Humanity is used in "justifying no fewer than nine of the fourteen ethical duties."[33] As Manfred Kuehn as pointed out correctly, Wood here misconstrues the purpose of Kant's discussion: Kant is not attempting to derive duties from the formulas of the Categorical Imperative; rather, "he [Kant] 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."[34]

Kant's Own Descriptions of his Project

Still, one might remain unmoved by this absence of a test function of the Categorical Imperative in the *Doctrine of Virtue* and respond that support for the standard reading should come from Kant's own description of his project of a metaphysics of morals in the *Groundwork* and in the *Metaphysics of Morals*. In the *Groundwork*, Kant remarked that he had to limit the scope of this work to the identification of the highest principle of morality, and that its application would be left for a future metaphysics of morals. There, he says:

---

[33] Wood 1997 pp. 11-12
[34] Kuehn 2010 p. 24

No doubt my assertions on this important and central question, discussion of which has till now been far from satisfactory, would receive a great deal of light from the application of the same principle to the whole system, and of confirmation through the adequacy that it would everywhere show; but I had to forgo this advantage [...] (GMS 4:392)

Mary Gregor has interpreted this passage as suggesting the standard reading. In the introduction to her translation of the *Metaphysics of Morals*, she writes: "As the tide of the *Grundlegung* implies, its purpose is to lay the foundation for the projected *Metaphysik der Sitten*, to investigate the supreme principle of morality, from which, in the subsequent work, *'the whole system' of duties will be derived*."[35] However, note that in the passage quoted above, Kant does not claim that the 'entire system' will be derived from the highest principle of morality. Technically speaking, Kant says that in a future metaphysics of morals, the highest principle of morality will be applied *to the* entire system of duties. Of course, if we gratuitously presuppose the standard reading, we will likely also interpret this passage as suggesting a derivation of the entire system from the Categorical Imperative. But if we do not so presuppose this standard reading, then the passage remains simply inconclusive. Rather than supporting the standard reading, the passage begs the question what it would be to apply the insights from the *Groundwork* 'to the entire system.'[36] Of course, this is not to say that this small formulation can carry all the weight in convincing readers to reevaluate their assumptions about the *Metaphysics of Morals*; but neither does this small passage provide the support for the standard reading that e.g., Gregor's gloss suggests it does.

Another passage that has been interpreted as support of the standard reading comes from Kant's introduction to the *Metaphysics of Morals*. There, Kant compares his project to a metaphysics of nature, and briefly indicates how it relates to anthropology.

---

[35] Gregor 1963 p. xi, emphasis added
[36] I address this question in some detail in the final section of this essay.

But just as there must be principles in a metaphysics of nature for applying those highest universal principles of a nature in general to objects of experience, a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles. But this will in no way detract from the purity of these principles or cast doubt on their a priori source. - This is to say, in effect, that a metaphysics of morals cannot be based upon anthropology but can still be applied to it. (MS 6:216-7)

As in the case of the passage from the *Groundwork*, we might initially be tempted to interpret this as suggesting that a metaphysics of morals will apply the categorical imperative to our knowledge of human nature, and thus derive a system of duties from it. However, the difficulty is that, technically speaking, Kant here talks about is the application *of the Metaphysics of Morals* to anthropology – not of the Categorical Imperative to anthropological knowledge. Thus, this passage suggests that the *Metaphysics of Morals* itself has principles of application. Especially regarding the case of the *Doctrine of Right*, this seems quite intuitive. After all, Kant believes that his doctrine of right is only an outline, and that this work cannot actually provide the entirety of all our duties of right. Because the concept of 'right' applies to practice, Kant thinks that a metaphysics of right would need to account for the entire manifold of empirical cases in order to be complete. But clearly, this is impossible to provide. Hence, Kant's discussion of right is merely a metaphysical first principles [Anfangsgründe] of right: "It [the *Doctrine of Right*] will be dealt with as in the (earlier) *Metaphysical First Principles of Natural Science*: namely, that right which belongs to the system outlined a priori will go into the text, while rights taken from particular cases of experience will be put into remarks, which will sometimes be extensive […]." (MS 6:205). Unless we gratuitously presuppose the standard reading, I believe that this passage is just as inconclusive about the purpose of Kant's *Metaphysics of Morals* as his remark in the preface to the *Groundwork*. Rather than supporting the standard reading, Kant's remark in the

introduction raises the (now slightly different) question what it means to apply a metaphysics of morals to anthropological knowledge.[37]

A third passage potentially supporting the standard reading might be in the preface to the *Doctrine of Virtue.* There, Kant introduces his work by discussing the need for a metaphysics of morals. Although Kant thinks that few people will appreciate such technical philosophy, he also believes that it is still important to demonstrate that the basis of our moral knowledge lies in reason and not in feeling. In this context, he remarks:

> […] a popular teacher can indeed be content to rely on a certain feeling which, because of the results expected from it, is called moral, insofar as he insists that the following lesson be taken to heart, as the touchstone for deciding whether or not something is a duty of virtue: "How could a maxim such as yours harmonize with itself if everyone, in every case, made it a universal law?" But if it were mere feeling that made it our duty even to use this proposition as the touchstone, this duty would not be dictated by reason but would be taken to be a duty only instinctively, and hence blindly. But in fact no moral principle is based, as people sometimes suppose, on any feeling whatsoever. Any such principle is really an obscurely thought metaphysics that is inherent in every human being because of his rational predisposition, as a teacher will readily grant if he experiments in questioning his pupil socratically about the imperative of duty and its application to moral appraisal of his actions. - The way the teacher presents this (his technique) should not always be metaphysical nor his terms scholastic, unless he wants to train his pupil as a philosopher. But his thought must go all the way back to the elements of metaphysics, without which no certitude or purity can be expected in the doctrine of virtue, nor indeed any moving force. (6:376)

Kant here indicates that the formulae of the Categorical Imperative have heuristic value. And indeed, when Kant is saying that the application of the 'touchstone' must be supported by critical metaphysics, it might sound like he is also describing his project in the *Metaphysics of Morals.* However, in this passage Kant refers to the application of the touchstone as an edifying method of a

---

[37] For a detailed commentary on all of Kant's remarks in the introduction to the *Metaphysics of Morals* and the *Doctrine of Virtue* see Zöller 2013.

'popular teacher,' whereas metaphysics is what *precedes* this and gives the popular teacher the adequate resources for their work by making explicit the rational basis of our (common) moral knowledge. And in the preface to the *Metaphysics of Morals*, one would expect Kant to use the notion of 'metaphysics' to refer to the project of the book with this title. Once again, Kant's introductory remark appears far less conclusive than one might have hoped. Rather than supporting the standard reading, this passage raises the question what Kant means by a touchstone in the popular teaching of ethics, and what practical purpose this touchstone serves.[38]

What, then, should we make of these difficulties for the standard assumption about the purpose of Kant's *Metaphysics of Morals?* Do we have to follow the readers who have posited a break between Kant's earlier work and the *Metaphysics of Morals?* Ottfried Höffe, for instance, has recently pointed out correctly that one looks in vain for a 'test function' of the Categorical Imperative in the *Doctrine of Virtue*, and consequently claimed that this absence marks a significant break between Kant's *Groundwork*, the second *Critique* and the *Metaphysics of Morals*.[39] However, I believe that no such thing is necessary. As I have argued elsewhere, we have good reason to doubt the common assumption that Kant intended his discussion of the formulas of the Categorical Imperative in the *Groundwork* to supply a test or decision-procedure.[40] And as I have argued in the previous chapter, we have good reason to understand Kant's project of a 'metaphysics of morals' in section II of the

---

[38] As I have argued in 'Conceptions of Moral Knowledge and the Practical Purpose of Moral Philosophy,' the practical purpose of Kant's ethics is best understood as therapeutic. In this way, the formulations of the Categorical Imperative help make explicit the implicit moral cognition of rational agents in order to help them overcome temptation against their better moral knowledge.

[39] "Kant veröffentlicht seine systematische Moralphilosophie, die zweiteilige *Metaphysik der Sitten*, zwölf Jahre nach der ersten systematisch relevanten Schrift, der *Grundlegung* (1785), und knapp ein Jahrzehnt nach der *Kritik der praktischen Vernunft* (1788). In beiden Texten, am stärksten in der *Grundlegung*, spielt der kategorische Imperativ in seinen beiden Funktionen eine Rolle, zum einen als semantisches bzw. Wesenselement – die Moral nimmt für sinnliche Vernunftwesen die Gestalt einer uneingeschränkten Verbindlichkeit an -, zum anderen als Test: Maximen sind moralisch, wenn sie sich nach dem Muster eines Naturgesetzes streng verallgemeinern lassen. In der *Tugendlehre* sucht man die Testfunktion vergeblich." (Höffe 2020 p. 163)

[40] Ramsauer 2024 and chapter 4.

176

Groundwork as an attempt to analyze the rational principles that underlie our moral judgment and their systematic interconnection. In the following section, I argue that the *Metaphysics of Morals* provides the long-promised elaboration of that very project.

An Alternative Reading: Systematizing Moral Knowledge

A metaphysics of morals is possible if there is (i) moral cognition that qualifies as knowledge, (ii) this cognition is unified into a coherent system, and (ii) its unifying principle has a priori status. In what way could Kant's *Metaphysics of Morals* be 'applying' a priori principles such that it unifies a body of (common) moral knowledge into a system? I want to propose the following reading: Kant's *Metaphysics of Morals* is trying to make explicit the rational basis of our common moral knowledge by showing how the highest principle of morality underlies a system of commonly accepted moral duties as it can be found in treaties on ethics from his time. In this way, neither the *Doctrine of Right* nor the *Doctrine of Virtue* are attempting to derive a body of duties from the Categorical Imperative, as it were, for the first time, with a new 'method of ethics.' The form of practical reason allows us to unify a system of common moral knowledge that covers those two aspects of the moral world: our moral cognition of the rightfulness of external actions, and our moral cognition of the ends that humans should adopt. Attempting to derive an entire system of duties from a decision-procedure may be one way of making explicit the rational basis of a body of moral knowledge. But depending on how confident one is about common moral judgment, one might also work in the other direction: starting with the specific duties and moral ends we already acknowledge, and trying to highlight the a priori structure that underlies our knowledge of these duties and ends. Let me briefly consider each of the three requirements as it relates to my proposed reading in turn.

177

First, The *Metaphysics of Morals* aims to provide a system of moral cognition that qualifies as *knowledge*. Moral knowledge in the *Metaphysics of Morals* is union of a priori cognition and empirical (anthropological) knowledge about human nature. Knowledge about our nature as finite, and dependent beings can be both subjectively and objectively justified, as can knowledge about social practices and institutions. In other words, we can plausibly make objectively valid judgments about human nature, social practices and institutions that are discursive and can be communicated via concepts. We can also, according to Kant, gain insights into the a priori principles of practical reason (i.e., the kind of principles that make representations of actions as good so much as possible), which are also the subject of the *Groundwork* and the *Critique of Practical Reason*.

Second, the *Metaphysics of Morals* can present a *system* of moral knowledge if they are unified by an idea or principle. The unifying principle of the *Metaphysics of Morals* is the form of practical reason of a free, rational being: the Categorical Imperative. As Kant tells us in his introduction to the work, there are several concepts which both parts of the *Metaphysics of Morals* have in common. The first one mentioned in Kant's section on 'preliminary concepts of the metaphysics of morals' (MS 6:221) is the concept of freedom, which makes practical lawgiving possible. The form of this lawgiving of a free being is the Categorical Imperative, "which as such only affirms what obligation is" (MS 6:225). The form of pure practical reason figures prominently throughout the entire work – especially in the *Doctrine of Virtue*. However, as I pointed out in section three above, the way in which key concepts like humanity figure in the *Metaphysics of Morals* cannot adequately be described as applications of a test. More adequately, these references can be characterized as systematizing our common moral knowledge by making explicit the a priori elements of conventional duties. In this way, a system of moral knowledge is possible because the form of pure practical reason allows us to

connect this knowledge into a coherent system. As Kant claims in the preface to the *Metaphysics of Morals*, there is "a single system that connects all duties of virtue by one principle" (MS 6:207).

Lastly, this system of moral knowledge is unified by *a priori* principles, concepts of pure practical reason. Because our knowledge of the form of pure practical reason is cognition a priori, the system of moral knowledge unified by this principle constitutes a metaphysics of morals.

In this way, the alternative reading can equally well explain how a metaphysics of morals is possible, meeting all three requirements. Moreover, we also avoid the initial puzzle about the *Metaphysics of Morals*: we do not need to assume any change of mind or senile relapse, because Kant's early and late works on moral philosophy are compatible: in his earlier works, Kant attempted to analyze the form of pure practical reason that is contained in our common moral cognition, and to demonstrate the concept's applicability to finite rational beings; in his later work, he starts from a body of moral knowledge and attempts to demonstrate how the a priori form of practical reason unifies this moral knowledge into a coherent system. Therefore, it is just not the case that "everyone who comes to the old Kant's *Metaphysics of Morals* from the *Groundwork* and the *Critique of Practical Reason* will not be able to avoid the question if here is really offered the earlier-announced, named after its idea, 'Metaphysics of Morals.'"[41] Quite to the contrary, on the reading I propose, the *Metaphysics of Morals* is precisely what Kant promises us in the *Groundwork*. There, Kant suggests that a future metaphysics of morals would apply the highest principle of morality to the whole system of ethics.

> No doubt my assertions on this important and central question, discussion of which has till now been
> far from satisfactory, would receive a great deal of light from the application of the same *principle to the*

---

[41] Anderson 1923 p. 41, translation mine. I discuss this alleged tension in detail in chapter 5, 'A Sudden Change of Mind?'

whole system, and of confirmation through the adequacy that it would everywhere show; but I had to forgo this advantage […] (GMS 4:392)

Although many commentators have read this passage as if it talked about an application to our knowledge of human nature from which we then derive the entire system, I believe that we can take Kant's remark here perfectly literally: already in 1785, Kant intended to produce a metaphysics of morals in which the rational basis of common moral knowledge would be demonstrated. And we find further support in section II of the *Groundwork*, in Kant's infamous four examples of a 'derivation of imperatives of duty.' After telling us that he will now enumerate a couple of duties according to the (at his time) commonly recognized divisions, Kant adds that: "it must be noted here that I reserve the division of duties entirely for a future Metaphysics of Morals, so that the division here stands only as one adopted at my discretion (for the sake of arranging my examples)." (GMS 4:421) As this indicates, already when writing the *Groundwork* Kant thought that at least one central purpose of the *Metaphysics of Morals* would be to provide a rationally determined classification of our ethical duties into separate divisions.[42]

Moreover, on the reading I propose, there is nothing odd or even 'uncritical' about Kant's reliance on earlier authors. Throughout the *Metaphysics of Morals*, Kant often closely follows the content of other treaties on ethics. For instance, Kant's discussion in the *Doctrine of Right* overlaps largely in content with in the topics covered by Gottfried Achenwall, in his *Jus naturae in usum auditorum* (first published in 1763). And Kant's discussion in the *Doctrine of Virtue* overlaps significantly with Alexander Gottlieb Baumgarten's *Ethica philosophica* (Halle 1751) and the *Initia philosophiae practicae primae acroamatice* (Halle 1760).[43] That Kant's moral and legal philosophy was

---

[42] For a similar claim, see Theunissen 2013 p. 111
[43] For a detailed summary of the overlap between Kant's *Metaphysics of Morals* and the ethics of Baumgarten and Wolff see especially Anderson 1923 pp. 55-61

significantly influenced by both Achenwall and Baumgarten is clear from the fact that Kant used their works as textbooks for his lectures on ethics and jurisprudence.[44] This overlap in content has led several readers to suggest that Kant's *Metaphysics of Morals* is not sufficiently 'critical,' and that it might represent a relapse of his old age.[45] More sympathetic readers, like Manfred Kuehn, have suggested 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."[46] However, I believe that it would be a mistake interpret this the overlap between the contents of Kant's *Metaphysics of Morals* and earlier works like Achenwall's *Jus naturae* and Baumgarten's *Initia* and *Ethica* as evidence of anything 'uncritical.' Whether or not Kant's *Metaphysics of Morals* can be considered part of his critical philosophical system depends on whether or not we accept the standard reading. And as I have argued, we have good reason to be cautious about this traditional reading of Kant's *Metaphysics of Morals*. Consequently, the overlap in content between Kant and earlier writers may not be problematic at all. Indeed, on the reading I propose, Kant's reliance on both Achenwall and Baumgarten makes perfect sense: if the *Metaphysics of Morals* is trying to make explicit the rational basic of our common moral knowledge as we find it in existing doctrines of ethics, then we might expect that Kant would not be averse to relying on earlier writers regarding the specific duties, virtues and vices of human beings. We might in fact expect that Kant would take the content of systematic treaties on ethics from his own time, and would move through the topics in there, making explicit along the way how the form of practical reason underlies each of these duties. This, I suggest, is what Kant means by applying the Categorical Imperative 'to the entire

---

[44] See the, *Vorlesungsverzeichnisse der Universität Königsberg* (1720-1804) edited by Oberhausen and Pozzo 1999.
[45] See for instance Ritter 1971 and Ilting 1983. Cf. Busch 1979.
[46] Kuehn 2010 p. 20, emphasis added.

system:' we start with the content of what is at the time supposed to be a coherent, systematic treatment of human duties, revise them in light of Kant's analysis in the *Groundwork* and second *Critique*, and show how the form of practical reason underlies these duties.

On my proposed reading, the most interesting aspect of Kant's *Metaphysics of Morals* is that it demonstrates Kant's commitment to the tangibility of the ideal. Kant is rightly remembered for locating ethics in an unreachable realm of the highest good and acting from duty, a final goal we cannot reach in our finite life, and in a way of acting that we can never be certain we have actually achieved. To many readers, this can make it seem as if ethical life is necessarily an alienated life on Kant's view. While I believe that Kant had good reasons to believe that our true motivations are opaque to finite, rational beings like us; and that he had good reasons to believe that the highest good – happiness in accordance with complete virtue – is realistically unreachable for us in this life, Kant's moral agent is not the alienated agent. Already in the *Groundwork*, Kant shows how his analysis of the form of pure practical reason underlies our common moral knowledge through his four examples of 'derivations of imperatives of duty;' the *Metaphysics of Morals* shows how the form of pure practical reason underlies and systematizes the moral knowledge contained in ethical treatise of his time into a coherent system. The virtues and practices Kant discusses in the *Metaphysics of Morals* are not alien to his readers. In fact, they already form part of their lifeworld in some sense already. Thus, we have a sense of what moral progress would look like, right here, right now: for things to go well for beings like us would be for us to feel home in our ethical life. This is a goal we cannot ever realistically believe to have achieved, but it is a goal we can all realistically hope to make progress in, right here, right now. In this way, Kant's *Metaphysics of Morals* is a worthwhile project, deserving of its protracted origins.

Conclusion

Kant's last work on moral philosophy is less utopian than archeological. The *Metaphysics of Morals* is not an exercise in deriving a novel system of duties from the Categorical Imperative; rather, it is and attempt to make explicit the rational basis and structure of a system of duties and moral ends we are already familiar with. Thus, it seems misleading to characterize the *Metaphysics of Morals* as a work of what we would nowadays calls 'normative ethics.'[47] The difference between unearthing what is rational in our commonly accepted moral knowledge and deriving a system of duties and ends from a supreme principle is perhaps subtle, but philosophically and exegetically significant: if we read Kant's *Metaphysics of Morals* as unearthing the rational basis of a systematic body of moral knowledge, then there is no tension between the *Groundwork* and the *Metaphysics of Morals*, and we need not choose which work to emphasize and which one to downplay.

---

[47] As e.g., Thomas Hill (2010) has suggested regarding the *Doctrine of Virtue.*

Chapter 7

Kant's Casuistical Questions

Introduction

One of the most interesting features of Kant's late practical philosophy are the casuistical questions in the *Metaphysics of Morals*. These are eight small sections Kant added to each of his discussions of specific ethical duties in the first half of the *Doctrine of Virtue*. In one of them, for instance, Kant asks us to consider the following situation: "An author asks one of his readers 'How do you like my work?' […] The author will take the slightest hesitation in answering as an insult. May one, then, say what is expected of one?" (MS 6:4319) In another, shortly thereafter, Kant asks his readers: "How far should one expend one's resources in practicing beneficence?" (MS 6:454) Whether or not we believe that we have good answers, Kant's casuistical questions point to ethical difficulties that all of us have faced and will continue to face in our life. Strikingly, these difficulties are familiar to everyone with even a modest amount of ethical sensitivity despite, on the face of it, being embedded in a long-gone cultural and historical context. The most perspicuous difficulty demonstrated by Kant's casuistical questions is the temptation to rationalize away the requirements of our immediate moral judgment in favor of our apparently reasonable inclinations. Thus, one of Kant's many insights in the *Metaphysics of Morals* is that moral challenges cannot be resolved by downplaying their difficulty; only by taking their difficulty *for us* serious can we make any attempt at better understanding – and dealing – with the challenges of our ethical lives.

Kant's casuistical questions have received comparatively little scholarly attention so far.[1] Moreover, they have traditionally been treated in ways that make it difficult to appreciate the different moral challenges they raise. On a common view of Kant's ethics as a theory primarily aimed at helping us figure out what is truly the right thing to do, it is plausible to expect that the casuistical questions are practice examples for the application of the theory's test or decision-procedure for right action; and on this view, the casuistical questions have usually been understood as either difficult cases of moral conflict (which do not lend themselves to easy resolution by maxim-testing)[2] or as rhetorical-questions that test the reader's comprehension of the decision-procedure.[3] However, Kant's casuistical questions do not easily fit this picture: the questions neither seem to be mere rhetorical questions, nor do the difficulties they highlight concern cases of moral conflict. Moreover, if read in this light, Kant's casuistical questions would not add anything to what either Kant's own earlier discussion or other philosophers have already pointed out. Consequently, if read in this light, we also lose sight of their arguably unique contribution.

Instead, I argue that we should read the casuistical questions against the background of the therapeutic purpose of Kant's moral philosophy. Locating Kant's conception of moral philosophy within a longstanding tradition of ethical reflection – one aimed at shaping human character and freeing us from our own forms of self-entrapment – helps us understand the point of the discussion of casuistical questions as it emerges at the specific juncture in the *Doctrine of Virtue*: their point is to articulate exemplary instances of the characteristic sort of moral difficulties which human beings

---

[1] To my knowledge, the only scholarly literature exclusively focusing on Kant's casuistical questions in German or English are Kittsteiner 1988, James 1992, Unna 2003, Kim 2009, Schüssler 2012, Patrone 2013. See also Di Giulio and Frigo (eds) 2020. Additional discussions can be found in Gregor 1963, Timmermann 2000, Forkl 2001, O'Neill 2002, Koch 2003, Esser 2008 and Theunissen 2013.
[2] Esser 2008 and Koch 2003 pp. 121-188. Unna 2003 takes this position regarding the duty of self-preservation and James 1992 regarding the questions on suicide.
[3] Schüssler 2012

face – difficulties which can be traced to certain sources. Like in the practice of a therapist, the first part of dealing with this sort of problem lies in first achieving an understanding of its origins, which in turn moves us closer to making genuine progress with respect to it. On my proposed reading, Kant seeks to provide an overview of the different species of moral difficulties which human beings face. Thus, the purpose of Kant's casuistical questions lies not so much in their potential answers but in the difficulties they highlight.

In the first section of this chapter, I provide a concise overview of the initial puzzle about Kant's casuistical questions and their contemporary reception. In section two, I outline the historical context of Kant's treatment of casuistry, which will provide a helpful background for my alternative interpretation of the casuistical questions and their role in the *Doctrine of Virtue*. For ease of discussion, section three reproduces Kant's casuistical questions. There, I also argue that Kant's questions are remarkably diverse and defy the categorizations available in previous literature. Finally, section four argues that Kant's casuistical questions should be read, against the background of the therapeutic purpose of Kant's ethics, as highlighting the diverse difficulties human beings face in their ethical lives.

Kant's Casuistical Questions: an Inconsistency or a Puzzle?

On the face of it, Kant's casuistical questions either give rise to an inconsistency or a puzzle. In an introductory remark added to his discussion of the divisions of a *Doctrine of Virtue,* Kant defines 'casuistry' as the practice of making determining judgments in the context of imperfect duties. Since imperfect duties like 'be benevolent' or 'help your neighbor' allow a certain latitude in how they are to be complied with, they further require determining judgment, i.e., 'casuistry.'

[E]thics, because of the latitude it allows in its imperfect duties, unavoidably leads to questions that call upon judgment to decide how a maxim is to be applied in particular cases, and indeed in such a way that judgment provides another (subordinate) maxim (and one can always ask for yet another principle for applying this maxim to cases that may arise). So ethics falls into a casuistry which has no place in the doctrine of right. (MS 6:411)

As Kant makes clear in this introductory passage, a treatment of ethics must contain a discussion of casuistry because ethics contains imperfect duties. The *Doctrine of Right*, however, which does not contain imperfect but only perfect duties, has 'no place' for casuistry. Kant thus explicitly rules out 'casuistry' in the case of perfect duties. And yet, six out of eight sets of casuistical questions are devoted to perfect duties.

One possible way of avoiding this seeming inconsistency is by differentiating 'casuistry' as discussed – and ruled out in the case of perfect duties – in the passage above from 'casuistical questions.' Not only would such a distinction avoid an inconsistency, it also fits with Kant's use of the terms – which consistently refers to the sets of questions added to his discussions of duties as 'casuistical questions,' not as 'casuistry.' However, this differentiation opens up a puzzle: if the casuistical questions are not part of what Kant described as 'casuistry', what are they? And what purpose are they supposed to serve within Kant's practical philosophy?

On the assumption that Kant's casuistical questions were instances of casuistry, we would have had a straightforward answer to what casuistical questions and their scholarly purpose in a metaphysics of morals would be – Kant tells us what purpose casuistry serves in the passage quoted above, and he goes on to describe how casuistry fits into the *Metaphysics of Morals*: "casuistry is not so much a doctrine about how to find something as rather a practice in how to seek truth. So it is woven into ethics in a fragmentary way, not systematically (as dogmatics would have to be), and is added to ethics only by way of scholia to the system." (MS 6:411) But as we saw, on this assumption

we also end up with a stark inconsistency: Kant explicitly ruled out casuistry in the case of perfect duties, and yet six out of eight sets of casuistical questions concern such perfect duties. By contrast, if we want to avoid this inconsistency and differentiate between casuistry and casuistical questions according to Kant's terminology, we have to reconsider the question what casuistical questions are and what their purpose in Kant's *Metaphysics of Morals* could be.

This initial puzzle about the possible purpose of casuistical questions about perfect duties is particularly salient in light of Kant's background assumptions on moral judgment. There is a long tradition of casuistry in the history of western philosophy, but the perceived need for casuistry usually arose for ethical theories that restricted full moral knowledge to an elite of trained experts, who would work out all the minute rules through difficult casuistry. By contrast, Kant notoriously believed that no one needs moral philosophy or instruction in order to know what is right and wrong;[4] and in his examples of moral judgment, he emphasizes that common moral agents can make moral judgments 'on the spot,' without the need for any elaborate casuistical method.[5] If Kant's casuistical questions are not about developing an expert taxonomy of derivative rules and exceptions, what purpose could the casuistical questions possibly serve?

How, then, does the existing literature navigate between the seeming inconsistency and the general puzzle about Kant's casuistical questions? So far, the most popular solution has been to attribute confusion to Kant. Thus, David James has speculated that Kant really meant to say 'broad' duties in his explanation of 'casuistry,' rather than what he actually said (namely 'imperfect' duties).[6]

---

[4] See KrV A43/B61-A44, KrV A830/B858-A831/B859, GMS 4:403-404, TP 8:288, HN 20:44.
[5] See TP 8:284-287, MS 6:480-81, KpV 5:36.
[6] In his explanation of 'casuistry,' Kant says that imperfect duties fall into a casuistry because they are indeterminate; thus, they require 'subsidiary maxims,' or maxims about maxims, in order to specify more precisely what one should do. Thus, James claims that, really, all broad duties require further subsidiary maxims, and thus all broad duties can lead to casuistry. However, not all casuistical questions can be said to concern situations in which we need maxims for maxims –

Similarly, Yvonne Unna has suggested that perfect duties may be included in Kant's casuistical questions because they prohibit particular actions but do not prescribe particular actions,[7] and may thereby still allow for some latitude in terms of which actions the agent should perform – an aspect shared with imperfect duties.[8] Consequently, on Unna's reading too, Kant was simply confused in excluding imperfect duties from casuistry in his explanation of casuistry in the introduction to the *Doctrine of Virtue*.[9] Tatiana Patrone (following James's suggestion just mentioned) even dismisses the seeming inconsistency by saying that "Kant's take on his own position with respect to casuistry is misleading."[10]

To my knowledge, Rudolf Schüssler has so far been the only interpreter who has made a serious effort at taking Kant by his word(ing). On Schüssler's reading, Kant's casuistical questions are not instances of Kantian 'casuistry.' In large part, this interpretation fits well with the text of the *Metaphysics of Morals*: throughout the *Doctrine of Virtue*, Kant is consistent in never calling his casuistical questions instances of casuistry. Distinguishing between casuistry (which is the search for 'subsidiary maxims' in the case of imperfect duties, which allow significant latitude) and casuistical questions helps explain why there can be casuistical questions regarding perfect duties. Kant's explanation of casuistry, which explicitly rejects the possibility of casuistry regarding perfect duties, simply does not apply to casuistical questions.[11]

---

which James admits (James 1992). So, even if we say that Kant misspoke and meant to say 'broad' duties instead of 'imperfect,' that would still not make sense of all the casuistical questions, as James himself admits.

[7] Which, of course, Kant's ethics does not do, for Kant's ethics is about both maxims and ends, not action-types; specific action-types are the matter of the *Doctrine of Right*.

[8] Unna 2003.

[9] Unna's implicit response to the first puzzle becomes more perplexing in light of Unna's final argument: she demonstrates that Kant in fact answers four of the five casuistical questions on suicide negatively in his lectures on ethics without reference to subsidiary maxims, so Kant obviously discussed casuistical questions that do not fit her own explanation for why some perfect duties can lead to casuistry.

[10] Patrone 2013 p. 490

[11] As Schüssler points out, this reading is further supported by the fact that Kant does not address subordinate maxims in the casuistical questions – even in the questions concerning imperfect (broad) duties to others (Schüssler 2012 pp. 84-

However, taking seriously this distinction between casuistry and casuistical questions now leaves us with the more general puzzle about Kant's casuistical questions: if the questions are not part of what Kant described as casuistry, what purpose are casuistical questions supposed to serve within Kant's practical philosophy? On Schüssler's proposed reading, Kant's casuistical questions are closed and merely rhetorical, intended to test the reader on whether they have really understood Kant's ethics. Although this answer has the considerable virtue of being theoretically consistent, it still has an awkward shortfall. Not only did Kant show a remarkable amount of sympathy (if not admiration) for e.g., some acts of suicide, like that of Jean-Marie Roland de la Platière during the Terror in Paris (Anth 7:259).[12] When Kant discusses the casuistical questions as a method of teaching children ethics, he explicitly claims that casuistical questions are genuinely difficult (MS 6:483). And more importantly, numerous Kant-scholars have found the casuistical questions difficult (or even impossible) to answer. Some authors have taken Kant's casuistical questions to be closed but difficult questions about cases of moral conflict – or, in Kant's own jargon, about conflicting grounds of obligation;[13] some have argued that only some questions are closed rhetorical ones while other ones are open;[14] and yet others have thought that the casuistical questions are about cases of moral conflict and genuinely open, and have thus taken Kant's casuistical questions in the *Doctrine of*

---

85). Indeed, the casuistical questions concerning imperfect duties to other do not even seem to invite questions about subordinate maxims in Kant's discussion. But that is exactly what we would have expected if the casuistical questions were indeed instances of casuistry. Further support for this reading can be found in Kant's writings on pedagogy. As Schüssler notes, there Kant talks about using casuistical questions *of right* in the education of children (Päd 9:490). Therefore, Kant clearly believed that casuistical questions can apply to the realm of right. Also here, Kant does not talk about 'casuistry,' but just about a catechism of right and questions.

[12] Schüssler 2012 p. 78 n26. See also Unna 2003 pp. 462-463.

[13] Esser 2008 and Koch 2003 pp. 121-188. Similarly, Unna 2003 and James 1992 have taken this position regarding the duty of self-preservation (Unna 2003), the questions on suicide (James 1992). Note that Unna and James come to different results regarding suicide: James thinks it can be permitted, Unna thinks Kant denied it ever being permitted.

[14] James 1992 has interpreted some of Kant's casuistical questions as rhetorical and closed, serving a pedagogical purpose of inducing students into liking the study of ethics. The remaining questions, on James's view, are open and concern maxims about maxims.

Elements as demonstrating that his moral philosophy is sensitive to ethical difficulty.[15] Unless we want to summarily dismiss these scholars as seriously confused, the very fact that so many senior Kant-scholars have found the casuistical questions difficult strongly suggests, at least to my mind, that they are unlikely to be merely rhetorical. After all, it would surely be awkward to suppose that Kant's casuistical questions test the reader on whether they have really understood the scope of the duties discussed, only for Kant scholars themselves to consistently fail this test.

Fortunately, I believe that neither interpretive option – that Kant would have been confused about his own discussion of casuistry, or that Kant scholars would be consistently confused about the substance of Kant's ethics – is necessary. On the alternative interpretation I offer below, we can differentiate between casuistry and casuistical questions as well as provide a more sympathetic answer to what purpose ties together the casuistical questions. Before moving on to this option, it will be helpful to briefly consider the historical context of Kant's discussion of casuistry in the *Metaphysics of Morals*.

Casuistry and the Purpose of Moral Philosophy

Kant wrote his casuistical questions during a time of popular disdain for the practice of casuistry. The analysis of cases of conscience had been a common practice already in scholastic philosophy, but the rise of the early modern practice of casuistry set in especially after the Council of

---

[15] O'Neill 2002 p. 343, Timmermann 2000 p. 49, Kim 2009, and Gregor 1963. Because of this reading, Kim (2009) has even suggested that Kant's *Metaphysics of Morals* presents a significant 'correction' or 'alteration' of his earlier moral theory from a formal to a more material conception of ethics in response to criticism; and that the casuistical questions, with their attention to the specifics of particular cases, are a result of this alleged change of mind (Kim 2009 pp. 335-338). In this view, Kim joins authors who have taken Kant's *Metaphysics of Morals* to be incompatible with Kant's *Groundwork* and the *Critique of Practical Reason*. Even more extreme, Mary Gregor claimed that Kant's casuistical questions concern the task of finding exceptions to general rules (Gregor 1963 pp. 141-142).

Trent (1545-1563), which had decided that sinners had to confess all the circumstances that might influence the nature of their sinful actions. Consequently, the second half of the 16th and the 17th century saw a surge of commentaries dealing with collections of various cases of conscience, taking into consideration all the possible, minute background circumstances of an individual's action.[16] Intimately connected with this practice of casuistry was the doctrine of probabilism, developed by Bartolomé de Medina in his 1577 commentary on Aquinas's *Summa Theologiae*. There, de Medina argued that in difficult ethical situations one could permissibly follow an opinion deemed less likely to be correct without committing a sin, if only that opinion had been stated or defended at some point by a serious authority with good arguments.[17] This theoretical background of probabilism gave rise to an, allegedly, widespread practice of justifying people's unconscientious actions by retrospectively finding some professional opinion that would justify it. Thus, the practice of casuistry soon acquired the reputation of attempting to justify lax morals, or even immoral conduct, by recourse to hair-splitting and factitious distinctions.[18] By the middle of the 17th century, casuistry had become both firmly identified with the doctrine of probabilism and been subjected to harsh criticism from rigorist thinkers. Partly due to prominent criticisms of authors like Pascal and Rousseau, among many others, the term 'casuistry' became almost synonymous with spurious moral reasoning

---

[16] As Di Giulio and Frigo note, the *Resolutiones Morales* in 9 volumes by Antonio Diana, written between 1629 and 1659, contain over 28,000 cases.

[17] Although it might sound odd to think that one should be morally permitted to follow a less 'probable' opinion in ethical difficulties, de Medina's original thesis is quite intuitive in its original context: Medina did not talk about the case of an individual contemplating their action, weighing different options and choosing the less probable opinion. (This would indeed sound irrational.) Rather, Medina considered the case of judging *another's* action. E.g., the case of a confessor who believed that a particular action was prohibited by a certain law, while the penitent in good faith, and in line with the opinion of other confessors or ecclesiastical authorities, had believed that the action was permissible. Although the confessor might find his own opinion more sound, Medina argued that he would not need to reprimand the penitent for following a probable opinion in good faith – even though another opinion might be more probable. For a historical discussion of this kind of probabilism and its relation to casuistry, see Jonsen and Toulmin 1988 pp. 164-175.

[18] Since the probabilism underlying the practice of casuistry was defended by important Jesuit authors, casuistry in the pejorative sense became identified primarily with the Society of Jesus.

by the mid to late 18th century.[19] (Although casuistry has since made something of a revival in philosophy through applied ethics – especially in medical ethics – the word still carries negative connotations today.[20])

Rousseau gives particularly memorable expression to this popular antipathy toward casuistry in his *Emile*, first published in 1762. In a section entitled 'The Creed of a Savoyard Priest,' he writes:

> […] I do not derive these rules [of conduct] from the principles of the higher philosophy, I find them in the depths of my heart, traced by nature in characters which nothing can efface. I need only consult myself with regard to which I wish to do; what I feel to be right is right, what I feel to be wrong is wrong; *conscience is the best casuist*; and it is only when we haggle with conscience that we have recourse to the subtleties of argument. […] Conscience is the voice of the soul, the passions are the voice of the body. Is it strange that these voices often contradict each other? And then to which should we give heed? Too often does reason deceive us; we have only too good a right to doubt her; but conscience never deceives us; she is the true guide of man; it is the soul what instinct is to the body; he who obeys his conscience is following nature and he need not fear that he will go astray.[21]

As this passage from *Emile* reminds us, Rousseau dismissed casuistry not merely because he was an epistemic egalitarian about moral knowledge; Rousseau also had a very pessimistic opinion of philosophy, which he believed could do more evil than good for people's moral compass. It is 'only when we haggle with conscience' that we start relying on philosophical arguments. In this way,

---

[19] For a concise overview of the history of probabilism and casuistry leading up to Kant, see Di Giulio and Frigo 2020 pp. 1-12. For a historical overview of the practice of casuistry, its demise and its revival, see especially Jonsen and Toulmin 1988, Knebel 2000, Carraud 1996, Tutino 2018, and Schüssler 2019. For a detailed discussion of the demise of probabilism in German speaking countries in the mid 18th century see De Franceschi 2020. For very brief overviews of the historical background to Kant's views on casuistry see also Kittsteiner 1988 and Schüssler 2012 pp. 90-94. A particularly insightful discussion of Kant's complex views on probabilism can be found in Schüssler 2020. For an influential account of the history of Jesuit casuistry see Stäudlin 1808 (chapters 7-10) pp. 460-572, and for discussion of Pascal's criticism of the practice of casuistry see especially Jonsen and Toulmin 1988 pp. 231-249, Krailsheimer 1967 and Bell 1998.

[20] According to the Oxford English Dictionary, 'casuistry' is "often (and perhaps originally) applied to a quibbling or evasive way of dealing with difficult cases of duty; sophistry." (Online, accessed December 4th 2021)

[21] Rousseau 1961 pp. 249-250, emphasis added

Rousseau took the popular antipathy against 'casuistry' to an extreme, turning it against popular moral philosophy in general.

Kant famously followed Rousseau in his epistemic egalitarianism about moral knowledge;[22] but, as John Callanan has shown, Kant did not follow Rousseau in his more pessimistic belief that philosophy would corrupt the common moral knowledge that is already available to people in their uncorrupted, natural state.[23] While Rousseau believed that natural human feeling would be enough as a moral guide in human conduct, Kant believes that reason – and importantly, philosophy – still has a significant role to play.

According to Kant, the project of unearthing the supreme principle of practical reason is also a therapeutic exercise, bringing us closer to who we truly are, and helping us overcome the subconscious, corrosive parts of our finite, rational nature. Kant most clearly expresses this therapeutic conception of the practical purpose of moral philosophy in the *Groundwork*. There, Kant had claimed that beings like us are easily led astray by the 'natural dialectic,' which arises from the conflicting demands of our inclinations and the demands of morality, and which can lead people to rationalizing away the requirements of morality.

> The human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect - the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name happiness. Now reason issues its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with disregard and contempt for those claims, which are so impetuous and besides so apparently equitable (and refuse to be neutralized by any command). But from this there arises a natural dialectic, that is, a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them

---

[22] See also Kant's remarks at RGV 6:185.
[23] Callanan 2019

194

Rousseau took the popular antipathy against 'casuistry' to an extreme, turning it against popular moral philosophy in general.

Kant famously followed Rousseau in his epistemic egalitarianism about moral knowledge;[22] but, as John Callanan has shown, Kant did not follow Rousseau in his more pessimistic belief that philosophy would corrupt the common moral knowledge that is already available to people in their uncorrupted, natural state.[23] While Rousseau believed that natural human feeling would be enough as a moral guide in human conduct, Kant believes that reason – and importantly, philosophy – still has a significant role to play.

According to Kant, the project of unearthing the supreme principle of practical reason is also a therapeutic exercise, bringing us closer to who we truly are, and helping us overcome the subconscious, corrosive parts of our finite, rational nature. Kant most clearly expresses this therapeutic conception of the practical purpose of moral philosophy in the *Groundwork*. There, Kant had claimed that beings like us are easily led astray by the 'natural dialectic,' which arises from the conflicting demands of our inclinations and the demands of morality, and which can lead people to rationalizing away the requirements of morality.

> The human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect - the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name happiness. Now reason issues its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with disregard and contempt for those claims, which are so impetuous and besides so apparently equitable (and refuse to be neutralized by any command). But from this there arises a natural dialectic, that is, a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them

---

[22] See also Kant's remarks at RGV 6:185.
[23] Callanan 2019

better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity - something that even common practical reason cannot, in the end, call good. In this way common human reason is impelled […] to go out of its sphere and to take a step into the field of practical philosophy […]. (GMS 4:405)[24]

Kant's account of this natural dialectic is not simply a conflict between inclination and reason; rather, this natural dialectic is itself a rational process. As finite beings, we necessarily take an interest in our happiness. And our happiness is a genuinely valuable end – which makes pursuing our happiness also reasonably appears so 'equitable.' And since it is hardly an option for finite, dependent beings like us to shape our inclinations at will, we take a pathological interest in justifying them. In other words, we are prone to rationalizing in favor of our inclinations and against our moral judgment not simply because we are suffering from overly strong inclinations, but precisely because we are rational beings who take an interest in everything that is reasonable (as our happiness is).[25]

Moral philosophy can help us overcome this corrosive aspect of our rational, finite nature because it excavates and brings closer to our feeling the supreme principle of practical reason, thus bringing us closer to who we truly are. We need the resolve to trust our capacity for pure practical reason and pursue what is intrinsically good even when it may conflict with our happiness. Kant believed that philosophy could help people find this resolve to rely on their rational capacities and to trust their immediate moral judgments, instead of rationalizing away the requirements of morality and relegating their decisions to inclinations. Moreover, while bad moral theory risks reinforcing our rationalizations, moral philosophy done well deprives us of such tools and may even help us at

---

[24] As John Callanan (2019) has compellingly argued, this passage at the beginning of the *Groundwork* is a response to Rousseau.
[25] For related discussion of Kant's notion of 'rationalizing' (vernünfteln) see also Sticker 2021, especially pp. 29-30.

catching ourselves out.[26] In this vein, Ido Geiger has argued that Kant's formulas of the Categorical Imperative in the *Groundwork* are best understood as "provid[ing] a heuristic tool for moral self-criticism and so for fighting temptation."[27]

Thus, while Kant agrees with Rousseau about the accuracy of common moral knowledge, he did not believe that we should therefore either give up or turn against moral philosophy. On Kant's view, moral philosophy can be practical because it is therapeutic. Both the highest principle of practical reason as well as its main obstacles are already within us, simply because we are finite, rational beings. Thus, the project of unearthing the supreme principle of practical reason which allows us to represent actions and ends as intrinsically good, and to demonstrate the possibility of pure reason being practical, is not a merely metaethical or speculative exercise: for Kant, it is also a therapeutic exercise, bringing us closer to who we truly, and helping us overcome the subconscious, corrosive parts of our finite, rational nature.

With this approach, Kant is in venerable company: "Empty is that philosopher's argument," Epicurus tells us in a fragment, "by which no human suffering is therapeutically treated. For just as there is no use in a medical art that does not cast out the sickness of bodies, so too there is no use in philosophy, unless it casts out the suffering of the soul."[28] On this conception of the practical purpose of moral philosophy, we are less in need of a 'method of ethics' and instead require resources for a therapeutic way of reflection: a kind of reflection that can help us overcome our self-incurred forms of unenlightenment or self-entrapment.

---

[26] For recent discussion of Kant's account rationalizing and self-deceit, see Papish 2018 and. Wehofsits 2020.
[27] Geiger 2010 p. 286. Similarly, Samuel Fleischacker has pointed out that: "Never does Kant take as the paradigm moral question, "what should all people do?," or "what should so-and-so over there do?" The question is always, "what should *I* do?," or, more strictly, "how can I convince myself of what I ought to do?" For the agent in the examples is never really in doubt as to the nature of the appropriate moral laws […]" (Fleischacker 1991 p. 263).
[28] Us. 211, Porphyrius ad Marcellam 31 p. 209, 23 Nauck. Translation by Martha Nussbaum (2018 p. 102).

Because of their view of moral cognition as egalitarian (i.e., accessible to every human being

of a healthy mind and heart), Kant and Rousseau both rejected the idea of a practice of casuistry that

would allow people to quibble with their conscience by means of hair-splitting and factitious

distinctions. Thus, the Vigilantius's lecture notes record Kant as saying that probabilist casuistry is

the "practice of cheating or chicaning conscience through sophistry, insofar as one makes it

[conscience] appear mistaken." (V-MS/Vigil 27:620, translation mine)[29] Indeed, for Kant, probabilist

casuistry is the perfect example of the human tendency to rationalize against the moral law.[30] But

unlike Rousseau, Kant also leaves room for 'casuistical questions.' What, then, could provide a more

promising starting point for explaining their different respective attitudes toward casuistry and

casuistical questions than their different views about the practical purpose of moral philosophy? As I

suggest in the remainder of this essay, this therapeutic purpose of Kant's moral philosophy –

through which he distinguished himself from Rousseau in the *Groundwork* – allows us to trace an

alternative reading of Kant's casuistical questions.

Kant's Questions

Kant adds eight small sections of 'casuistical questions' to some (albeit not all) discussions of

specific ethical duties and corresponding vices. For ease of discussion, it is worth restating Kant's

---

[29] See also V-MS/Vigil 27:557. On the orthodox reading of Kant's late conception of 'conscience,' conscience cannot be mistaken or false. It can only be ignored. Because conscience is not the practical syllogism itself, or the subsumption of cases under rules but the second-order cognition of whether one's actions are in conformity with one's practical judgment. Although one's practical judgment can be mistaken, one's conscience (i.e., knowledge of my actions' conformity to my moral judgment) can only be ignored. Cf Sticker 2020.

[30] Thus, pace Kittsteiner (1988 p. 193), Kant's casuistical questions in the *Doctrine of Virtue* are not "the traditional casuistry which lived on as an appendix to the doctrine of virtue." Rather, as some have programmatically put it, for Kant the 17th century probabilist casuist is the personification of radical evil. (Di Giulio and Frigo 2020 p. 11, compare also Di Giulio 2020 pp. 268-278.)

casuistical questions with reference to the specific duties/section titles to which they are added. (Readers already familiar with them may skip the following four pages.)

1. *On killing oneself*

    -   Is it murdering oneself to hurl oneself to certain death (like Curtius) in order to save one's country? – or is deliberate martyrdom, sacrificing oneself for the good of all humanity, also to be considered an act of heroism?

    -   Is it permitted to anticipate by killing oneself the unjust death sentence of one's ruler - even if the ruler permits this (as did Nero with Seneca)?

    -   Can a great king who died recently be charged with a criminal intention for carrying a fast-acting poison with him, presumably so that if he were captured when he led his troops into battle he could not be forced to agree to conditions of ransom harmful to his state? – for one can ascribe this purpose to him without having to presume that mere pride lay behind it.

    -   A man who had been bitten by a mad dog already felt hydrophobia coming on. He explained, in a letter he left, that, since as far as he knew the disease was incurable, he was taking his life lest he harm others as well in his madness (the onset of which he already felt). Did he do wrong?

    -   Anyone who decides to be vaccinated against smallpox puts his life in danger, even though he does it in order to preserve his life; and, insofar as he himself brings on the disease that endangers his life, he is in a far more doubtful situation, as far as the law of duty is concerned, than is the sailor, who at least does not arouse the storm to which he entrusts himself. Is smallpox inoculation, then, permitted? (MS 6:423-4)

2. *On defiling oneself by lust*

    -   Nature's end in the cohabitation of the sexes is procreation, that is, the preservation of the species. Hence one may not, at least, act contrary to that end. But is it permitted to engage in this practice (even within marriage) without taking this end into consideration?

    -   If, for example, the wife is pregnant or sterile (because of age or sickness), or if she feels no desire for intercourse, is it not contrary to nature's end, and so also contrary to one's duty to oneself, for one or the other of them, to make use of their sexual attributes – just as in unnatural lust?

- Or is there, in this case, a permissive law of morally practical reason, which in the collision of its determining grounds makes permitted something that is in itself not permitted (indulgently, as it were), in order to prevent a still greater violation?

- At what point can the limitation of a wide obligation be ascribed to purism (a pedantry regarding the fulfillment of duty, as far as the wideness of the obligation is concerned), and the animal inclinations be allowed a latitude, at the risk of forsaking the law of reason? (MS 6:426)

3. *On stupefying oneself by the excessive use of food or drink*

- Can one at least justify, if not eulogize, a use of wine bordering on intoxication, since it enlivens the company's conversation and in so doing makes them speak more freely?

- Or can it even be granted the merit of promoting what Horace praises in Cato: virtus eius incaluit mero?

- The use of opium and spirits for enjoyment is closer to being a base act than the use of wine, since they make the user silent, reticent and withdrawn by the dreamy euphoria they induce. They are therefore permitted only as medicines. – But who can determine the measure for someone who is quite ready to pass into a state in which he no longer has clear eyes for measuring?

- Although a banquet is a formal invitation to excess in both food and drink, there is still something in it that aims at a moral end, beyond mere physical well-being: it brings a number of people together for a long time to converse with one another. And yet the very number of guests […] allows for only a little conversation (with those sitting next to one); and so the arrangement is at variance with that end […] How far does one's moral authorization to accept these invitations to intemperance extend? (MS 6:428)

4. *On lying*

- Can an untruth from mere politeness (e.g., the "your obedient servant" at the end of a letter) be considered a lie? No one is deceived by it.

- An author asks one of his readers "How do you like my work?" […] The author will take the slightest hesitation in answering as an insult. May one, then, say what is expected of one?

- If I say something untrue in more serious matters, having to do with what is mine or yours, must I answer for all the consequences it might have?

- For example, a householder has ordered his servant to say "not at home" if a certain human being asks for him. The servant does this and, as a result, the master slips away and commits a serious crime, which would otherwise have been prevented by the guard sent to arrest him. Who (in accordance with ethical principles) is guilty in this case? Surely the servant, too, who violated a duty to himself by his lie, the results of which his own conscience imputes to him. (MS 6:431)

5. *On avarice*

   - [M]iserliness is not just mistaken thrift, but rather slavish subjection of oneself to the goods that contribute to happiness which is a violation of duty to oneself since one ought to be their master. It is opposed to liberality of mind (liberalitas moralis) generally (not to generosity, liberalitas sumptuosa, which is only an application of this to a special case), that is, opposed to the principle of independence from everything except the law, and is a way in which the subject defrauds himself. But what kind of a law is it that the internal lawgiver itself does not know how to apply?

   - Ought I to economize on food or only in my expenditures on external things? in old age, or already in youth? Or is thrift as such a virtue? (MS 6:433-4)

6. *On servility*

   - Is not the human being's feeling for his sublime vocation, that is, his elation of spirit (elatio animi) or esteem for himself, so closely akin to self-conceit (arrogantia), the very opposite of true humility (humilitas moralis), that it would be advisable to cultivate humility even in comparing ourselves with other human beings, and not only with the law?

   - Or would not this kind of self-abnegation instead strengthen others' verdict on us to the point of despising our person, so that it would be contrary to our duty (of respect) to ourselves? Bowing and scraping before a human being seems in any case to be unworthy of a human being.

   - Preferential tributes of respect in words and manners even to those who have no civil authority – […] does not all this prove that there is a widespread propensity to servility in human beings? (MS 6:437)

7. *The duty of beneficence*

- How far should one expend one's resources in practicing beneficence? Surely not to the extent that he himself would finally come to need the beneficence of others.

- How much worth has beneficence extended with a cold hand (by a will to be put into effect at one's death)?

- If someone who exercises over another (a serf of his estate) the greater power permitted by the law of the land robs the other of his freedom to make himself happy in accordance with his own choices, can he, I say, consider himself the other's benefactor because he looks after him paternalistically in accordance with his own concepts of happiness? Or is not the injustice of depriving someone of his freedom something so contrary to duty of right as such that one who willingly consents to submit to this condition, counting on his master's beneficence, commits the greatest rejection of his own humanity, and that the master's utmost concern for him would not really be beneficence at all? Or could the merit of such beneficence be so great as to outweigh the right of human beings? – I cannot do good to anyone in accordance with my concepts of happiness (except to young children and the insane), thinking to benefit him by forcing a gift upon him; rather, I can benefit him only in accordance with his concepts of happiness.

- Having the resources to practice such beneficence as depends on the goods of fortune is, for the most part, a result of certain human beings being favored through the injustice of the government, which introduces an inequality of wealth that makes others need their beneficence. Under such circumstances, does a rich man's help to the needy, on which he so readily prides himself as something meritorious, really deserve to be called beneficence at all? (MS 6:454)

8. *The duty of sympathy*

- Would it not be better for the well-being of the world generally if human morality were limited to duties of right, fulfilled with the utmost conscientiousness, and benevolence were considered morally indifferent? It is not so easy to see what effect this would have on human happiness. But at least a great moral adornment, benevolence, would then be missing from the world. This is, accordingly, required by itself, in order to present the world as a beautiful moral whole in its full perfection, even if no account is taken of advantages (of happiness).

- [U]niversal love of one's neighbor can and must be based on equality of duties, whereas in gratitude the one put under obligation stands a step lower than his benefactor. Is it not this,

namely pride, that causes so much ingratitude? seeing someone above oneself and feeling resentment at not being able to make oneself fully his equal (as far as relations of duty are concerned)? (MS 6:458)

Kant's casuistical questions are remarkably diverse. The first six duties to which he adds casuistical questions are perfect duties to oneself, while the last two duties are imperfect duties to others. On first sight, some of Kant's casuistical questions seem to pose genuinely difficult, if not unanswerable questions about conflicting grounds of obligation, like some of the questions concerning suicide (1). By contrast, some casuistical questions seem to concern conflicting grounds of obligation but also appear to be answerable, like the questions concerning lying (4). The questions regarding lust (2) can seem difficult to answer from a Kantian perspective, but they do not obviously concern any conflicting grounds of obligation. Still other questions appear to be both answerable and also do not obviously concern conflicting grounds of obligations, like the questions regarding the use of food and drink (3) and the questions about servility (6). In addition to these differences, some of Kant's casuistical questions regarding the duty of beneficence (7) seem clearly answerable, while other questions regarding the same duty seem both difficult and even concern conceptual rather than practical difficulties. And some entire sets of casuistical questions do not appear to raise any specific practical questions at all but merely theoretical ones, like the questions regarding sympathy (8). Lastly, the questions on avarice (5) may sound as if they were about an imperfect duty although they explicitly concern a perfect duty toward oneself.

Although Kant's questions regarding suicide present genuinely difficult situations of conflicting grounds of obligation, Kant himself answers at least three out of the five questions on

suicide in his own writings.[31] Similarly, Kant also provides an immediate answer to his third

casuistical question on beneficence – whether one may deprive others of their freedom of choice in

order to paternalistically impose one's own conception of happiness on them: "I can benefit him

only in accordance with *his* concepts of happiness." And Kant also provides at least part of an

answer to his first casuistical question on beneficence – how much of one's resources one should

spend for beneficence: "Surely not to the extent that he himself would finally come to need the

beneficence of others."

Finally, Kant's casuistical questions do not seem to be examples of 'casuistry,' i.e., cases in

which we need to find 'subsidiary' maxims for the application of other, more general, maxims. In his

introduction to the *Metaphysics of Morals*, Kant had stated that ethics, "because of the latitude it allows

in its imperfect duties, unavoidably leads to questions that call upon judgment to decide how a

maxim is to be applied in particular cases, and indeed in such a way that judgment provides another

(subordinate) maxim (and one can always ask for yet another principle for applying this maxim to

cases that may arise). So ethics falls into a casuistry […]." (MS 6:411-12) While some of Kant's

casuistical questions might be read in a way that makes them fit this picture – like, for instance, the

questions concerning how much of one's resources should be spent in beneficence, or how much

alcohol one can consume without stupefying oneself through excess – most of Kant's casuistical

questions do not fit this explanation of casuistry. Interestingly, the casuistical questions which we

would have expected to fit this paradigm most of all – namely the questions concerning *imperfect*

duties – do not fit this explanation at all. Thus, Kant's casuistical questions are most likely no

---

[31] Anth 7:259, Refl 15:972-976, and V-MS/Vigil 27:603 and 27:629 respectively. For discussion see Unna 2003. As Unna has pointed out, the questions on suicide also perspicuously involve different conflicting motivations: the welfare of others, honor, and the preservation of life.

'casuistry' – indeed, Kant is very consistent in his use of these two terms for apparently different practices.[32]

After this brief, initial look at Kant's casuistical questions, we can already make the following preliminary observation. Kant's casuistical questions have little, if anything, in common: neither the type of difficulty involved, nor the answerability of the question, nor even the nature of the question (whether practical or theoretical) remains consistent throughout. Consequently, it would be misleading to describe them as instances of rhetorical questions that test the reader's understanding of Kant's ethics (as some of them remain genuinely difficult even to Kant scholars), or as instances of conflicting grounds of obligation (as not all are about moral conflicts) or uncertain cases that require 'subsidiary' maxims for their solution (since only few might require subsidiary maxims), or examples of simply unanswerable ethical questions (as some of them can be answered).

One thing, however, that Kant's casuistical questions share, is their sensitivity to ethical difficulties.

Highlighting the Complexity of Moral Life

One likely reason why Kant's casuistical questions have appeared so strange to commentators is that most recent readers have also assumed a specific view about the purpose of Kant's moral philosophy. On this commonly assumed view, Kant's ethics is a theory of right action – that is, a method of ethics (to use Sidgwick's term) which will provide a test or decision-procedure

---

[32] For discussion, see Schüssler 2012.

and thus help us resolve difficult questions about what we should do.[33] On this view, it is reasonable to expect that the casuistical questions are exercise-examples (whether easy or tricky) which the theory's test for right action will allow us to answer. However, as we saw above, this way of reading the casuistical questions forces a uniformity on them that the casuistical questions simply do not have. And importantly, this way of reading the casuistical questions also blinds out the questions themselves: on such a view, their purpose would merely lie in their answer (whether open or closed) because they are supposed to be exercise-examples for a test.

By contrast, I suggest that we see the significance of the casuistical questions not in their potential answers but in the questions themselves and in the difficulties they highlight. If moral philosophy is to cast out suffering of the soul, it must first be sensitive to the kinds of suffering beings like us face. And arguably, Kant's casuistical questions demonstrate just that: the complexity of moral life does not merely consist in agents puzzling about what they should do.

One such human difficulty that plays a central role in Kant's casuistical questions is the natural dialectic: the human propensity to place their inclinations over their knowledge of what is moral, and to rationalize away the requirements of morality because of the demands of our inclinations "which are so impetuous and besides so apparently equitable." (GMS 4:405) Kant had already articulated this difficulty for finite, dependent beings like us in the *Groundwork*, and he picks it up again in the casuistical questions. For instance, Kant's first three questions concerning lust are paradigmatic examples of this difficulty. The first and last question concerning the excessive use of drink also deal with this propensity of human nature to succumb to the natural dialectic. Similarly, the second question on lying involves a case in which one is tempted to tell a falsehood in order to

---

[33] For representative examples, see e.g., Korsgaard 1998 pp. xii, Rawls 2000 p. 163, Schneewind 1991 p. 289 and Sticker 2015 p. 980.

avoid inadvertently insulting an author, and thus to avoid an uncomfortable situation. One might also add the second question regarding suicide to this category. Moreover, these questions highlight not merely the difficulty of following through on a moral demand in the light of contravening inclinations; they also highlight that the demands of our inclinations are not merely "impetuous" but represented to us as "so apparently equitable." After all, is the pleasure of the kind of lively conversation that only happens after a couple of drinks really not a perfectly reasonable enjoyment? And aren't many forms of sexual pleasure genuinely good even though they do not take 'nature's end' into account? On Kant's view, the inclinations that tempt us to these options are not merely strong, they also appear so reasonable because there is nothing in itself wrong with pleasure. This makes it all the more tempting to rationalize away the arduous requirements of morality in favor of our inclinations.

Another version of this difficulty comes up in Kant's last question regarding suicide. There, he asks if smallpox inoculation is morally permissible.[34] Although this question is sometimes read as a moral conflict insofar as the obligation in question seems to pull in two different ways, it cannot be a conflict between different grounds of obligation. For the only obligation under consideration here is to preserve one's life. The casuistical question asks whether one may pursue this moral end by taking the comparatively small risk of inoculation with a mild variant of smallpox, or whether one should instead avoid such risks and hope not to contract smallpox at all. Thus, on first sight the difficulty in question concerns the appropriateness of specific actions in the pursuit of one and the

---

[34] Kant's opposition to smallpox inoculation might seem bewildering in a time when ever more cranky conspiracy theories are trying to outdo each other in their fight against safe vaccines. However, smallpox inoculation during the 18th century was still a live attenuated vaccine. Although the likelihood of dying from inoculation with a mild version of the live virus was significantly lower than from uncontrolled contraction of smallpox (20-40%), the 18th century smallpox inoculation still had the drawback of developing into the full-blown disease in some cases and killing the recipient. (In addition, inoculation with the live virus also made the recipient an active transmitter of the disease – with potentially disastrous public-health consequences.) For a brief history of the smallpox vaccines see Kotar and Gessler 2013.

same moral end. Kant answers this question in his unpublished notes on Anthropology, claiming that the maxim of a person who risks their life either willingly or by indifference can be accused of suicide.[35] However, in these notes Kant also points out that the conflict underlying this casuistical question is actually one between the 'prudence in consideration of one's advantage' and the requirements of morality not to endanger one's life. Kant's answer is that considerations of advantage (increasing one's chances of surviving smallpox) cannot outweigh requirements of morality (not to risk one's life). As one would expect from the rigorist philosopher, Kant thought that no matter how certain the gained advantage might be, means contrary to the moral law must not be used in pursing one's advantage. Thus, although this casuistical question initially seems to concern a moral conflict or the appropriateness of a particular action to a moral end (i.e., an obligation pulling us in two different directions), the underlying difficulty turns out to be the human propensity to place foreseen advantage over morality. Ironically, the very fact that the conflict appears, on first sight, to be moral underscores Kant's claim that the natural dialectic makes us rationalize away the requirements of morality in light of the demands of our inclinations.

Another difficulty thematized in Kant's casuistical questions is our frailty. In his third question on suicide, Kant refers to the belief that Frederick II of Prussia had carried with him a fast-acting poison, as Kant remarks "presumably so that if he were captured when he led his troops into battle he could not be forced to agree to conditions of ransom harmful to his state." As so many others, this casuistical question is not about conflicting grounds of obligation. From a Kantian perspective, the duties at stake appear to be to (i) not kill oneself, and (ii) not to agree to anything harmful to the state and citizens. But these grounds of obligations are not themselves contradictory.

---

[35] Note besides that even in his discussion of this casuistical question in his notes Kant does not refer to the Categorical Imperative in any formulation but simply claims that intentionally risking one's life is a case of suicide.

Rather, the problem arises because of human vulnerability and frailty. The two obligations in question only give rise to a conflict when we assume that human frailty would make it impossible to comply with the latter duty without ending one's life. Of course, only the most pathologically megalomaniac among us could seriously believe that they could withstand the pressure of psychological and/or physical torture. What actions, then, are morally permissible when we expect that harm might potentially befall others because of our frailty? We might not share Kant's seeming optimism about the capacity of rational beings to overcome their limitations by the force of pure reason; but even if we do not, Kant's question remains perfectly reasonable. We might excuse, understand, even admire the suicide of someone who fears that they would not withhold torture and think suicide the only realistic option for protecting others. But Kant's question perspicuously highlights the moral difficulty: is it merely to accept one's vulnerability and frailty if one does not even try to overcome it? Or is an intention that presumes we would not be able to overcome it the same as to renounce our ability to do the right thing altogether?

A further difficulty concerns self-conceit. In his first two question on servility, Kant points out that self-esteem is dangerously similar to arrogance. In almost proto-Freudian fashion, Kant here suggests that one easily deceives oneself about the actual nature of one's feeling. How, then, should we deal with such a propensity for self-conceit? Kant first asks if it would not be advisable for humans to cultivate humility by comparing ourselves with others in order to counteract our propensity to portray our arrogance as self-esteem. But he then asks if such a habit of cultivating humility by comparing ourselves with others might not make others view us as inferior, and thus appear in conflict with our duty of self-respect. The options adumbrated by Kant in these two questions – comparing ourselves with others, and making others thereby despise us – are far less significant than the underlying problem. Indeed, once we take the problem seriously, we can see that

the intrinsically appropriate response to the problem is neither to cultivate humility by comparing ourselves with others, nor would such a habit of comparing ourselves with others be inappropriate because of how others may consequently choose to despise us. Rather, the appropriate response is, first of all, to take the difficulty seriously, and to cultivate a habit of reflecting especially on those feelings that are so prone to self-deception.

Some of Kant's casuistical questions also deal with the (seemingly more banal) difficulty of drawing a line. In his last question regarding lust, Kant asks at what point one may allow a latitude to our inclinations, and at what point compliance with a wide duty becomes pedantry. Similarly, in his second question on avarice and miserliness, Kant asks on which expenditures, and at what point in life, one ought to economize. And regarding beneficence, Kant asks how much of one's resources one should give to others. All of these questions concern the necessity of a determination in situations that do not seem to have a clear limit or border. As such, the difficulty in these questions might seem quite banal on first sight. But what is more interesting are the competing considerations. For what makes these cases difficult is not simply the indeterminacy of the situation but the conflicting goals in question. We want to give lust a latitude because of our sexual drive; we want to minimize beneficence because we desire our own comfort; we economize on expenditures for ourselves because we want to keep or accumulate wealth. Unlike some other casuistical questions, Kant neither answers them in the *Metaphysics of Morals* nor anywhere else, nor does he seem to provide the resources for answering them. Evidently, what matters for Kant in these questions is not a solution but an accurate awareness of the difficulty. In other words, what appears to matter in these casuistical questions is not simply the difficulty of drawing a line but a more complete awareness what makes the difficulty of drawing a line an ethical difficulty for human beings: the

tension between rational principles and conflicting inclinations that reason rightly represents as 'so equitable.'

A final practical difficulty concerns conflicting grounds of obligation. So far, this type of difficulty has received the most attention in the secondary literature. In two of the most famous casuistical questions, Kant asks if one may not sacrifice oneself for the good of all humanity (like in the legend of the Roman soldier Marcus Curtius) or to kill oneself at the onset of hydrophobia in order to protect others. Contrary to the questions discussed above, these two casuistical questions do not concern the natural dialectic, human frailty or our propensity for self-deceit. Instead, these two questions concern two equally applicable grounds of obligation: self-preservation and the well-being of others. Interestingly, Kant also answered these questions negatively in his lectures.[36] Although conflicting grounds of obligation only make for two casuistical questions, and although Kant eventually answers them in his lectures, they demonstrate that Kant was aware of the fact that conflicting grounds of obligation are a genuine moral difficulty – even where they can allegedly be resolved.

Several casuistical questions are not about practical matters at all but concern theoretical questions about moral obligation and political institutions. For instance, in the final two questions regarding lying, Kant asks if one can be held legally responsible for the consequences of a lie. Although Kant's example in the last casuistical question on lying concerns a lie by a servant who may feel contractually obliged to obey their employer's order, there can be little doubt that Kant did not think of this situation as one of moral conflict, or giving rise to any significant moral difficulty at

---

[36] We might not be surprised by this, since Kant took self-preservation to be a perfect duty, while the furthering the happiness and well-being of others is only an imperfect duty. For a discussion of Kant's answers to his questions regarding suicide, see Unna 2003. Note again that in his discussion of these questions in his lectures, Kant does not answer them by applying the Categorical Imperative as a test.

all. Instead, Kant's focus is on the relation between the immorality of lying, and the permissibility of holding people legally accountable for their lie. Moving even further away from directly ethical concerns, several casuistical questions concern the correct understanding of abstract concepts, as well as the correct classification of specific types of actions and attitudes. For instance, in his brief discussion of avarice Kant asks in what sense we should understand the term 'law' in our talk about the moral law, given that the 'internal lawgiver' themselves can have difficulties in applying the law. In his discussion of servility, Kant asks if empirical evidence does suggest that servility might be a widespread, and presumably unconscious, propensity in human beings. Kant also asks if testamentary bequests can be classified as beneficence, and whether the philanthropic endeavors of rich people in a deeply unjust society can be considered beneficence at all. Finally, Kant asks if pride might be the unconscious cause of ingratitude.

So far, commentators have often tended to ignore Kant's theoretically-oriented casuistical questions. However, I believe these questions are just as insightful as Kant's practical questions. After all, Kant's moral philosophy seeks to be therapeutic by unearthing the rational principles that underlie our moral cognition and their systematic interconnection. Consequently, we might expect that Kant's casuistical questions highlight not merely the difficulties that his moral philosophy should help us to resolve, but also some of the difficulties in producing such a philosophy.

Conclusion

Kant's casuistical questions demonstrate the complexity of moral life. This complexity is not exhausted by puzzled agents asking themselves what they should do. It also involves the self-deceived agent, the frail agent, those struggling with the natural dialectic, the morally unlucky who

cannot act without going against some ground of obligation, and those struggling to find the right theoretical understanding of their practical situation. If moral philosophy is to cast out suffering of the soul, it must first be sensitive to the specific difficulties beings like us face. Kant's casuistical questions highlight just those difficulties.

Since many of Kant's casuistical questions emphasize the natural dialectic as a cause for moral difficulty they also forge a direct bridge between the first and final pages of Kant's moral philosophy. In section I of the *Groundwork*, Kant had remarked that "[t]he human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect [...]." And therefore, practical wisdom "needs science, not in order to learn from it but in order to provide access and durability for its precepts." (GMS 4:405) With this view about the practical purpose of moral philosophy, Kant distanced himself significantly from thinkers like Rousseau, who shared Kant's epistemic egalitarianism about moral cognition and his rejection of baroque casuistry, but who also had a much more pejorative view about the role of scientific and philosophical reasoning for ethical life.[37] Kant's project of a metaphysics of morals is an attempt to analyze the a priori principles that underlie our common moral cognition and their systematic interconnection. Such an analysis should provide us with the resources for a therapeutic way of reflecting, a reflection that brings us closer to our true nature as rational beings. But in order to do this, Kantian moral philosophy must be sensitive to the specifically human difficulties we face in our ethical life. Only if we take these diverse difficulties seriously can moral philosophy get closer to an improvement of our situation.

---

[37] For a recent discussion of the relation between Rousseau's views about moral philosophy and Kant's remarks about the natural dialectic in the *Groundwork*, see Callanan 2019.

PART III

WHY MORE THAN KANTIAN THERAPY?

Chapter 8

Kant's Racism as a Philosophical Problem

Introduction

Immanuel Kant was possibly both the most influential racist as well as the most influential moral philosopher in the history of modern, western thought. On the one hand, Kant's work and handwritten remains contain outrageously racist remarks. Moreover, Kant was arguably the first European thinker to produce an entire theory of race.[1] On the other hand, Kant's conception of all rational beings as "ends in themselves" is widely considered a paradigm of moral egalitarianism. Even more, his insistence on the inviolable dignity of each person and his ideas for "cosmopolitan" rights continue to be a major influence on contemporary moral and political philosophy. This stark contrast between Kant's racist remarks and his moral egalitarianism has inspired a long-overdue debate about the problem of Kant's racism: How could an ardent advocate of the universal dignity of all human beings simultaneously hold such despicable views about non-White people? And does the combination of these two facts indicate a failure of Kant's moral philosophy itself? Or does it merely indicate a failure of the person Immanuel Kant?

---

[1] For recent discussions of Kant's theory of race, see Bernasconi (2001, 2002, 2006, 2010), Boxill (2017), Eberl (2019), Eze (1995), Herb (2018), Lagier (2004), Larrimore (1999, 2006), Leutgöb (2015), Malter (1990), McCarthy (2009, chapters 2 and 5), Mikkelsen (2013), Sandford (2018), Shell (2006) and Zammito (2006). For interpretations that highlight the particular historical context of Kant's theory of race, see especially Eberl (2019) and Sandford (2018) who compellingly argue that Kant's theory of race must be understood in the historical context of his rejection of popular accounts of polygenism and the prevailing discourse on "barbarism." See also Geier (2022), who equally emphasizes the ongoing debates about "race" in the 1770s and consequently defends a (comparatively) less negative view of Kant. For a recent collection dealing with the emergence of the idea of race in 18th century Germany more generally, see Eigen & Larrimore (2006). For a discussion of Kant's racism in relation to his philosophy of history and teleology, see Boxill (2017) and Sutter (1989, pp. 258-259). Finally, for recent literature dealing more specifically with Kant and colonialism, see especially Eberl (2019), Hedrick (2008), Williams (2014), and Flikschuh and Ypi (2014).

After a long period of almost complete complacency about Kant's racism in European and Anglo-American academia, recent scholarship has started to take Kant's racism seriously and profoundly deepened our understanding of both the historical development of Kant's views on race and the nature of his racism. Within the emerging literature on Kant's racism, scholars still virtually unanimously agree over framing the problem as choice between two interpretations of Kant's moral philosophy: either Kant was an *inconsistent egalitarian*, or he was a *consistent inegalitarian*.[2] Undoubtedly the majority of Kant scholars believe that Kant was an inconsistent egalitarian. On their view, Kant was simply inconsistent in that he failed to draw the necessary conclusions from his own moral philosophy.[3] By contrast, some scholars have argued that Kant was a *consistent inegalitarian*. On their view, Kant's racist remarks are not incompatible with his moral philosophy, because when Kant wrote about the dignity of all persons as ends in themselves, he did not mean to include non-White people.[4]

Both sides of the debate agree that Kant's moral philosophy and his racist beliefs are not merely in tension but in obvious contradiction.[5] However, this assumption faces significant textual difficulties. Of the two most intuitive candidates for a possible contradiction, neither are as obvious

---

[2] The first scholar to characterize the debate in these terms was Pauline Kleingeld (2007), but her characterization also fits contributions to the debate that do not use this explicit wording.

[3] Versions of this view have most recently been defended by Allais (2016), Bernasconi (2011), Hill & Boxill (2001), Kleingeld (2007, 2014, 2019), Kaufmann (2019), McCabe (2019), Mensch (2017) and Terra (2013). For an interpretation of Kant's theory of race on which even Kant's use of teleological principles in anthropology contains significant anti-racist elements, see Malter (1990) and Dörflinger (2001).

[4] For recent defenses of this view, see Eze (1995), Serequeberhan (1996) and Mills (2005, 2014). For a similar, but somewhat less radical, argument that Kant's view was less universalist than usually assumed, see Larrimore (1999), who argues that Kant had a "two-stage view of raciation insulated form history and ethics," but "should the international federation of republics he looked forward to ever be achieved, the members of (non-white) races would have no place." (ibid. p. 125). Similarly, Serequeberhan also argues that Kant's notion of enlightenment and historical progress excludes non-White people and thus manifests an inegalitarian "metaphysics" under the cover of universalism. Cf. Hedrick (2008).

[5] See, for example, Mills (2014, p. 22) (speaking about a "contradiction so flagrant"), Allais (2016, pp. 1, 8) ("startling contrast" and "obvious contradictions"), McCabe (2019, p. 7) ("gross incompatibility"), Kleingeld (2007, p. 584) ("genuine contradiction").

as commentators have suggested. First, Kant's outrageous remarks about non-White people in his lesser-known writings never go as far as to literally deny their moral personhood. While his texts evince both bigoted prejudice and the belief that people of different "races" would have different psychological and physical characteristics that could be ranked, Kant's comments are never in outright contradiction with his view of each human being as an end in itself. Second, Kant explicitly denied that equal legal and political rights would follow from his conception of equal moral status: In the *Doctrine of Right*, Kant both explicitly affirmed the equal moral status of women while also denying them equal legal and political rights. If there really is any genuine contradiction between Kant's anthropological racism and his abstract egalitarianism, this contradiction is much less obvious than is usually assumed – if it exists at all.

Fortunately, these two interpretive options are not exhaustive, though the correct answer is much more worrisome for Kantian ethics than the two options presented so far. On my view, the verdict about Kant's racism and its relevance for his moral philosophy must be far more cynical than the two dominant interpretive options in the literature allow. Kant was arguably a consistent formal egalitarian: Kant's characterization of the equal dignity of all persons is so abstract that it remains largely useless as an antidote to racism.[6] Consequently, we have to seriously question if Kant's

---

[6] To my knowledge, the only authors to openly (albeit passingly) acknowledge this possibility are Hedrick (2008), Basevich (2020) and Lu-Adler (2022a). I discuss their respective views in more detail below. Although my argument is in a similar spirit as the famous "empty-formalism" charges by Hegel and Mill – and one may read this paper as a Hegelian-inspired formalism worry – my primary targets in this paper are contemporary readers of Kant and their largely unexamined assumption that Kant's abstract moral egalitarianism and his racism are in obvious contradiction. On my reading, Hegel's empty-formalism charge is directed against any attempt to develop a moral philosophy through a conceptual separation of form from content; by contrast, my argument here is significantly less ambitious, claiming merely that Kant's moral egalitarianism is too formal (or "thin") to clearly rule out his deeply racist views. Thus, my argument also does not rely on some of the topics usually discussed in the context of the empty-formalism charge (like maxim-description or the alleged universalization test). My argument also differs structurally in crucial ways from Hegel's classic "empty-formalism" charge: the latter accuses Kant of not being entitled to otherwise acceptable conclusions, while I argue that Kant's abstract egalitarianism appears compatible with completely unacceptable conclusions.

language of universal human dignity really is the powerful tool against racism and misogyny that it is frequently said to be.

In the course of treating Kant's views as contradictory, most commentators still deny that Kant's racism demonstrates a *philosophical* problem – that is, a problem truly pertaining to the kind of moral philosophy Kant was committed to doing. Those who argue that Kant was an inconsistent egalitarian inadvertently portray the issue of his racism as a personal failure of Immanuel Kant but of no larger significance for his moral philosophy; on their view, the problem of Kant's racism is simply that Kant himself was capable of moral failure and cognitive dissonance, but his personal cognitive dissonance does not demonstrate any shortfall of his moral philosophy. Those who argue that Kant was a consistent inegalitarian inadvertently also risk portraying the issue of Kant's racism as primarily a personal failure. On their view, the problem of Kant's racism is that he did not include enough individuals in the community of persons, but once we admit that really all human beings are persons the story of Kantian abstract egalitarianism allegedly has a happy ending. By contrast, this paper joins a minority of commentators in arguing that Kant's racism really is a philosophical problem: It neither demonstrates the failure to overcome cognitive dissonance nor the failure to recognize some people as fully human; rather, Kant's racism demonstrates a failure of Kant's moral philosophy. However deep his insights were in other regards, his egalitarianism remained so abstract as to be almost entirely useless as an antidote to his own racism. In this way, the seeming compatibility between Kant's abstract egalitarianism and his racism also highlights a perennial philosophical difficulty: How do we get from abstract principles to substantive ideals in a way that doesn't merely polish up our existing prejudices?

In section two, I briefly highlight the standard assumption that unifies the great majority of contemporary interpretations of Kant's racism by recourse to two opposing (and equally invaluable)

217

recent accounts. Section three argues that this common assumption faces textual difficulties by taking a closer look at the two most intuitively plausible candidates for a contradiction between Kant's account of moral equality and his racism. Finally, section four argues that this seeming compatibility between Kant's racism and his abstract egalitarianism is of wider philosophical significance.

The Standard Assumption

So far, the standard approach to Kant's racism has been based in the assumption that Kant's racist remarks are in obvious contradiction with his moral egalitarianism, and that Kant's language of universal human dignity offers a powerful tool against racism and misogyny. Consequently, a central question of recent scholarship and popular commentary is how we should make sense of the apparent contradiction behind Kant's racism.[7] According to the more critical voices within the standard reading, the best way of making sense of Kant's racism is to see him as a consistent inegalitarian: When Kant spoke of the dignity of all human beings as ends in themselves, he did not actually mean it. But according to the majority of voices within the standard approach, the best way of making sense of Kant's racism is to see him as an inconsistent egalitarian: When he made his outrageous remarks about non-White people, he was not being a good Kantian.

Over the last decades, the consistent-inegalitarian reading has prominently been defended by Emmanuel Eze and Charles Mills. While Eze's interpretation of Kant's racism (on which Kant

---

[7] This standard assumption of a contradiction between Kant's racist remarks and his moral egalitarianism also permeates popular culture. See for instance the recent contributions by Markus Willaschek, Bernd Dörflinger, Marina Martinez Mateo, among others, in the Frankfurter Allgemeine Zeitung; and Stefan Gosepath and Gabi Wuttke, Kant und die Rassismus-Debatte: „Die Vertreter der Aufklärung sind nicht unschuldig" Deutschlandfunk Kultur, June 16, 2020.

explicitly assigns different moral worth to different human beings) has not found many, if any, sympathetic readers,[8] Charles Mills has since powerfully argued that Kant was a consistent inegalitarian. Like Eze, Mills claims that the dignity of each person as an end in itself was only meant to apply to White people, and that Kant simply didn't consider non-White people to be persons. But unlike Eze, Mills claims that Kant had a silent taxonomy of moral worth, and thus implicitly distinguished not merely between persons (i.e., those with the capacity for practical reason) and things, but also between those human beings who are persons and those humans who are considered less than persons, or "sub-persons."

> The "sub-person" category is, admittedly, a reconstruction of the normative logic of racial and gender subordination in his [Kant's] thought, a reconstruction that is certainly not openly proclaimed in the articulation of his conceptual apparatus, and may seem, prima facie, to be excluded by it. [...] Nonetheless, I would claim that it is the best way of making sense of the actual (as against officially represented) logic of his writings, taken as a whole, and accommodates the sexist and racist declarations in a way less strained than the orthodox reading.[9]

Mills provides two further claims to support this interpretation of Kant's thought. First, Mills points out that words can assume different meanings depending on their context, and invites us to consider that Kant simply did not really mean to include non-White people in his category of persons. Thus, although the Formula of Humanity ("So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means." GMS 4:429) might sound egalitarian to us, Kant himself might not have meant it to include non-White people. Second, Mills has argued that ascribing such a hierarchy of persons and sub-

---

[8] Eze (1995) also claims to have shown that Kant's racist theory of race is "intimately" connected with the basic tenets of Kant's transcendental philosophy, and thereby compromises Kant's entire transcendental idealism. Specifically, Eze claims that on Kant's view "racial differences and racial classifications are based a priori on the reason of the natural scientist," and that for Kant the "classification of humans according to race [...] is a priori, transcendentally grounded and immutable." (Eze, 1995, pp. 122 and 124). For a discussion of Eze's argument, see Hill & Boxill (2001, pp. 452-459). Cf Kaufmann (2019, pp. 191-192).
[9] Mills (2005, p. 106).

persons to Kant's moral theory is, in fact, the only reasonably charitable interpretation of Kant's writings. In response to two authors who had previously defended the inconsistent egalitarian reading, Mills contends:

> How could it be more plausible to attribute to Kant the degree of cognitive dissonance requisite for the genuinely universalist reading of his work to be correct? I will put the contrast in the following stark form, to bring home (what I see as) the absurdity involved. [...]
> Unqualified Universalism: All biological humans/all races, as full persons, must be treated as ends, never as mere means.
>
> Racist Particularism: The races of Blacks and Native Americans may be colonized and enslaved. I submit to the reader that a contradiction so flagrant would have been noticed by anyone of the most minimal intelligence, let alone one of the smartest minds in the more than two thousand-year history of the western philosophical tradition. [...] So, faced with the alternatives of a Kant blind to this flagrant contradiction and a Kant for whom there was no contradiction given the extent of radical interracial differentiation within the human race, the far more plausible interpretation seems to me that humanity was normatively divided for him.[10]

This, it seems to me, is the strongest argument for ascribing to Kant the view of a consistent inegalitarian. As Mills knows, the actual text of Kant's moral philosophy cannot be squared with the idea that some human beings are not persons, that is, that some human beings do not have a dignity as ends in themselves. By tying the moral status of persons to their capacity for practical reason, and by further characterizing human beings as practically rational animals, Kant is explicitly committed to the view that all human beings have an absolute moral value that is entirely incommensurate.[11]

---

[10] Mills (2014, p. 22).

[11] In the *Religion*, Kant states this notion of the human being as consisting of "animality" (as a living being), "humanity" (as a rational being) and "personality" (as a responsible being). The difference between "rationality" and "responsibility" consists in rationality that is merely theoretical or means-end rationality and practical rationality, that is, rationality that is itself capable of determining the power of choice. (Rel 6:26). Thus, on first sight, it may seem odd that "humanity" and practical rationality can come apart on Kant's view; however, Kant explicitly lists "animality," "humanity," and "personality" as "elements of the determination of the human being." Consequently, what Kant somewhat unhelpfully labels "humanity" in this passage is merely one aspect of being human; the human being not only has "humanity" in light of her theoretical reason but also "responsibility" in light of her practical reason. For a recent discussion of the humanity-personality connection in Kant's ethics, see especially Geiger (2020). For an insightful overview of Kant's conception of the person as linked to a special Gattungswesen (rather than an empirical conception of the person

The strength of Eze's and Mills' claim that Kant was a consistent inegalitarian ultimately comes from the apparent cognitive dissonance that Mills so poignantly describes in the passage above.

Although most Kant scholars have not followed Mills radical conclusion, they share Mills' view about the "flagrant contradiction." But instead of concluding that Kant must have been a secret inegalitarian, they conclude that Kant must have been an inconsistent egalitarian. Lucy Allais, for instance, has recently argued that we should see Kant's racism as a lesson about the nature of racism in general – namely, that racism can involve a lot of cognitive dissonance and even the rationalizing of entirely irrational antipathies. As she puts it:

> Kant's practical philosophy cannot be made compatible with Mills' Untermensch postulation. However, I accept Mills' point about the dramatic and important inconsistency this requires ascribing to Kant, and how striking it is to think Kant could have not noticed such an obvious problem. I argue that rather than trying to make Kant consistent, we can use the example of Kant's racism to tell us something about the nature of racism.[12]

Thus, Allais agrees not only that there is a "dramatic inconsistency" but also that it is striking that Kant did apparently not notice this "obvious problem." This fact, she believes, can "tell us something about the nature of racism: How pervasive it can be in a person's belief system and resistant to evidence – as shown by the possibility of a person's not noticing obvious contradictions in their thinking."[13] On this view, it now seems that Kant's racism might be an interesting empirical phenomenon, and moral psychology might have a place in explaining it. But on this view, it also seems like Kant's racism is not a problem for his moral egalitarianism. While Allais's emphasis that racism can involve deep-seated cognitive dissonance provides an important lesson, it also misses

---

focused on the specific individual's capabilities), see Kaufmann (2019, pp. 200-202). And for a discussion of the historical, scholastic context of Kant's conception of the person as Gattungswesen, see Kobusch (1993, pp. 129-157).
[12] Allais (2016, p. 8).
[13] Allais (2016, p. 20).

what is most worrying about the famous moral philosopher's racism: namely, that whatever the alleged contradiction between his abstract egalitarianism and his racism may be, it is simply not as "obvious," "dramatic" and "flagrant" as either Mills, Allais or other contemporary Kant scholars present it. And this, as I will suggest below, has significant consequences for how we should understand Kant's ethics, and ultimately also for how we engage in moral and political philosophy more generally.

The Problems for the Standard Approach

Although Kant scholars have devoted significant efforts to making sense of the alleged contradiction between Kant's racism and his abstract egalitarianism, they have so far devoted little effort to examining the precise nature of the contradiction. One reason for this lack of interest might be that the basic outlines of Kant's moral philosophy, and in particular Kant's insistence on the inviolable dignity of each person, have become almost synonymous with moral egalitarianism in general. Thus, to question the idea that racist ideology necessarily contradicts Kant's abstract egalitarianism is a counterintuitive route, to put it mildly.[14] But if we really want to figure out if Kant's racism entails a failure of Kantian moral philosophy itself, rather than a mere failure of Kant the person, we had better pursue this route. As this section demonstrates, what seems intuitively contradictory to many contemporary readers did not seem so to Kant; and as I will emphasize in the following section, this lack of an obvious contradiction can tell us something significant about the

---

[14] Of course, philosophy can become complacent about philosophical difficulties exactly where things are too intuitive: Adorno and Horkheimer warned us long ago that "by assuming the unity of humanity to have been already realized in principle, the liberal thesis serves as an apology for the existing order." (Horkheimer and Adorno, 2002, p. 138).

potential shortfalls of Kant's account of moral equality and how we should do moral philosophy more generally.

Here, then, are some of the worst remarks we find in Kant's work about non-White people:

(A) Humanity has its highest degree of perfection in the White race. The yellow Indians have a somewhat lesser talent. The Negroes are much lower, and lowest of all is part of the American races (PG 9:316).[15]

(B) Whites contain all the impulses [Triebfedern] of nature in affects and passions, all talents, all dispositions to culture and civilization and can obey as well as govern. They are the only ones who always advance toward perfection (Refl 15:878, translation mine).

(C) [T]he Hindus have incentives, but they have a strong degree of compo- sure, and they all look like philosophers. Despite this, they are nevertheless very much inclined toward anger and love. As a result, they acquire culture in the highest degree, but only in the arts and not in the sciences. They never raise it up to abstract concepts (V-Anth/Mensch 25:1187).

(D) The Negroes of Africa have by nature no feeling that rises above the ridiculous. Mr. Hume challenges anyone to adduce a single example where a Negro has demonstrated talents, and asserts that among the hundreds of thousands of Blacks who have been transported elsewhere from their countries, although very many of them have been set free, nevertheless not a single one has ever been found who has accomplished something great in art or science or shown any other praiseworthy quality [...] So essential is the difference between these two human races, and it seems to be just as great with regard to the capacities of mind as it is with respect to color (GSE 2:253).[16]

(E) Americans and Blacks cannot govern themselves. They thus serve only for slaves (Refl 15:878, translation mine).

(F) To adduce only one example: One makes use of the red slaves (Americans) in Surinam only for labors in the house because they are too weak for field labor, for which one uses Negroes (GSE 2:438n).

---

[15] Compare also Refl 15:877.
[16] Translation modified.

Quote A comes from Volume II of the *Physical Geography*, published in 1802.[17] Quote B and

E are both taken from Kant's sketches for his lectures on anthropology during the 1780s, which are

contained in his "Reflections" on anthropology. Quote C is taken from a student's notes on one of

Kant's lectures on anthropology,[18] and quotes D and F come from Kant's (pre-critical) work

Observations on the Feeling of the Beautiful and Sublime of 1764.

Quotes A-D clearly demonstrate that Kant believed in a racial hierarchy. Evidently, he

believed that only (male) White Europeans have all the psychological and physiological "talents" that

make human beings excel in work, in the sciences, and the arts. Quote E clearly indicates that Kant

believed that some non-White people cannot politically govern themselves. And most importantly,

quotes E and F might also suggest that Kant condoned slavery. In this context, is also worth noting

that before writing the *Metaphysics of Morals*, Kant never explicitly condemned slavery. Although Kant

was clearly aware of the heated debates between defenders and critics of the institution of chattel

slavery,[19] as well as the tortious treatments slaves were subject to,[20] he did not take an explicit stance

on this topic in any of his published texts until the *Doctrine of Right*.

If Kant had condoned slavery for non-White people while also rejecting it implicitly through

the Formula of Humanity and explicitly in his *Metaphysics of Morals*, this would indeed be a plausible

---

[17] Because there are similarly racist remarks in other works, I assume that this is an adequate representation of Kant's thought at the time. However, it is open to debate to what extent the *Physical Geography* can be said to be Kant's text at all. For a discussion, see the editorial preface to the translation of the *Physical Geography* in the CUP edition's volume on Natural Science (2012 pp. 434-437). For a detailed discussion of Kant's *Physical Geography*, see Elden and Mendieta (eds., 2011).

[18] This Academy Edition's text of these lecture notes is based on the book Immanuel Kant's *Menschenkunde oder philosophische Anthropologie*, edited by Johann Bergk (under the pseudonym of Friedrich Starke) in 1831. To my knowledge, we do not know where Bergk got the manuscript from and who originally took these notes.

[19] See ÜGTP 8:174n where Kant approvingly quotes a German text by Matthias Sprengel (1786) as written by a "knowledgeable man," in which Sprengel had paraphrased the pro-slavery polemic by James Tobin (Tobin, 1785) against the abolitionist James Ramsey. For a discussion of the debate between Tobin and Ramsey, see Shyllon (1977, pp. 59-70). Kant was thus clearly aware of the arguments for and against the institution of slavery at the time.

[20] See Kant's remarks in his essay *Toward Perpetual Peace* at ZeF 8:359.

contradiction. Although it is notoriously difficult to extrapolate what Kant might have meant exactly

by treating someone as an end in themselves, and almost equally difficult to know what it means to

treat someone merely as a means,[21] it is uncontroversial that treating someone as a living tool is

incompatible with treating them as an end in themselves. However, Kant also denies that slavery

(whether for non-White people or White Europeans) could be a rightful institution in his *Metaphysics*

of Morals.[22] Already in the introduction to the *Doctrine of Right*, Kant claims that this volume deals

with the relations between human beings who have both rights and duties vis-à-vis each other. By

contrast, the book does not deal with humans' relationship to a being that has only rights but no

duties (God), for a duty to such a being would be "transcendental," that is, no corresponding

external subject to whom this duty is owed could be found. And, most importantly, Kant also says

that his book does not deal with relations to beings who do not have rights but merely duties ("serfs,

slaves"). But Kant does not merely reject slavery as a potential topic of the book, leaving open that it

could still be an otherwise acceptable institution. Kant explicitly rejects this potential topic as empty,

because it would imply that there are human beings without personality. And because all human

beings are persons a doctrine of right need not concern itself with such empty topics.[23] Moreover, in

section I of the *Doctrine of Right* (which deals with the acquisition of a right to property in external

objects), Kant explicitly denies that anyone could own other human beings. "So someone can be his

own master (sui iuris) but cannot be the owner of himself (sui dominus) (cannot dispose of himself

---

[21] For recent discussions of what it means to treat someone as a mere means, see Denis (2007), Kerstein (2009, 2013), Formosa (2014), Papadaki (2016), Patrone (2018) and Kleingeld (2020).

[22] For the purpose of this paper, I leave aside the question how Kant's account of the *Groundwork* and second *Critique* allows the move to his rejection of slavery in the *Doctrine of Right*. The precise nature of the relation between these works remains a topic of considerable controversy. This of course leaves open the possibility that Kant did not believe legal equality follows from moral equality. For the purpose of this paper, I assume that legal equality of at least able-minded adults would follow from moral equality in order to assume the strongest case for the standard reading, since such a continuity between moral and legal equality would then provide a plausible case of a contradiction in Kant.

[23] See also Kant's discussion of innate right as "innate equality" at MS 6:237 and Kant's remarks in his preparatory notes for *Toward Perpetual Peace* VAZeF 23:174.

as he pleases) – still less can he dispose of others as he pleases – since he is accountable to the humanity in his own person." (MS 6:270).[24]

Despite rejecting slavery in general, Kant also makes an exception in the case of a criminal who has "forfeited his personality by a crime." (MS 6:283). As Kant puts it, "Certainly no human being in a state can be without any dignity, since he at least has the dignity of a citizen. The exception is someone who has lost it by his own crime, because of which, though he is kept alive, he is made a mere tool of another's choice (either of the state or of another citizen)." (MS 6:329–30).[25] Notwithstanding this exception, Kant makes clear that such a loss of personality can only occur through a particularly grave crime. Thus, in the *Doctrine of Right*, Kant regards the legal institutions of slavery or serfdom as incompatible with the personhood of the subjugated. And however racist Kant's views of non-White people were, he clearly did not believe that not being White was literally a crime that would forfeit one's personality. As mentioned above, even the most critical interpreters like Charles Mills point out that Kant's work is incompatible with the denial of personhood to any human being merely because of the color of their skin.

Therefore, we might conclude that Kant's remarks on slavery in his unpublished notes for his lectures on anthropology (written sometime during the 1770s-1780s) are incompatible with his account of moral equality in the *Groundwork* as well as his substantive political and legal philosophy developed in the *Doctrine of Right* in the late 1790s. But consider again the following two remarks from Kant's preparatory notes on anthropology and his *Observations on the Feeling of the Beautiful and Sublime*:

(E) Americans and Blacks cannot govern themselves. They thus serve only for slaves.

---

[24] See also MS 6:283, 6:359-60 and V-NR/Feyerabend 27:1319.
[25] Kant also denied that enemy combatants could be rightfully enslaved in war (MS 6:348-349).

(F) To adduce only one example: One makes use of the red slaves (Americans) in Surinam only for labors in the house because they are too weak for field labor, for which one uses Negroes.

These two statements demonstrate Kant's uncanny indifference to the suffering of non-White people as well as his beliefs in both racially determined characteristics (both physical and psychological) and a racial hierarchy. The even more troubling fact, however, is that neither of these racist diatribes is, strictly speaking, incompatible with Kant's formalistic account of moral equality. Kant's view of moral equality in the *Groundwork* is based on the thought that all mature rational beings have the capacity for autonomous practical reason and are thus capable of setting themselves moral ends. And Kant believed that this claim can be derived from an analysis of our practical reason: In representing some of my ends as moral ends, I represent them according to a particular form (i.e., the Categorical Imperative); and this form can be expressed in different ways in order to highlight the commitments that result from this form of representing ends. One of these commitments is that insofar as I take myself to be capable of representing moral ends through my practical reason, I must also accept the capacity of every other rational being to set themselves ends through practical reason (i.e., as an end in itself). This special capacity for setting ends makes a living being into a person; and for Kant, this is what a person's dignity consists in.[26] Quotes A-D clearly display Kant's belief in a racial hierarchy based on the idea that only (male) White Europeans have all the psychological and physiological "talents."[27] However, this belief by itself does not entail that

---

[26] Compare also Kant's definition of "person" in the introduction to the *Metaphysics of Morals* at MS 6:223.

[27] It is worth repeating that in this regard, Kant was not simply a child of his time. Kant's remarks in his review of the first two volumes of Herder's I*deen zur Philosophie der Geschichte der Menschheit* (1784 and 1785) make clear that by the mid 1780s, Kant had become aware of conflicting anthropological evidence available at the time, and the importance of picking one's sources carefully: "But now from a multiplicity of descriptions of countries one can prove, if one wants to, [...] that Americans and Negroes are each a race, sunk beneath the remaining members of the human species in their mental predispositions, but on the other side by just as apparent records that as regards their natural predispositions, they are to be estimated equal to every other inhabitant of the world; so it remains to the choice of the philosopher whether he wants to assume differences of nature or wants to judge everything in accordance with the principle *tout comme chez nous*, so that all his systems he erected on so shaky a foundation must take on the look of rickety hypotheses."

Kant also believed non-White people are not in possession of practical reason and thus ends in themselves. Moreover, quote E does not claim that slavery or serfdom are rightful institutions (pace Mills' suggestion in the passage quoted above); rather, E claims that an entire group of people (most non-White people) would not be good at anything except serving as slaves.[28] The same applies to quote F. Notwithstanding its hideousness, it does not logically contradict either the view that all human beings are ends in themselves qua their capacity for practical reason, nor does it contradict Kant's claim that slavery cannot be a rightful legal institution.[29]

But would it not be more plausible to think Kant changed his mind between writing the racist remarks E and F quoted above and writing his *Doctrine of Right* with its strict rejection of slavery? That Kant had "second thoughts" on race? Although I am quite skeptical of this conjecture,

---

(RezHerder 8:62). For a discussion of Kant's debate with Herder, and his ensuing dispute with Georg Forster, see especially Eberl (2019, pp. 401–407).

[28] Note that this is not exactly the same claim that readers often attribute to Kant, namely that Americans and Black people are "natural slaves" (Mills, 2014, p. 22) or "made to be" slaves (McCabe, 2019, p. 6). At least if "natural slave" is taken in Aristotle's sense of the term, this view clearly cannot be attributed to Kant. Kant's racist diatribe claims that Americans and Black people are only really good at the work of slaves, and do not excel at other work; his claim is not that there is some moral or teleological principle determining that they ought to serve as slaves.

[29] This has also been pointed out by Boxill (2017, p. 46). One might object to Boxill and my interpretation of quotes E and F by claiming that this is too charitable to Kant. After all, both quotes might sound like implicit condonements of the enslavement of non-white people. Thus, one might be tempted to follow Kleingeld and think that at the time he wrote these passages, Kant still approved of slavery and changed his view later (in the 1790s) when writing the *Metaphysics of Morals*. However, I believe we must reject such a reading for two reasons. First, reading these early passages E and F as condonements of slavery would also presuppose either a consequentialism or teleological thinking that Kant rejected – for Kant, whether or not someone is good at a task is simply irrelevant to whether a particular social relation is morally wrong. And second, Kant's handwritten remains also contain notes more explicitly critical of slavery, like the following from between 1772 and 1777: "It is well possible that human beings are ruled over as slaves or minors through coercion, status and prejudice; but all these evils must come to an end one day, and philosophy must provide the principles for this, if it should have any use at all." (Refl 15:230, translation mine). Compare also Kant's discussion of suicide in the Vigilantius lectures from 1793/94. There, Kant discusses two casuistical questions about suicide in which he wants to exemplify both the immorality of suicide while also highlighting why suicide can seem so reasonable. One example concerns a slave who wants to end their life because they do not consider the life of a slave worth living (V-MS/Vigil 27:603). This example strongly suggests that Kant slavery was an unjust institution. And finally, in Kant's *Observations on the feeling of the beautiful and the sublime* from 1764, he describes the disposition of a virtuous person: "He [the person of a melancholic frame] has a lofty feeling for the dignity of human nature. He esteems himself and holds a human being to be a creature who deserves respect. He does not tolerate abject submissiveness and breathes freedom in a noble breast. All shackles, from the golden ones worn at court to the heavy irons of the galley-slave, are abominable to him." (GSE 2:221) For further discussion of this topic cf. Lu-Adler (2022b), who argues that Kant "was morally indifferent to [slavery] (as an institution), and so he neither straightforwardly endorsed it as morally permissible nor condemned it as morally wrong" (ibid 269).

Kant may well have changed his mind.[30] Whether (and to what extent) he might have done so we will never know for sure. But whether or not he did so is, I believe, orthogonal to the present question – however interesting it might otherwise be. For the present question is whether there really is an obvious contradiction between Kant's racism and his abstract egalitarianism and what consequences this might have for our understanding of Kant's ethics and Kantian moral and political philosophy in general. If there is no such obvious contradiction, then this fact remains of philosophical significance whether or not (and to whatever extent) Kant might have changed his mind.[31]

Alternatively, could it not be that when Kant denied that slavery could be a rightful institution in the *Doctrine of Right*, he held on to a silent taxonomy of persons, according to which some human beings are not persons (as Mills has suggested)? As I already pointed out above, the reason for ascribing such a silent taxonomy – contrary to what Kant explicitly says about all human beings being ends in themselves – has been the assumption that there is a "flagrant" contradiction between Kant's racist remarks and his abstract egalitarianism. However, Kant's racist remarks do not explicitly condone slavery, despite their hideousness; nor do they deny the personhood of non-White people. Thus, Mills overstates the reason for attributing such a silent taxonomy to Kant. If Kant's statements are not obviously contradictory, then we also have no obvious reason to assume a silent taxonomy either; and surely, we cannot presuppose a silent taxonomy in Kant only to then justify our attribution of this silent taxonomy to Kant by pointing to an alleged contradiction that itself relies on presuming that very silent taxonomy.

---

[30] For a detailed discussion of this question, see especially Kleingeld (2007), Bernasconi (2011), Kleingeld (2014) and Lu-Adler (2022b).

[31] One limited connection between the present topic and the question whether Kant's views changed is that if there is no obvious contradiction, then we also incidentally have fewer reasons to think that Kant would have changed his mind.

Although Kant's racist remarks do not, technically speaking, imply an acceptance of slavery, they still demonstrate his belief in a racial hierarchy – only (male) "Whites" are said to have all the psychological and physiological "talents" that make human beings excel. Importantly, quote (E) suggests that non-White people could not successfully govern themselves politically. Even if this is not an endorsement of slavery, at the very least it suggests that Kant did not endorse the idea of equal political and legal rights; and for the sake of the argument, we may plausibly take it to suggest that Kant also would not have endorsed equal rights to political participation and representation.[32] Could this be the "flagrant contradiction" that the standard reading assumes? Unfortunately, it cannot. The inconvenient truth is that Kant's abstract account of the moral equality of all persons in the *Groundwork* and the *Critique of Practical Reason* is just not obviously incompatible with inegalitarian substantive political doctrines. In fact, Kant himself explicitly considered the question what rights followed merely from the status as a person, and denied that this status would be sufficient for equal legal and political rights. For instance, in the *Doctrine of Right*, Kant makes clear that he believes women should not be accorded the same legal and political rights as men, while also explicitly affirming women's moral equality as ends in themselves.

> For instance, in a discussion of citizenship, Kant says:
>
> [B]eing fit to vote presupposes the independence of someone who, as one of the people, wants to be not just a part of the commonwealth but also a member of it, that is, a part of the commonwealth acting from his own choice in community with others. This quality of being independent, however, requires a distinction between active and passive citizens, though the concept of a passive citizen

---

[32] To my knowledge, there is no explicit discussion in Kant's work of the political and legal rights that non-white people should be accorded within the state. In the *Doctrine of Right*, Kant eventually discusses the rights of indigenous peoples in the context of international law, in his discussion of colonialism and the settlements of land (and affirms their rights to land). One might perhaps take the claim that all peoples have a right to their land, and Kant's rejection of colonialism, as incompatible with the idea that women, non-white people, and dependent male citizens should not have full political rights within the same polity. However, this alleged incompatibility is, once again, not self-evident and would require more complex elaboration.

seems to contradict the concept of a citizen as such. - The following examples can serve to remove this difficulty: […] a minor (naturaliter vel civiliter); all women and, in general, anyone whose preservation in existence (his being fed and protected) depends not on his management of his own business but on arrangements made by another (except the state). (MS 6:314)

In the paragraph immediately following Kant's characterization of women as necessarily "passive citizens" who should not be allowed to vote, he continues:

This dependence upon the will of others and this inequality is, however, in no way opposed to their freedom and equality as human beings, who together make up a people […] But not all persons qualify with equal right to vote within this constitution, that is, to be citizens and not mere associates in the state. For from their being able to demand that all others treat them in accordance with the laws of natural freedom and equality as passive parts of the state it does not follow that they also have the right to manage the state itself as active members of it […] (MS 6:315, emphasis added)

Given his misogynistic discussion of the status of women,[33] Kant clearly did not believe that abstract moral equality, that is, the mere status of every person as an end in themselves, entailed that every person should also have even remotely similar legal and political rights – as we saw, Kant explicitly denied this. Instead, Kant's discussion of women and the family in the *Doctrine of Right* makes clear that he viewed some human beings as akin to children: Although they have dignity and are ends in themselves, he also believed that they would not have the physical and psychological abilities to successfully govern a body politic, a household, or even themselves.[34]

It is tempting to think that Kant might here simply contradict himself once again. However, to deny the relevance of Kant's misogynistic views in the *Doctrine of Right* for the present question by presuming that Kant must there simply contradict himself once again would beg the question why

---

[33] A discussion of Kant's views on women would be far beyond the scope of this paper. Important passages in this context are GSE 2:230ff and 2:236, MS 6:276-280, 6:314 and 6:358, Anth 7:303-309, WA 8:35, MAM 8:113, and TP 8:295. For recent discussions of Kant's views of women, see especially Varden (2017) and Kleingeld (2019).
[34] Note that Kant at times expressed (surprising) confidence in the moral knowledge of children. See, for example, MS 6:480-81 and TP 8:286.

we should discount Kant's explicit discussion of the unequal legal and political rights of moral equals as evidence for his own conception of moral equality. Surely, we cannot simply preclude Kant's views on women's rights as evidence for his conception of moral equality only to then justify this preclusion by saying that his views on women's rights are incompatible with his moral egalitarianism. As Kant did not believe that it was at all contradictory to claim that women should have significantly fewer rights than men, it seems difficult – however tempting it may be – to claim that there is really an obvious contradiction between Kant's own abstract moral egalitarianism and his racist beliefs about the alleged unsuitedness of non-White people to be active members of a body politic.[35]

Similarly, one might object that Kant's abstract egalitarianism is just not compatible with the idea of passive citizenship in the *Doctrine of Right*. Although most commentators on Kant's political philosophy take for granted that some citizens (paradigmatically children) should not have active rights of political participation, one may take issue either with the notion of passive citizenship in general or with the way Kant draws the distinction. Over the last decades, existing literature on Kant's *Doctrine of Right* has become truly voluminous – including ample controversy over Kant's notion of passive citizenship.[36] However, the crucial point is this: If passive citizenship in general or Kant's way of drawing the line are incompatible with Kant's abstract moral egalitarianism, then this incompatibility is again far from obvious – for, this incompatibility is precisely one of the open questions in contemporary Kant scholarship. Whatever the potential contradiction between Kant's

---

[35] In fact, Kant's political inegalitarianism goes much further than his racism and misogyny. As the passage quoted above (MS 6:314) already indicated, Kant also argued that white males are not fit for being active citizens if they are dependent on someone else through their employment. In this, Kant is entirely unoriginal. For another politically inegalitarian view on which all manual laborers are allegedly unfitness for being citizens, see Aristotle (1998, 1278a20-21, 1328b33-1329a39).

[36] For recent discussion, see especially Davies (2021), Moran (2021) and Vrousalis (2022).

formal egalitarianism and the idea of passive citizenship, it is not so "flagrant," "startling," "obvious" or "gross" as to prevent a considerable amount of philosophical debate.

Of course, none of this is to say that a Kantian egalitarianism could not, when further developed, turn out incompatible with Kant's racist (and misogynist) commitments. Rather, I am suggesting that any such incompatibility is far less obvious and flagrant than commentators so far assume; that, further, Kant explicitly considered the possibility of such an incompatibility and denied it (in the case of the political rights of women); and that this lack of any obvious contradiction should be problematized in discussions of Kant's ethics because, as I argue in the following section, it highlights a potential blind spot in a particular way of doing moral and political philosophy more generally.

Kant's Racism as a Philosophical Problem

What, then, should we make of Kant's moral egalitarianism? Nothing that was said so far entails that Kant's abstract analysis of practical reason in the *Groundwork* and the *Critique of Practical Reason* was, technically speaking, wrong; rather, I have suggested that is not enough. Kant's moral egalitarianism might be necessary for a substantive account of moral equality, but it is far from obvious that it would be sufficient. While the idea of human beings as ends in themselves can rule out some inegalitarian accounts like Aristotle's defense of slavery, it is not clear that Kant's abstract egalitarianism should also be enough to rule out bigoted views of the type we encounter in Kant's own writings, as most authors still habitually assume.

But how is any of this of philosophical interest? In other words, why might anyone outside of Kant scholarship care? Kant's racism is philosophically interesting because it demonstrates a

potential blind spot within formalistic approaches to moral and political philosophy, and should

serve as a warning to other moral and political philosophers whose ideal vision of society is

supposedly derived from abstract moral principles. As I have argued above, Kant's abstract

egalitarianism seems prima facie compatible with conceptions of what ethical and political life

should be that (hopefully) most of us would intuitively judge morally repugnant. (Indeed, I suspect

that this seeming compatibility of Kant's egalitarianism with such a wide variety of substantive

conceptions of ethical and political life is part of the explanation for its immense popularity.) And if

Kant's egalitarianism is prima facie compatible with stark forms of racism, then relying on Kant's

conception of moral equality by itself cannot be enough to show that we are not merely varnishing

our existing prejudices and parochial intuitions.[37]

So far, few commentators have openly acknowledged the prima facie compatibility between

Kant's racism and his abstract egalitarianism, and none of them have yet explicitly problematized

this prima facie compatibility to the extent I believe we should. Todd Hedrick, toward the end of his

discussion of the place of racial differences in Kant's teleological conception of world history, briefly

mentions this problem in passing. As he points out, Kant is "maddeningly sanguine about

permitting substantive inequalities within structures of formal equality, as his insistence that

restricting citizenship to economically self-sufficient men is perfectly compatible with the moral

demand for civic equality and freedom for all attests." Hedrick continues: "Many of us today would,

I think, be much quicker than Kant to claim that substantive inequalities often render formal

---

[37] One may even argue that Kant's abstract egalitarianism – if it's limits are not taken seriously – can be a powerful ideology in the pejorative sense. Although Marx himself identified ideology with false beliefs that are in the interest of one social class, and whose existence can be explained functionally as helping maintain a system of material and sexual exploitation, we might follow Michael Forster's broader conception of ideology as not necessarily involving falsehood (Forster, 2015). From this perspective, an abstract conception of equality might be ideology par excellence, insofar as it might be mistaken as justifying every other conservative conception of a "just society."

equality empty, although the point at which substantive inequality undermines formal equality is more often than not debatable."[38] Unfortunately, Hedrick does not discuss this problem any further in his discussion of Kant's cosmopolitanism, and his mention of the problem is not quite as explicit as could be. By contrast, Elvira Basevich perspicuously points out that it is not at all obvious – contrary to the frequent claims by Kant scholars – how Kant's abstract egalitarianism should contradict his racism.[39] But unfortunately, Basevich also does not problematize this potential consistency itself. Instead, Basevich (re)interprets Kant's theory of race as part of an alleged "non-ideal theory," that is, as meant to "illustrate the human species" tendency to deviate from the requirements of justice,"[40] and Kant's abstract egalitarianism as part of an alleged Kantian "ideal theory." Thus, on her view, the problem is not any potential contradiction between Kant's alleged "non-ideal theory" (i.e., his theory of race) and his alleged "ideal theory" (i.e., his abstract universalism and mature political philosophy) but the implausibility of Kant's alleged "non-ideal theory."[41] Consequently, Basevich's response is to provide a new "Kantian" non-ideal theory, as an alternative to Kant's own bigoted theory of race: namely, a Kantian approach to racial justice reform. Although I find Basevich's proposal for a Kantian approach to racial justice reform compelling, this way of framing Kant's racism also retains a shortfall: From this perspective, the problem of the prima facie compatibility between Kant's abstract egalitarianism and his racism drops out of view. For now, we simply focus on finding a plausible "non-ideal" theory. But if Kant's alleged "non-ideal" (i.e., racist) theory was compatible with Kant's "ideal theory" (i.e., his abstract egalitarianism and mature political philosophy) to begin with, what, we may ask, is the substantive

---

[38] Hedrick, 2008, p. 267.
[39] Basevich, 2020, pp. 227-228.
[40] Ibid. 227.
[41] This interpretation, taking Kant's abstract universalism to be his "ideal theory" and his theory of race part of his "non-ideal theory," is also adopted by Lu-Adler (2022a pp. 272-273). Since Basevich's treatment of this topic is significantly more elaborate my discussion focuses on Basevich (2020).

content of the (presumably) new Kantian "ideal theory" in light of which we judge the current state of affairs as unjust in the first place? Unfortunately, Basevich's account of Kantian racial justice reform does not discuss this question. Thus, despite her astute observation that Kant scholars have been too quick in assuming an obvious contradiction in Kant's thought, on Basevich's account, the philosophical problem highlighted by Kant's racism – the relation between an abstract principle of equality and substantive ideals – drops out of view.[42]

As mentioned above, the fact that Kant's racism highlights this philosophical problem does not imply that his abstract analysis of practical reason was wrong; rather, my argument has suggested that this analysis and its resulting moral egalitarianism are not enough to provide the antidote to Kant's racism they are often hoped to be. Consequently, Kant's moral philosophy need not be outright rejected – rather, what must be rejected is (merely) the mistaken belief that a correct, abstract formula will, by itself, give us a full understanding of how to collectively build a practice of moral equality.

This thought prompts several important questions, two broad and two narrow. If Kantian abstract egalitarianism is not enough to provide an antidote to racism by itself, what else is needed to respond to racism in a philosophically satisfying way? And how should philosophy as a discipline deal with racism? Even if the literature within contemporary critical philosophy of race were less extensive than it is, it would be immodest to suggest or outline where exactly this question should lead us at the end of an essay on Kant. Much closer to the spirit of this paper are the following two narrower questions: What might it look like to respond to Kant's racism within a broadly Kantian

---

[42] In this way, Basevich follows authors like Mills and Allais who hold on to Kant's abstract egalitarianism in different ways and thus deny that Kant's racism points to a deeper problem with abstract egalitarianism. Indeed, Basevich explicitly claims to "tackle Charles Mills's suggestion to theorize a 'black radical Kantianism', because following Mills, it is instructive to illustrate the extent to which Kant is [a] helpful resource for racial justice." (Basevich, 2020, p. 224 n7).

framework? And what challenges might such approaches face? While a full discussion of even these narrower questions would go beyond the scope of this paper too, a brief outline is still worthwhile.

One tempting possibility might be to think that Kant's abstract egalitarianism simply needs to be furnished with adequate empirical knowledge about human nature.[43] In his *Observations on the Feeling of the Beautiful and Sublime*, Kant concurred with Hume in the bizarrely false assertion that "among the hundreds of thousands of Blacks [...] not a single one has ever been found who has accomplished something great in art or science or shown any other praiseworthy quality." In light of such statements, it might be tempting to think of the required supplement as simply consisting of adequate empirical knowledge.

However, this approach faces its own challenges. First, one might worry that if we conceded the task of countering racism to empirical investigation, we would also have to concede that philosophy itself could not explain why a certain type of racism is wrong. Because by conceding the question whether or not some persons are fit for substantive moral and political equality, we would also concede to the racist what they want most: namely, the air of moral legitimacy for an empirical investigation into the characteristics of any singled-out group. Once the task of determining whether or not some person or persons are fit for moral and political equality is conceded to empirical investigation it is difficult to see how, on that basis alone, it could ever be taken away. Without some further philosophical resources, one could presumably only counter the racist with existing empirical evidence and argue that enough empirical investigation has taken place. However, one may think that such a response would omit an important moral criticism. More importantly still, this way of supplementing Kant's abstract egalitarianism would not so much answer the philosophical question

---

[43] See, for example, Hill & Boxill (2001).

about the relation between abstract principles and substantive ideals that Kant's racism highlights as push it to a different place. In other words, this approach only "supplements" Kant's abstract egalitarianism on the assumption that people's natural abilities are decisive for determining their rights as moral equals. However, whether or not, for example, rights to political participation should be distributed in a way that tracks people's natural abilities and talents – and if so, which ones? – is itself a normative question that is not obviously answered by Kant's abstract egalitarianism. And this is to say nothing yet of the epistemic difficulties of estimating people's "natural abilities and talents." Because surely the mere absence of certain types of achievements cannot be taken as evidence of lacking abilities and talent. As Emmanuel Eze has perspicuously pointed out in the context of Kant and 18th century colonialism: "Kant notes that some races, as if by right, 'have educated the others and controlled them with weapons.' [...] He does not raise the question of whether it is the 'education' and superior weapons of the conquering tribe that produce the 'immaturity' of the conquered."[44] Under conditions of inequality, one may see a lack of talents and achievements where one is really staring at the work of oppression.

Another possible way of supplementing Kant's abstract egalitarianism within a Kantian framework might be to start from a postulate of political, rather than merely moral, equality and to rely on some hermeneutic device like a contractualist procedure and normative assumptions about people's substantive interests to guide our political imagination. The most formidable recent attempt to develop such an anti-racist Kantianism is undoubtedly Charles Mills' political philosophy.[45] Conceiving of existing political reality as embodying an exclusionary racial contract, we may ask what a truly universal social contract may look like and how reparative justice might address existing

---

[44] Eze (2001,p. 81).
[45] See especially Mills (2017, 2018).

and previous wrongs. Of course, an essay on Kant's racism is not the place to provide a full discussion of Mills' political philosophy. Here, I only want to mention one difficulty this approach would have to respond to: Even a postulate of political equality cannot ensure that the philosophical difficulty of moving from abstract to substantive egalitarianism will not return. In the preface to his theory of political equality, Charles Beitz draws our attention to a version of this problem for theories of democracy and political participation. Speaking in the context of egalitarian institutional reforms in the United States from the 1960s to late 1980s, Beitz points out:

> All of these reforms were defended by their proponents as requirements of political equality. Yet, as continuing controversy in the courts, the legislatures, and the political parties attests, there is no consensus about the meaning of this principle. Even the most ardent supporter of the reforms is bound to be troubled by this questioning: while we feel confident that political equality means something, it is surprisingly difficult to give it a clear explanation and defense.[46]

Beitz's himself proposes for a contractualist account he labels "complex proceduralism," relying both on a hypothetical social contract procedure and a presumed set of substantive, regulative interests of individual citizens that help him determine what might count as "sufficient reason to refuse to accept" procedures of political participation. Here, the main challenge is to supply both the structure of the procedure and the normative assumptions about people's interests in a way that does not smuggle our existing prejudices back into our substantive account of equality. But whatever we might think about Beitz's concrete proposal, he perspicuously identifies the underlying philosophical difficulty in the context of political participation:

> The most natural thought is that a requirement of procedural equality is compelled by some version of the more basic principle that persons have a right to be treated as equals. But […] there are very deep difficulties in this relationship, and its plausibility fades on analysis. To anticipate, equal treatment might be seen either as an abstract moral requirement or as a concrete rule with

---

[46] Beitz, 1989, p. x.

determinate institutional content. If the idea is regarded abstractly enough to be noncontroversial, then its application to institutional questions will be uncertain without controversial intervening premises. […] If, on the other hand, the principle is taken to specify a determinate institutional right – for example, a right to have one's expressed interests given equal weight in the determination of policy – then it will fail to settle the issue that provoked it.[47]

As I hope to have demonstrated, the problem Beitz identified in his 1989 book (as well as the surprisingly common disinterest in this problem[48]) appears to be merely one instance of a larger phenomenon: the philosophical difficulty of getting from abstract principles to substantive ideals. And as examples like Mills' "Black radical Kantianism" and Basevich's proposal for Kantian racial justice reform show, some authors may have compelling political ideals while still leaving open the question how, exactly, their substantive ideals and postulates are supposed to emerge from their abstract Kantian egalitarianism.

As already mentioned above, this is not to rule out once and for all the possibility that a further developed Kantian egalitarianism might turn out incompatible with Kant's racist and misogynist commitments. Rather, I have argued that such an incompatibility is far less obvious than readers have traditionally assumed and highlighted some problems for developing the Kantian approach further. A full discussion of how we should go beyond Kant's abstract egalitarianism would go far beyond the scope of this paper – and possibly also beyond the scope of a single book. Because if my argument was correct, then we should not expect such a discussion to be reducible to yet another concise formula unless we want to rehearse the problem of abstract egalitarianism yet again. Thus, my comparatively modest aim throughout this paper has been to point out that Kant's

---

[47] Beitz, 1989, pp. 6-7.
[48] As Beitz points out in the context of theories of democracy: "Not everybody will agree that there is any need for a theory of political equality. In fact, the most widely held view of the subject is to deny it. [...] Something like this conception of political equality represents a persistent conviction among contemporary democratic theorists; indeed, it has become a kind of philosophical orthodoxy, perhaps because it has seemed to express so obvious a truth as not to require systematic defense." (Beitz, 1989, pp. 4-5).

racism highlights a genuine philosophical difficulty: How do we get from abstract principles to

substantive ideals in a way that doesn't merely polish up our existing prejudices?[49]

Corresponding to this philosophical difficulty of getting from principles to substantive ideals,

there is a potential blind spot in moral and political philosophy when it relies too complacently on

abstract principles. And getting the seeming compatibility between Kant's racism and his abstract

egalitarianism in view demonstrates this blind spot in one of our most prominent accounts of moral

equality. In this way, Kant's racism is philosophically significant quite unlike, say, Aristotle's defense

of slavery. Aristotle's views on the permissibility of slavery are deeply problematic, but hardly

anyone would treat Aristotle's ethics as an uncontroversial starting point for a progressive account

of moral and political equality. By contrast, as Robert Bernasconi has aptly put it, what makes the

racism of the European Enlightenment philosophers so troubling is that 'they join their racism to

the new universalism or cosmopolitanism, which is supposed to be one of the great achievements of

the Enlightenment and an antidote to racism'.[50] And this remains a popular view about Kant's ethics

today: Similar to Mills' and Basevich's attempt to theorize a "Black radical Kantianism," Jennifer

Mensch has recently claimed that "we must take Kant's universalism for what it is: an approach to

---

[49] Of course, such polishing up of existing prejudices with abstract egalitarian formulas is not the only way in which philosophy – and entire philosophical traditions – can be implicated in racism. For a discussion of another way, see Katrin Flikschuh's "Philosophical Racism" (2018). Flikschuh identifies a form of philosophical racism consisting in "unstated background assumptions about which contexts and domains of human experience are or are not worthy of philosophical reflection." (Ibid. 103). Because of this type of academic racism, Flikschuh ultimately suggests: "Philosophical discussions about racial injustice assume that the discipline is capable of offering theoretical solutions to it; my argument has been that our inherited terms of philosophical discourse are themselves a likely source of the problem." (Ibid. 107). Thus, our arguments reach similar conclusions, albeit for different (if mutually compatible) reasons.
[50] Bernasconi (2002, pp. 146–147). Although Bernasconi's remark captures perfectly the general reaction to Kant's and other Enlightenment philosophers' racism, it is somewhat misleading to call the universalism and cosmopolitanism of the Enlightenment "new." Earlier versions of universalism were a central tenet of ancient Stoicism. See for instance Cicero, 1991, pp. 108–111. For a detailed discussion of Stoic cosmopolitanism, see especially Schofield (1991). For a comparison of Kant's and the Stoic's cosmopolitanism, and Kant's reception of the latter, see Nussbaum (1997). Christian philosophers had also long since believed in a basic moral equality of all humans. See, for instance, Augustine (1998, pp. 942–945).

morality whose own claims undermine the racist provincialism on display in Kant's works" because "the language of universal moral worth […] still offers us a conceptual tool of the kind that can be helpful when combatting the ongoing problems of racism and misogyny today."[51] However, if my argument was correct, then we have to take seriously the possibility that Kant's abstract egalitarianism that dominates much of contemporary moral and political philosophy cannot, by itself, provide this antidote.

---

[51] Mensch (2017, pp. 142 and 143).

Conclusion

Practical Therapy and Collective Neuroses

Empty is that philosopher's argument by which no human suffering is therapeutically treated. For just as there is no use in a medical art that does not cast out the sickness of bodies, so too there is no use in philosophy, unless it casts out the suffering of the soul.

-- Epicurus, Us. 221[1]

[B]y assuming the unity of humanity to have been already realized in principle, the liberal thesis serves as an apology for the existing order.

-- Horkheimer and Adorno, *The Dialectic of Enlightenment,* p. 138

Kant's moral philosophy forms part of a longstanding tradition of ethical reflection aimed at shaping human character and freeing us from our own forms of self-entrapment. Getting clear about the practical purpose of moral philosophy in Kant shows how central this therapeutic conception of moral philosophy is not merely on an individual level, but on a social level too – and how central this therapeutic conception of moral philosophy remains today. Whether one likes it or not, Kant's moral vocabulary has become part of the cultural repertoire through which we make sense of our own ethical and political lives. Therefore, even critics of Kant would be well-advised to take the insights and limits of Kant's moral philosophy seriously. To follow Kantian ethics without any such appreciation is to risk turning ideals of reason into a collective neurosis: a pathological way of making incomplete sense of our practical lives, and perpetuating such an incomplete

---

[1] Porphyrius ad Marcellam 31 p. 209, 23 Nauck. Translation by Martha Nussbaum (2018 p. 102).

understanding as a society. As against such collective neuroses, moral philosophy ought to be therapy: helping us achieve a better understanding of the concepts through which we already make sense of our moral lives, helping us avoid pathologies, and follow the ideals we are already committed to as finite, rational beings.

This book has aimed to clarify Kant's account of practical therapy, his corresponding ambitions in a metaphysics of morals, and the underlying formalism of Kant's ethics. With this larger picture in view, we can now appreciate in more nuance the specific limits of Kant's moral philosophy. The formality of the a priori requirements of morality – in other words, the fact that the requirements of morality are the form of pure practical reason – allows Kant to champion a strict epistemic egalitarianism about moral judgment. Having a good will does not depend on one's upbringing, one's psychological traits or talents, or on one's education. Everyone can have a good will and comply with the requirements imposed by the form of practical judgment. And it is this universal ability for autonomous practical judgment of all rational beings that allows Kant to insist that every rational being – and thus all humans as dependent, rational beings – have an inviolable dignity as ends in themselves. This is the promise of Kant's ethics as building up from the form of practical reason: a truly universal egalitarianism. But precisely because of this focus on the formal requirements of autonomous practical reason, Kant's resulting moral egalitarianism of human beings as ends in themselves seemingly imposes comparatively thin requirements on the treatment of others. As Kant's view of the psychological traits and talents of non-white people and his misogynistic remarks reveal, Kant believed that some people are inflicted by strong natural inclinations clouding the requirements of practical reason – so much that they would remain in a perpetual state of immaturity, unlikely to ever reach a virtuous disposition and would be, like children, unfit for governing others and even themselves.

Thus, the insight and limits of Kant's moral egalitarianism go hand in hand. Kant's view that every rational being can act in accordance with the formal requirements of autonomous practical judgment allowed him to be a truly universal egalitarian; but precisely because of its formality, Kant's attempt to be a truly universal egalitarian made his egalitarianism more toothless than is commonly admitted. If we expect too much from Kant's moral philosophy it can ultimately only fail us. Conversely, only if we take this limitation seriously can we make room for more, and think about what else might be needed.

In this way, Kant's racism also brings to light just how strong Hegel's objection to Kant's formalism really was. As I have argued earlier, Hegel's discontent with Kant's formalism was essentially a discontent with the limited ambition of Kant's moral philosophy. While an appreciation of the limited ambitions of Kant's moral philosophy might provide a promising defense of Kant against some old as well as contemporary objections, Kant's racism can make us doubt whether we should be content with these limited ambitions. After all, we are now in a position to see vividly that how ambitious moral philosophy should be is not just a question of taste. If the limited ambitions of Kant's moral philosophy led to an account of ethical life that is seemingly compatible with the bigotry that we find in Kant's remarks, we might be tempted to conclude that his moral philosophy has ultimately fallen short of an important philosophical aim.

However, far from being a reductio, this shortfall poses a serious Kantian philosophical task. If, as I am inclined to think, Kant was on the right track in his analysis of moral judgment, and if this book's basic outline of his moral philosophy is correct, then we can neither overcome this shortfall by turning the Categorical Imperative into a method of ethics nor by rejecting Kant's ethics in favor of an entirely different moral outlook altogether. Moreover, an answer to the question what else

245

might be needed cannot dispense with the Kantian perspective, for the problem only comes fully in view through this very Kantian perspective.

Any answer to the question what else might be needed should not fit into a neat formula – neither in a mere conclusion nor perhaps in a single book – unless we want to risk repeating the problem of formal moral philosophy once again. But to pursue Kantian ethics without a deep appreciation of its limits would be to risk turning ethical ideals into collective neuroses: to expect something from Kant's framework it cannot provide, to create a moral discourse in our cultural repertoire that makes incomplete sense of our ethical lives and obscures the sources of self-entrapment from ourselves. Not just against individual forms of self-entrapment, but also against such collective neuroses, moral philosophy ought to be therapy.

Bibliography

Acton, H. B., 1970, *Kant's Moral Philosophy*, London: Macmillan

Allais, Lucy, 2016, 'Kant's Racism,' *Philosophical Papers* vol. 45, no. 1, pp. 1-36

Allison, Henry E., 1990, *Kant's Theory of Freedom*, Cambridge: Cambridge University Press

Allison, Henry E., 2011, *Kant's Groundwork for the Metaphysics of Morals: A Commentary*, Oxford: Oxford University Press

Althusius, Johannes, 1964, *The Politics of Johannes Althusius. An abridged translation of the Third Edition of* Politica Methodice Digesta, Atque Exemplis Sacris Et Profanis Illustrata, translated by Frederick S. Carney, Boston: Beacon Press

Altman, Matthew C., 2011, *Kant and Applied Ethics: The Use and Limits of Kant's Practical Philosophy*, Wiley-Blackwell

Ameriks, Karl, 1996, 'On Schneewind and Kant's Method in Ethics/ Sobre Schneewind y el Método de Kant en Etica,' *Ideas Y Valores*, vol. 102, pp. 28-53

Ameriks, Karl, 2000, *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy*, Cambridge: Cambridge University Press

Ameriks, Karl, 2012, "Kant, Human Nature, and History after Rousseau," In *Kant's Observations and Remarks: A Critical Guide*, edited by Susan Meld Shell and Richard Velkley, Cambridge: Cambridge University Press, pp. 247–65

Anderson, Georg, 1921, 'Die „Materie" in Kants Tugendlehre und der Formalismus der kritischen Ethik, *Kant-Studien* vol. 26, no. 1/2, pp. 289-311

Anderson, Georg, 1923, 'Kants Metaphysik der Sitten – ihre Idee und ihr Verhältnis zur Ethik der Wolffschen Schule,' *Kant-Studien* vol. 28, no. 1, pp. 41-61

Annas, Julia, 1993, *The Morality of Happiness*, Oxford University Press

Arendt, Hannah, 1992, *Lectures on Kant's Political Philosophy*, edited by Ronald Beiner, Chicago: University of Chicago Press

Aristotle, 1998, *Politics*, translated by C. D. C. Reeve, Indianapolis and Cambridge: Hackett

Augustine, 1998, *The City of God against the Pagans*, edited and translated by R. W. Dyson, Cambridge: Cambridge University Press

Aquinas, Thomas, 2007, *Commentary on Aristotle's* Politics, translated by Richard J. Regan, Indianapolis and Cambridge: Hackett

Ayer, A. J., 1972, 'On the Analysis of Moral Judgments,' in *Philosophical Essays*, London: Palgrave Macmillan, pp. 231-249

Bartuschat, Wolfgang, 2009, 'Kant über Philosophie und Aufklärung,' in *Kant und die Zukunft der europäischen Aufklärung*, edited by Heiner F. Klemme, Berlin and New York: de Gruyter

Basevich, Elvira, 2020, 'Reckoning with Kant on Race,' *Philosophical Forum*, vol. 51, no. 3, pp. 221-245

Baum, Manfred, 2017, 'Drei kategorische Imperative bei Kant,' in *Das Verhältnis von Recht und Ethik in Kants Praktischer Philosophie*, edited by Bernd Dörflinger, Dieter Hüning and Günter Kruck, Hildesheim: Georg Olms Verlag, pp. 131-152

BBC Radio 4, "Kant's Categorical Imperative" in *In Our Time*, with John Callanan, Alison Hills and David Oderberg, September 21st 2017, https://www.bbc.co.uk/programmes/b0952zl3

Beck, Lewis White, 1963, *A Commentary on Kant's Critique of Practical Reason*, Chicago and London: University of Chicago Press

Beitz, Charles R., 1989, *Political Equality: An Essay in Democratic Theory*, Princeton: Princeton University Press

Bell, David F., 1998, 'Pascal: Casuistry, Probability, Uncertainty,' *Journal of Medieval and Early Modern Studies* vol. 28, no. 1, pp. 37-50

Bentham, Jeremy, 1907, *An Introduction to the Principles of Morals and Legislation*, Oxford: Clarendon Press

Berg, Anastasia N. A., 2020, 'Kant on moral self-opacity,' *European Journal of Philosophy* vol. 28, no. 3, pp. 567-585

Bernasconi, Robert, 2001, 'Who Invented the Concept of Race?' in *Race*, edited by Robert Bernasconi, Oxford: Blackwell, pp. 11-36

Bernasconi, Robert, 2002, 'Kant as an Unfamiliar Source of Racism,' in *Philosophers on Race: Critical Essays*, edited by Julie K. Ward and Tommy L. Lott, Oxford: Blackwell, pp. 145-166

Bernasconi, Robert, 2006, 'Kant and Blumenbach's Polyps: A Neglected Chapter in the History of the Concept of Race,' in *The German Invention of Race*, edited by Sara Eigen and Mark Larrimore, Albany: SUNY Press, pp. 73-90

Bernasconi, Robert, 2010, 'The Place of Race in Kant's *Physical Geography* and in the Writings of the 1790s,' in *Rethinking Kant vol. 2*, edited by Pablo Muchnik, Newcastle: Cambridge Scholars Publishing, pp. 274-290

Bielefeldt, Heiner, 2001, *Symbolic Representation in Kant's Practical Philosophy*, Cambridge: Cambridge University Press

Birken-Bertsch, Hanno, 2015, "Reflexion," in *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr and Stefano Bacin, Berlin: de Gruyter, pp. 1915-1917

Boxill, Bernard, 2017, 'Kantian Racism and Kantian Teleology,' in *Philosophy and Race*, edited by Naomi Zack, Oxford: Oxford University Press, pp. 44-53

Brandt, Reinhard, 2018, 'Was ist Aufklärung? Beobachtungen zur Kantischen Antwort,' *Kant-Studien,* vol. 109, no. 1, pp. 147-169

Brender, Natalie, 2004, 'What Is Disorientation in Thinking?' in *New Essays on the History of Autonomy: A Collection Honoring J. B. Schneewind*, edited by Natalie Brender and Larry Krasnoff, Cambridge: Cambridge University Press, pp. 154-180

Brewer, Talbot, 2002, 'Maxims and Virtues,' *The Philosophical Review* vol. 111, no. 4, pp. 539-572

Burlamaqui, Jean Jacques, 1747, *Principes du Droit Naturel*, Geneva: Barrillot & Fils

Busch, Werner, 1979, *Die Entstehung der kritischen Rechtsphilosophie Kants, 1762-1780*, Berlin and New York: De Gruyter

Butler, Joseph, 2017, *Fifteen Sermons Preached at the Rolls Chapel and other writings on ethics*, edited by David McNaughton, Oxford: Oxford University Press

Callanan, John J., 2019, 'Kant on Misology and the Natural Dialectic,' *Philosophers' Imprint* vol. 19, no. 47, pp. 1-22

Carraud, Vincent and Olivier Chaline, 1996, 'Casuistes et casuistique à l'époque moderne,' in *Dictionnaire d'éthique et de philosophie morale*, edited by Monique Canto-Sperber, Paris: Presses universitaires de France, pp. 213-222

Cassirer, Ernst, 1983, *Kant's Life and Thought*, translated by James Haden, New Haven: Yale University Press

Cicero, Marcus Tullius, 1991, *On Duties*, edited and translated by Miriam T. Griffin and E. Margaret Atkins, Cambridge: Cambridge University Press

Crusius, Christian August, 1767, *Anweisung vernünftig zu leben, Darinnen nach Erklärung der Natur des menschlichen Willens die natürlichen Pflichten und allgemeinen Klugheitslehren im richtigen Zusammenhange vorgetragen werden*, Leipzig: Johann Friedrich Gleditsch

De Franceschi, Sylvio Hermann, 2020, 'Le basculement de la théologie morale catholique allemande: Laxisme, probabilisme et casuistique indulgente dans le Saint-Empire du XVIIIᵉ siècle prékantien,' in *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 91-141

Denis, Lara, 2007, 'Kant's Formula of the End in itself: Some Recent Debates,' *Philosophy Compass* vol. 2, no. 2, pp. 244-257

Di Giulio, Sara (editor), 2020, *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited with Albert Frigo, Berlin and Boston: de Gruyter

Di Giulio, Sara and Alberto Frigo, 2020, 'Einführung,' in *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 1-15

Di Giulio, Sara, 2020, 'Video meliora proboque, deteriora sequor : Zur Irreführung des Gewissens bei Kant,' in *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 233-287

Dietrichson, Paul, 1964, 'When is a Maxim Fullly Universalizable?' *Kant-Studien* vol. 55, no. 2, pp. 143-170

Dietrichson, Paul, 1969, 'Kant's Criteria of Universalizability' in *Kant: Foundations of the Metaphysics of Morals: Text and Critical Essays*, edited by Robert Paul Wolff, Indianapolis: Bobbs-Merrill Company, pp. 230-244

Dörflinger, Bernd, 2013, 'Ethische Methodenlehre: Didaktik und Asketik (TL 6:477-485),' in *Kant's "Tugendlehre." A Comprehensive Commentary*, edited by Andreas Trampota, Oliver Sensen and Jens Timmermann, Berlin/Boston: de Gruyter, pp. 383-410

Ebbinghaus, Julius, 1967, 'Interpretation and Misinterpretation of the Categorical Imperative,' in *Kant: A Collection of Critical Essays*, edited by Robert Paul Wolff, New York: Anchor Books, pp. 211-227

Ebels-Duggan, Kyla, 2011, 'Kantian Ethics,' in *The Continuum Companion to Ethics*, edited by Christian Miller, London/New York: Continuum, pp. 168-189

Eberl, Oliver, 2019, 'Kant on Race and Barbarism: Towards a More Complex View on Racism and Anti-Colonialism in Kant,' *Kantian Review*, vol. 24, no. 3, pp. 385-413

Eigen, Sara and Mark Larrimore, 2006, *The German Invention of Race*, Albany: SUNY Press

Elden, Stuart and Eduardo Mendieta (eds.), 2011, *Reading Kant's Geography*, Albany: SUNY Press

Elsenhans, Theodor, 1904, *Kant's Rassentheorie und ihre bleibende Bedeutung. Ein Nachtrag zur Kant-Gedächtnisfeier*, Leipzig: Willhelm Engelmann

Engstrom, Stephen, 2002, 'Kant's Distinction between Theoretical and Practical Knowledge,' *Harvard Review of Philosophy* vol. 10, no. 1, pp. 49-63

Engstrom, Stephen, 2009, *The Form of Practical Knowledge: A Study of the Categorical Imperative*, Cambridge, MA: Harvard University Press

Engstrom, Stephen, 2015, 'Reflection and Reason in Hume and Kant,' *Hegel Bulletin* vol. 36, no. 1, pp. 15-32

Engstrom, Stephen, 2018, 'Virtue and Vice in Aristotle and Kant,' in *Evil in Aristotle*, edited by Pavlos Kontos, Cambridge: Cambridge University Press, pp. 222-239

Epicurus, 1887, *Epicurea*, edited by Hermann Usener, Leipzig: Teubner

Esser, Andrea, 2008, 'Kant on Solving Moral Conflicts,' in *Kant's Ethics of Virtue*, edited by Monika Betzler, Berlin and New York: de Gruyter, pp. 279-302

Eze, Emmanuel Chukwudi, 1995, 'The Color of Reason: The Idea of "Race" in Kant's Anthropology,' *The Bucknell Review* vol. 38, no. 2, pp. 200-241

Field, G. C., 1932, 'Kant's First Moral Principle,' *Mind* vol. 41, no. 161, pp. 17-36

Flatt, Johann Friedrich, 1786, 'Grundlegung zur Metaphysik der Sitten von Immanuel Kant,' *Tübingische Gelehrte Anzeigen* vol. 14, February 16th, pp. 105-112

Fleischacker, Samuel, 1991, 'Philosophy in Moral Practice: Kant and Adam Smith,' *Kant-Studien* vol. 82, no. 3, pp. 249-269

Flikschuh, Katrin, 2007, 'Kant's Indemonstrable Postulate of Right: A Response to Paul Guyer,' *Kantian Review* vol. 12, no. 1, pp. 1-39

Flikschuh, Katrin and Lea Ypi (eds.), 2014, *Kant and Colonialism: Historical and Critical Perspectives*, Oxford: Oxford University Press

Flikschuh, Katrin, 2018, 'Philosophical Racism,' *Proceedings of the Aristotelian Society, Supplementary Volume*, 91:1, 91-110

Forkl, Markus, 2001, *Kants System der Tugendpflichten*, Frankfurt am Main: Peter Lang

Formosa, Paul, 2014, 'Dignity and Respect: How to Apply Kant's Formula of Humanity,' *Philosophical Forum* vol. 45, no. 1, pp. 49-68

Forschler, Scott, 2013, 'Two Dogmas of Kantian Ethics,' *Journal of Value Inquiry* vol. 47, no. 3, pp. 255-269

Forster, Georg, 1958-, *Georg Forsters Werke*, edited by the Deutsche Akademie der Wissenschaften, Berlin: Akademie Verlag

Forster, Michael N., 2008, *Kant and Skepticism*, Princeton and Oxford: Princeton University Press

Forster, Michael N., 2015, 'Ideology,' in *The Oxford Handbook of German Philosophy in the Nineteenth Century*, edited by Michael N. Forster and Kristin Gjesdal, Oxford: Oxford University Press, pp. 806-827

Freyenhagen, Fabian, 2011, 'Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection,' *Bulletin of the Hegel Society of Great Britain* vol. 63, no. 1-2, pp. 163-186

Freyenhagen, Fabian, 2012, 'The Empty Formalism Objection Revisited: §135R and Recent Kantian Responses,' in *Hegel's Philosophy of Right: Essays on Ethics, Politics, and Law*, edited by Thom Brooks, London: Blackwell, pp. 43-72

Frigo, Alberto (editor), 2020, *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited with Sara Di Giulio, Berlin and Boston: de Gruyter

Galvin, Richard, 2009, 'The Universal Law Formulas,' in *The Blackwell Guide to Kant's Ethics*, edited by Thomas E. Hill, Oxford: Wiley-Blackwell, pp. 52-82

Geiger, Ido, 2010, 'What is the Use of the Universal Law Formula of the Categorical Imperative?' *British Journal for the History of Philosophy* vol. 18, no. 2, pp. 271-295

Geiger, Ido, 2015, 'How Do We Acquire Moral Knowledge? Is Knowing Our Duty Ever Passive? – Two Questions for Martin Sticker,' *British Journal for the History of Philosophy*, vol. 23, no. 5, pp. 990-997

Geiger, Ido, 2020, 'Humanity and personality – what, for Kant, is the source of moral normativity?' *Inquiry* pp. 1-24

Gerhardt, Volker, 2009, 'Die Menschheit in der Person des Menschen. Zur Anthropologie der menschlichen Würde bei Kant,' in *Kant und die Zukunft der europäischen Aufklärung*, edited by Heiner F. Klemme, Berling and New York: de Gruyter, pp. 269-291

Girtanner, Christopher, 1796, *Über das Kantische Prinzip für die Naturgeschichte*, Göttingen: Vandenhoefund Ruprecht

Glasgow, Joshua M., 2003, 'Expanding the Limits of Universalization: Kant's Duties and Kantian Moral Deliberation,' *Canadian Journal of Philosophy* vol. 33, no. 1, 23-47

Glass, Ronald, 1971, 'The Contradictions in Kant's Examples,' *Philosophical Studies* vol. 22, no. 5/6, pp. 65-70

Gorodeisky, Keren, 2021, "Re!ection (Re!exion, Überlegung)" in *The Cambridge Kant Lexicon*, edited by Julian Wuerth, Cambridge: Cambridge University Press, pp. 374-377

Gregor, Mary J., 1963, *Laws of Freedom: A Study of Kant's Method of Applying the Categorical Imperative in the Metaphysics of Morals*, New York: Barnes & Noble

Gregor, Mary J., 1991, 'Translator's Introduction,' in *Immanuel Kant: The Metaphysics of Morals*, Cambridge: Cambridge University Press

Grenberg, Jeanine, 2013, *Kant's Defense of Common Moral Experience: A Phenomenological Account*, Cambridge: Cambridge University Press

Griffin, James, 1996, *Value Judgment: Improving our Ethical Beliefs*, Oxford: Clarendon Press

Grotius, Hugo, 2005, *The Rights of War and Peace*, edited by Richard Tuck, from the edition translated by Jean Barbeyrac, Indianapolis: Liberty Fund

Guyer, Paul, 2005a, 'Kant's System of Duties,' in Paul Guyer, *Kant's System of Nature and Freedom: Selected Essays*, Oxford: Oxford University Press

Guyer, Paul, 2005b, 'Kant's Deductions of the Principles of Right' in Paul Guyer, *Kant's System of Nature and Freedom: Selected Essays*, Oxford: Oxford University Press

Hamann, Johann Georg, 1965, *Briefwechsel. Fünfter Band,* edited by Arthur Denkel, Frankfurt am Main: Insel

Harrison, Jonathan, 1957, 'Kant's Examples of the First Formulation of the Categorical Imperative,' *Philosophical Quarterly* vol. 7, no. 26, pp. 50-62. Reprinted in *Kant: A Collection of Critical Essays*, edited by Robert Paul Wolff, New York: Anchor Books, 1967, pp. 228-245

Hedrick, Todd, 2008, 'Race, Difference and Anthropology in Kant's Cosmopolitanism,' *Journal of the History of Philosophy*, vol. 46, no. 2, pp. 245-268

Hegel, Georg Wilhelm Friedrich, 2018, *The Phenomenology of Spirit*, translated and edited by Terry Pinkard, Cambridge: Cambridge University Press

Hegel, Georg Wilhelm Friedrich, 1991, *Elements of the Philosophy of Right*, translated by H. B. Nisbet, edited by Allen Wood, Cambridge: Cambridge University Press

Hegel, Georg Wilhelm Friedrich, 1999, 'On the Scientific Ways of Treating Natural Law, on its Place in Practical Philosophy, and its Relation to the Positive Sciences of Right,' (first published 1802-3), translated by H. B. Nisbet, in *Hegel: Political Writings*, edited by Lawrence Dickey and H. B. Nisbet, New York: Cambridge University Press, pp. 102-180

Heineccius, Johann Gottlieb, 1764, *Elementa Juris Naturae et Gentium, Commoda Auditoribus Methodo adornata*, Venice: Balleoniana

Heineccius, Johann Gottlieb, 2008, *A Methodical System of Universal Law: Or, the Laws of Nature and Nations, with Supplements and a Discourse by George Turnbull*, translated by George Turnbull, edited by Thomas Ahnert and Peter Schröder, Indianapolis: Liberty Fund

Henrich, Dieter, 1963, 'Das Problem der Grundlegung der Ethik bei Kant und im spekulativen Idealismus,' in *Sein und Ethos*, edited by Paulus Engelhardt OP, Mainz: Matthias-Grünewald Verlag

Henrich, Dieter, 1992, *Aesthetic Judgment and the Moral Image of the World: Studies in Kant*, Palo Alto: Stanford University Press

Herb, Karlfriedrich, 2018, 'Unter Bleichgesichtern. Kant's Kritik der kolonialen Vernunft,' *Zeitschrift für Politik*, vol. 65, no. 4, pp. 381-398

Herman, Barbara, 1989, 'Murder and Mayhem: Violence and Kantian Casuistry,' *The Monist* vol. 72, no. 3, pp. 411-431

Herman, Barbara, 1993, *The Practice of Moral Judgment*, Cambridge, Mass: Harvard University Press

Herman, Barbara, 2011, 'Embracing Kant's Formalism,' *Kantian Review* vol. 16, no. 1, pp. 49-66

Hill, Thomas E. and Bernard Boxill, 2001, 'Kant and Race,' in *Race and Racism*, edited by Bernard Boxill, Oxford: Oxford University Press, pp. 448-471

Hill, Thomas E., 2007, 'Kantian Normative Ethics,' in *The Oxford Handbook of Ethical Theory*, edited by David Copp, Oxford: Oxford University Press, pp. 480-514

Hill, Thomas E., 2010, 'Kant's *Tugendlehre* as normative ethics,' in *Kant's Metaphysics of Morals: A Critical Guide*, edited by Lara Denis, Cambridge: Cambridge University Press, pp. 234-255

Höffe, Otfried, 1977, 'Kants kategorischer Imperativ als Kriterium des Sittlichen,' *Zeitschrift für philosophische Forschung* vol. 31, no. 3, pp. 354-384

Höffe, Otfried, 1994, *Immanuel Kant*, translated by Marshall Farrier, Albany: State University of New York Press

Höffe, Otfried, (editor) 1999, *Immanuel Kant. Metaphysische Anfangsgründe der Rechtslehre*, Berlin: Akademie Verlag

Höffe, Otfried, 2020, 'Kasuistische Fragen in der *Tugendlehre* und das Gewissen (TL §§9-18),' in *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 163-176

Horkheimer, Max, and Theodor W. Adorno, 2002, *Dialectic of Enlightenment: Philosophical Fragments*, translated by Edmund Jephcott, edited by Gunzelin Schmid Noerr, Stanford CA: Stanford University Press

Hoy, David Couzens, 1989, 'Hegel's Critique of Kantian Morality,' *History of Philosophy Quarterly* vol. 6, no. 2, pp. 207-232

Huber, Ulric, 1735, *Praelectionum Juris Civilis*, Leipzig: Gleditsch & Söhne

Hume, David, 2007, *Dialogues concerning Natural Religion and Other Writings*, edited by Dorothy Coleman, Cambridge: Cambridge University Press

Ilting, Karl-Heinz, 1981, 'Gibt es eine kritische Ethik und Rechtsphilosophie Kants?' *Archiv für Geschichte der Philosophie*, vol. 63, no. 3, pp. 325-345

Iwasa, Noriaki, 2013, 'Reason Alone Cannot Identify Moral Laws,' *Journal of Value Inquiry*, vol. 47, no. 1-2, pp. 67-85

James, David N., 1992, 'Twenty Questions: Kant's Applied Ethics,' *The Southern Journal of Philosophy*, vol. 30, no. 3, pp. 67-87

Jauch, Ursula Pia, 1988, *Immanuel Kant zur Geschlechterdifferenz. Aufklärische Vorteilskritik und bürgerliche Geschlechtsvormundschaft*, Wien: Passagen Verlag

Joerden, Jan C., 1993, 'Was leisten Kants Beispiele bei der Anwendung des Kategorischen Imperativs? Zugleich eine Besprechung von: Christian Schnoor, "Kants Kategorischer Imperativ als Kriterium der Richtigkeit des Handelns,"' *Archiv für Rechts- und Sozialphilosophie* vol. 79, no. 2, pp. 257-258

Jonsen, Albert R. and Stephen Toulmin, 1988, *The Abuse of Casuistry: A History of Moral Reasoning*, Berkeley: University of California Press

Kant, Immanuel, 1785, *Grundlegung zur Metaphysik der Sitten*, Riga: Johann Friedrich Hartknoch

Kant, Immanuel, 1838, *Immanuel Kant's Grundlegung zur Metaphysik der Sitten, Kritik der praktischen Vernunft nebst den Abhandlungen zur Philosophie der Geschichte*, edited and with an introduction by Gustav Hartenstein, Leipzig: Modes und Baumann

Kant, Immanuel, 1895, *Fundamental Principles of the Metaphysic of Ethics*, translated by Thomas K. Abbott, London and New York: Longmans, Green and Co.

Kant, Immanuel, 1959, *Kritik der praktischen Vernunft*, edited by Karl Vorländer, Hamburg: Meiner

Kant, Immanuel, 2002, *Critique of Practical Reason*, translated by Werner S. Pluhar, Cambridge/Indianapolis: Hackett

Kant, Immanuel, 2002, *Groundwork of the Metaphysics of Morals*, edited and translated by Allen Wood, New Haven and London: Yale University Press

Kant, Immanuel, 2011, *Groundwork of the Metaphysics of Morals. A German-English Edition*, edited by Jens Timmermann, translated by Mary Gregor, revised by Jens Timmermann, Cambridge: Cambridge University Press

Kaufmann, Matthias, 2019, 'Wie gleich sind Personen – und Menschen? Kant über Geschlechter, Rassen und Kolonisierung,' *Jahrbuch für Recht und Ethik* vol. 27, pp. 183-204

Kemp, John, 1958, 'Kant's Examples of the Categorical Imperative' *The Philosophical Quarterly* vol. 8, no. 30, pp. 63-71. Reprinted in *Kant: A Collection of Critical Essays*, edited by Robert Paul Wolff, New York: Anchor Books, 1967, pp. 246-258

Kerstein, Samuel J., 2002, 'Introduction: Derivation, Deduction, and the Supreme Principle of Morality,' in *Kant's Search for the Supreme Principle of Morality*, Cambridge: Cambridge University Press, pp. 1-15

Kerstein, Samuel, 2009, 'Treating Others Merely as Means,' *Utilitas*, vol. 21, no. 2, pp. 163-180

Kerstein, Samuel, 2013, *How to Treat Persons*, Oxford: Oxford University Press

Kim, Soo Bae, 2009, 'The Formation of Kant's Casuistry and Method Problems of Applied Ethics,' *Kant-Studien* vol. 100, no. 3, pp. 332-345

Kisilevsky, Sara and Martin J. Stone (editors), 2017, *Freedom and Force: essays on Kant's Legal Philosophy*, Oxford: Hart Publishing

Kitcher, Patricia, 2004, 'Kant's Argument for the Categorical Imperative,' *Noûs* vol. 38, no. 4, pp. 555-584

Kittsteiner, H. D., 1988, 'Kant and casuistry,' in *Conscience and Casuistry in Early Modern Europe*, edited by Edmund Leites, Cambridge: Cambridge University Press, pp. 185-213

Kleingeld, Pauline, 2007, 'Kant's Second Thoughts on Race,' *The Philosophical Quarterly* vol. 57, no. 229, pp. 573-592

Kleingeld, Pauline, 2014, 'Kant's Second Thoughts on Colonialism,' in *Kant and Colonialism: Historical and Critical Perspectives*, edited by Katrin Flikschuh and Lea Ypi, Oxford: Oxford University Press, pp. 43-67

Kleingeld, Pauline, 2019, 'On Dealing with Kant's Sexism and Racism,' *SGIR Review* vol. 2, no. 2, pp. 3-22

Kleingeld, Pauline, 2017, 'Contradiction and Kant's Formula of Universal Law,' *Kant-Studien* vol. 108, no. 1, pp. 89-115

Kleingeld, Pauline, 2019, 'A Contradiction of the Right Kind: Convenience Killing and Kant's Formula of Universal Law,' *Philosophical Quarterly* vol. 69, no. 274, pp. 64-81

Kleingeld, Pauline, 2019, 'On Dealing with Kant's Sexism and Racism,' *SGIR Review* vol. 2, no. 2, pp. 3-22

Kleingeld, Pauline, 2020, 'How to Use Someone 'Merely as a Means',' *Kantian Review* vol. 25, no. 3, pp. 389-414

Klemme, Heiner, 'Moralized Nature, Naturalized Autonomy,' in *Kant on Moral Autonomy*, edited by Oliver Sensen, Cambridge: Cambridge University Press

Knebel, Sven, 2000, *Wille, Würfel und Wahrscheinlichkeit: Das System der moralischen Notwendigkeit in der Jesuitenscholastik 1550-1700*, Hamburg: Meiner

Knox, T. M., 1957, 'Hegel's Attitude to Kant's Ethics,' *Kant-Studien* vol. 49, pp. 70-81

Kobusch, Theo, 1993, *Die Entdeckung der Person. Metaphysik der Freiheit und modernes Menschenbild*, Darmstadt: Wissenschaftliche Buchgesellschaft

Koch, Lutz, 2003, *Kants ethische Didaktik*, Würzburg: Ergon Verlag

Kofman, Sarah, 1982, *Le Respect des femmes*, Paris: Galilée

Korsgaard, Christine M., 1996a, *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press

Korsgaard, Christine M., 1996b, *The Sources of Normativity*, Cambridge: Cambridge University Press

Korsgaard, Christine, 1998, 'Introduction,' in *Immanuel Kant: Groundwork of the Metaphysics of Morals*, translated and edited by Mary Gregor, Cambridge: Cambridge University Press, pp. vii-xxx

Korsgaard, Christine, 2008, *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*, Oxford and New York: Oxford University Press

Krailsheimer, A. J., 1967, 'Introduction' to *Pascal: The Provincial Letters*, translated by A. J. Krailsheimer, London: Penguin

Krasnoff, Larry, 2004, 'Pythagoras Enlightened: Kant on the Effect of Moral Philosophy,' in *New Essays on the History of Autonomy: A Collection Honoring J. B. Schneewind*, edited by Natalie Brender and Larry Krasnoff, Cambridge: Cambridge University Press, pp. 133-153

Kuehn, Manfred, 2001a, *Kant. A Biography*, Cambridge: Cambridge University Press

Kuehn, Manfred, 2001b, 'Kant and Cicero,' in *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses*, vol. III, edited by Volker Gerhardt, Rolf-Peter Horstmann and Ralph Schumacher, Berlin: de Gruyter, pp. 270-278

Kuehn, Manfred, 2010, 'Kant's *Metaphysics of Morals*: the history and significance of its deferral,' in *Kant's Metaphysics of Morals: A Critical Guide*, edited by Lara Denis, Cambridge: Cambridge University Press, pp. 9-27

Lagier, Raphaël, 2004, *Les races humaines selon Kant*, Paris: Presses universitaires de France

Langton, Rae, 1992, 'Duty and Desolation,' *Philosophy* vol. 67, no. 262, pp. 481-505

Langton, Rae, 2012, 'Maria von Herbert's Challenge to Kant,' in *Ethics: Essential Readings in Moral Theory*, edited by George Sher, New York and London: Routledge, pp. 377-386

Larrimore, Mark, 1999, 'Sublime Waste: Kant on the Destiny of the 'Races',' *Canadian Journal of Philosophy*, vol. 29, pp. 99-125

Larrimore, Mark, 2006, 'Race, Freedom and the Fall in Steffens and Kant' in *The German Invention of Race*, edited by Sara Eigen and Mark Larrimore, Albany: SUNY Press, pp. 91-120

Laurence, Ben, 2015, 'Juridical laws as moral laws in Kant's *The Doctrine of Right*,' in *Reasons and Intentions in Law and Practical Agency*, edited by George Pavlakos and Veronica Rodriguez-Blanco, Cambridge: Cambridge University Press

Laurence, Ben, 2018, 'Kant on Strict Right,' *Philosophers' Imprint* vol. 18, no. 4, pp. 1-22

Leutgöb, Andreas, 2015, *Der Rassismus und seine wissenschaftlichen Wurzeln bei Immanuel Kant. Eine Zeitreise in die Epoche der Aufklärung und seiner Rassismusdebatte*, Norderstedt: Books on Demand

Lo, Ping-cheung, 1981, 'A Critical Reevaluation of the Alleged "Empty Formalism" of Kantian Ethics,' *Ethics* vol. 91, no. 2, pp. 181-201

Locke, John, 1997, *An Essay Concerning Human Understanding*, edited by Roger Woolhouse, London: Penguin

Lottenbach, Hans and Sergio Tenenbaum, 1995, 'Hegel's Critique of Kant in the Philosophy of Right,' *Kant-Studien* vol. 86, no. 2, pp. 211-230

Ludwig, Bernd, 2013, 'Die Einteilungen der Metaphysik der Sitten im Allgemeinen und die der Tugendlehre im Besonderen (MS 6:218–221 und RL 6:239–242 und TL 6:388– 394, 410–413),' in *Kant's Tugendlehre: A Comprehensive Commentary*, edited by Andreas Trampota, Oliver Sensen and Jens Timmermann, Berlin: de Gruyter, pp. 59-84

Ludwig, Bernd, 2015a, 'Sympathy for the Devil(s)? Personality and Legal Coercion in Kant's Doctrine of Law,' *Jurisprudence* vol. 6, no. 1, pp. 25-44

Ludwig, Bernd, 2015b, 'Metaphysische Anfangsgründe der Rechtslehre (Die Metaphysik der Sitten, Erster Teil),' in *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr and Stefano Bacin, Berlin: de Gruyter, pp. 1552-1562

Ludwig, Bernd, 2020, *Aufklärung über die Sittlichkeit: Zu Kant's Grundlegung einer Metaphysik der Sitten*, Frankfurt am Main: Vittorio Klostermann

Ludwig, Ralf, 1991, *Kategorischer Imperativ und Metaphysik der Sitten: die Frage nach der Einheitlichkeit von Kants Ethik*, Frankfurt am Main: Lang

MacIntyre, Alasdair, 1998, *A Short History of Ethics. A History of Moral Philosophy from the Homeric Age to the Twentieth Century*, second edition, first published 1967, London: Routledge

Mahon, James Edwin, 2006, 'Kant and Maria von Herbert: Reticence vs. Deception,' *Philosophy* vol. 81, no. 317, pp. 417-444

Malter, Rudolf, 1990, 'Der Rassebegriff in Kants Anthropologie,' in *Die Natur des Menschen: Probleme der physischen Anthropologie und Rassenkunde (1750-1850)*, edited by Gunter Mann, Stuttgart and New York: Gustav Fischer Verlag, pp. 113-123

Marshall, John, 1994, *John Locke: Resistance, Religion and Responsibility*, Cambridge: Cambridge University Press

Matson, Wallace I., 1954, 'Kant as Casuist,' *Journal of Philosophy* vol. 51, no. 25, pp. 855-860

McCabe, David, 2019, 'Kant Was a Racist: Now What?' *APA Newsletter on Teaching Philosophy* vol. 18, no. 2, pp. 2-9

McCarthy, Thomas, 2009, *Race, Empire, and the Idea of Human Development*, Cambridge: Cambridge University Press

McGrath, Sarah, 2014, 'On Hasting Rashdall's "The Limits of Casuistry,"' *Ethics,* vol. 125, no. 1, pp. 196-199

McNair, Ted, 2000, 'Universal Necessity and Contradictions in Conception,' *Kant-Studien* vol. 91, no. 1, pp. 25-43

Mensch, Jennifer, 2017, 'Caught Between Character and Race: 'Temperament' in Kant's Lectures on Anthropology,' *Australian Feminist Law Journal*, vol. 43, no. 1, pp. 125-144

Merritt, Melissa, 2015, 'Varieties of Reflection in Kant's Logic,' *British Journal for the History of Philosophy*, vol. 23, no. 3, pp. 478-501

Merritt, Melissa, 2018, *Kant on Reflection and Virtue*, Cambridge: Cambridge University Press

Mikkelsen, Jon M., 2013, 'Translator's Introduction: Recent Work on Kant's Race Theory / The Texts / The Translations,' in *Kant and the Concept of Race: Late Eighteenth-Century Writings*, translated and edited by Jon M. Mikkelsen, Albany: SUNY Press, pp. 1-40

Mill, John Stuart, 2001, *Utilitarianism* and the *1868 Speech on Capital Punishment*, edited by George Sher, Indianapolis and Cambridge: Hackett

Mills, Charles W., 2005, 'Kant's Untermenschen,' in *Race and Racism in Modern Philosophy*, edited by Andrew Valls, Ithaca NY: Cornell University Press, pp. 169-193. Reprinted in Charles Mills, 2017, *Black Rights/White Wrongs: The Critique of Racial Liberalism*, Oxford: Oxford University Press, pp. 91-135

Mills, Charles W., 2014, 'Kant and Race, Redux,' *Graduate Faculty Philosophy Journal* vol. 35, no. 1, pp. 125-157

Munzel, G. Felicitas, 1999, *Kant's Conception of Moral Character The "Critical" Link of Morality, Anthropology, and Reflective Judgment*, Chicago and London: University of Chicago Press

Nance, Michael, 2012, 'Kantian right and the categorical imperative: Response to Willaschek' *International Journal of Philosophical Studies* vol. 20, no. 4, pp. 541-556

Neiman, Susan, 1994, *The Unity of Reason. Rereading Kant*, New York and Oxford: Oxford University Press

Nenon, Thomas, 1993, 'Freedom, Responsibility, Character: Some Reflections on Kant's Notion of the Person,' *Jahrbuch für Recht und Ethik*, vol. 1, pp. 159-168

Nietzsche, Friedrich, 2001, *The Gay Science. With a Prelude in German Rhymes and an Appendix of Songs*, edited by Bernard Williams, translated by Josefine Nauckhoff, Cambridge University Press

Nussbaum, Martha C., 1980, 'Shame, Separateness, and Political Unity: Aristotle's Criticism of Plato,' in *Essays on Aristotle's Ethics*, edited by Amélie Oksenberg Rorty, Berkeley and Los Angeles: University of California Press, pp. 395-435

Nussbaum, Martha C., 1997, 'Kant and Stoic Cosmopolitanism,' *Journal of Political Philosophy*, vol. 5, no. 1, pp. 1-25

Nussbaum, Martha C., 2018, *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*, Princeton University Press

O'Hagan, Timothy, 1987, 'On Hegel's Critique of Kant's Moral and Political Philosophy,' in *Hegel's Critique of Kant*, edited by Stephen Priest, Oxford: Clarendon Press, pp. 135-160

O'Neill, Onora, 1983, 'Kant after Virtue,' *Inquiry* vol. 26, no. 4, pp. 387-405

O'Neill, Onora, 2013, *Acting on Principle*, Cambridge: Cambridge University Press (first published as Onora Nell, *Acting on Principle*, 1975, Columbia University Press)

O'Neill, Onora, 1991, 'Kantian Ethics,' in *A Companion to Ethics*, edited by Peter Singer, Oxford: Blackwell, pp. 175-185

O'Neill, Onora, 1998, 'Consistency in Action,' in *Ethical Theory 2: Theories About How We Should Live*, edited by James Rachels, Oxford: Oxford University Press, pp. 256-281

O'Neill, Onora, 2002, 'Instituting Principles: Between Duty and Action,' in *Kant's Metaphysics of Morals: Interpretive Essays*, edited by Mark Timmons, Oxford: Oxford University Press, pp. 331-347

Oberhausen, Michael and Riccardo Pozzo (eds.), 1999, *Vorlesungsverzeichnisse der Universität Königsberg (1720-1804)*, 2 volumes, Stuttgart-Bad Cannstatt: frommann-holzboog

Papadaki, Lina, 2016, 'Treating Others Merely as Means: A Reply to Kerstein,' *Utilitas* vol. 28, no. 1, pp. 73-100

Papish, Laura, 2018, *Kant on Evil, Self-Deception, and Moral Reform*, Oxford University Press

Paton, Herbert James, 1948, *The Categorical Imperative: A Study in Kant's Moral Philosophy*, Tiptree, UK: Anchor Press

Patrone, Tatiana, 2013, 'Making Sense of Kant's Casuistry,' in *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses. Band 3*, edited by Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca and Margit Ruffing, Berlin/Bosten: de Gruyter, pp. 483-494

Patrone, Tatiana, 2018, 'Treating Others as Means: But Not Merely as Means: Re-Reading Kant's Formula of Humanity,' *Ethical Perspectives* vol. 25, no. 1, pp. 61-86

Paulsen, Friedrich, 1899, *Immanuel Kant. Sein Leben und seine Lehre*, Stuttgart: Frommanns

Pippin, Robert B., 1982, *Kant's Theory of Form: An Essay on the Critique of Pure Reason*, New Haven and London: Yale University Press

Pippin, Robert B., 1989, *Hegel's Idealism: The Satisfaction of Self-Consciousness*, Cambridge: Cambridge University Press

Pippin, Robert B., 1999, 'Dividing and Deriving in Kant's *Rechtslehre*,' in *Metaphysische Anfangsgründe der Rechtslehre*, edited by Ottfried Höffe, Berlin: Akademie Verlag

Pippin, Robert B., 2013, 'Reason's Form,' in *The Impact of Idealism: The Legacy of Post-Kantian German Thought. Volume I Philosophy and Natural Sciences*, edited by Karl Ameriks, Cambridge: Cambridge University Press, pp. 373-394

Pogge, Thomas W., 2002, 'Is Kant's Rechtslehre a 'Comprehensive Liberalism'?' in *Kant's Metaphysics of Morals*, edited by Mark Timmons, Oxford: Oxford University Press

Potter, Nelson, 1993, 'What is Wrong with Kant's Four Examples,' *Journal of Philosophical Research* vol. 18, pp. 213-229

Price, Richard, 1787, *A Review of the Principal Questions in Morals*, London: Cadell in the Strand

Pufendorf, Samuel, 1729, *Of the Law of Nature and of Nations*, translated by Basil Kenneth, with notes and a prefatory discourse by Jean Barbeyrac, London: J. Walthoe et al.

Pufendorf, Samuel, 1991, *On the Duty of Man and Citizen According to Natural Law*, translated by Michael Silverthorne, edited by James Tully, Cambridge: Cambridge University Press

Pütz, Peter, 1978, *Die Deutsche Aufklärung*, Darmstadt: Wissenschaftliche Buchgesellschaft

Ramsauer, Laurenz, 2023, 'Kant's Racism as a Philosophical Problem,' *Pacific Philosophical Quarterly*, vol. 104, no. 4, pp. 791-815

Ramsauer, Laurenz, 2024, 'Kant's Derivation of Imperatives of Duty,' *Kantian Review*

Rashdall, Hastings, 1894, 'The Limits of Casuistry,' *International Journal of Ethics*, vol. 4, no. 4, pp. 459-480

Rawls, John, 2000, *Lectures on the History of Moral Philosophy*, edited by Barbara Herman, Cambridge, MA: Harvard University Press

Reath, Andrews, 1989, 'The Categorical Imperative and Kant's Conception of Practical Rationality,' *The Monist* vol. 27, no. 3, pp. 384-410

Reath, Andrews, 1994, 'Legislating the Moral Law,' *Noûs* vol. 28, no. 4, pp. 435-464

Reid, Thomas, 2010, *Essays on the Active Powers of Man*, edited by Knud Haakonssen and James A. Harris, Edinburgh: Edinburgh University Press

Ripstein, Arthur, 2004, 'Authority and Coercion,' *Philosophy and Public Affairs* 32:1, pp. 2-35

Ripstein, Arthur, 2009, *Force and Freedom: Kant's Legal and Political Philosophy*, Cambridge MA: Harvard University Press

Ritter, Bernhard, 2021, 'Solace or Council for Death: Kant and Maria von Herbert,' in *Women and Philosophy in Eighteenth-Century Germany*, edited by Corey W. Dyck, Oxford: Oxford University Press, pp. 137-156

Ritter, Christian, 1971, *Der Rechtsgedanke Kants nach den frühen Quellen*, Frankfurt am Main: Klostermann

Rohden, Valerio, 2005, 'Somente uma Nova Fórmula: A propósito das críticas eudemonistas de Tittel à Filosofia Moral de Kant,' *Revista Portuguesa de Filosofia* vol. 61, no. 3/4, pp. 747-755

Rorty, Richard, 2004, 'Trapped between Kant and Dewey: The Current Situation of Moral Philosophy,' in *New Essays on the History of Autonomy: A Collection Honoring J. B. Schneewind*, edited by Natalie Brender and Larry Krasnoff, Cambridge: Cambridge University Press, pp. 195-214

Rousseau, Jean Jacques, 1961, *Émile*, translated by Barbara Foxley, London: Dent & Sons

Rühl, Ulli F. H., 2015, 'Metaphysische Anfangsgründe der Tugendlehre (Die Metaphysik der Sitten, zweiter Teil),' in *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr and Stefano Bacin, Berlin: de Gruyter, pp. 1562-1568

Sala, Giovanni B., 2004, *Kant's "Kritik der praktischen Vernunft": ein Kommentar*, Darmstadt: Wissenschaftliche Buchgesellschaft

Salam, Abdallah, 2014, 'Perfect and Imperfect Rights, Duties and Obligations: From Hugo Grotius to Immanuel Kant,' DPhil diss., University of Oxford

Sandford, Stella, 2018, 'Kant, race, and natural history,' *Philosophy and Social Criticism* vol. 44, no. 9, pp. 950-977

Scarre, Geoffrey, 1998, 'Interpreting the Categorical Imperative,' *British Journal for the History of Philosophy* vol. 6, no. 2, pp. 223-236

Schadow, Steffi, 2013, 'Recht und Ethik in Kants Metaphysik der Sitten (MS 6:218–221 und TL 6:390 f.),' in *Kant's Tugendlehre: A Comprehensive Commentary*, edited by Andreas Trampota, Oliver Sensen and Jens Timmermann, Berlin: de Gruyter, pp. 85-111

Schmidt, James 1989, 'The Question of Enlightenment: Kant, Mendelssohn, and the Mittwochsgesellschaft,' *Journal of the History of Ideas*, vol. 50, no. 2, pp. 269-291

Schmidt, James, 1992, 'What Enlightenment Was: How Moses Mendelssohn and Immanuel Kant Answered the Berlinische Monatsschrift,' *Journal of the History of Philosophy*, vol. 30, no. 1, pp. 77-101

Schmidt, James, 2003, 'Inventing the Enlightenment: Anti-Jacobins, British Hegelians, and the "Oxford English Dictionary",' *Journal of the History of Ideas* vol. 64, no. 3, pp. 421-443

Schmidt, James, 2006, 'What Enlightenment Was, What It Still Might Be, and Why Kant May Have Been Right After All,' *American Behavioral Scientist* vol. 49, no. 5, pp. 647-663

Schneewind, Jerome B., 1991, 'Natural Law, Skepticism, and Methods of Ethics,' *Journal of the History of Ideas* vol. 52, no. 2, pp. 289-308

Schneewind, Jerome B., 1998, *The Invention of Autonomy: A History of Modern Moral Philosophy*, Cambridge: Cambridge University Press

Schneewind, Jerome B., 2009, 'Kant against the 'spurious principles of morality',' in *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*, edited by Jens Timmermann, Cambridge: Cambridge University Press

Schneewind, Jerome B., 2010, *Essays on the History of Moral Philosophy*, Oxford: Oxford University Press

Schnoor, Christian, 1989, *Kants Kategorischer Imperativ als Kriterium der Richtigkeit des Handelns*, Tübingen: Mohr (Paul Siebek)

Schofield, Malcolm, 1991, *The Stoic Idea of the City*, Chicago and London: University of Chicago Press

Schopenhauer, Arthur, 2010, *The World as Will and Representation, Volume I*, translated and edited by Judith Norman, Alistair Welchman, and Christopher Janaway, Cambridge: Cambridge University Press

Schüssler, Rudolf, 2012, 'Kant und die Kasuistik: Fragen zur Tugendlehre,' *Kant-Studien*, vol. 103, no. 1, pp. 70–95

Schüssler, Rudolf, 2019, *The Debate on Probable Opinions in the Scholastic Tradition*, Leiden: Brill

Schüssler, Rudolf, 2020, ,Kant und der Probabilismus,' *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 193-216

Schwartz, Jeremy, 2017, 'Was Kant a 'Kantian Constructivist'?' *Kantian Review* vol. 22, no. 2, pp. 257-280

Sedgwick, Sally, 1988, 'On the Relation of Pure Reason to Content: A Reply to Hegel's Critique of Formalism in Kant's Ethics,' *Philosophy and Phenomenological Research* vol. 49, no. 1, pp. 59-80

Sedgwick, Sally, 1990, 'Can Kant's Ethics Survive the Feminist Critique?' *Pacific Philosophical Quarterly* vol. 71, no. 1, pp. 60-79

Sedgwick, Sally, 1996, 'Hegel's Critique of Kant's Empiricism and the Categorical Imperative,' *Zeitschrift für philosophische Forschung* vol. 50, no. 4, pp. 563-584

Sedgwick, Sally, 2011, 'Hegel on the Empty Formalism of Kant's Categorical Imperative,' in *A Companion to Hegel*, edited by Stephen Houlgate and Michael Baur, Wiley-Blackwell, pp. 265-280

Sedgwick, Sally, 2012, *Hegel's Critique of Kant*, Oxford: Oxford University Press

Seel, Gerhard, 2009, 'How does Kant justify the universal objective validity of the law of right?' *International Journal of Philosophical Studies* vol. 17, no. 1, pp. 71-94

Sensen, Oliver, 2013, 'Kant's Constructivism,' in *Constructivism in Ethics*, edited by Carla Bagnoli, Cambridge: Cambridge University Press, pp. 63–82

Serequeberhan, Tsenay, 1996, 'Eurocentrism in philosophy: The case of Immanuel Kant,' *The* Shell, Susan Meld, 1996, *The Embodiment of Reason. Kant on Spirit, Generation and Community*, Chicago: Chicago University Press

Shell, Susan Meld, 1996, *The Embodiment of Reason. Kant on Spirit, Generation and Community*, Chicago: Chicago University Press

Shell, Susan Meld, 2006, 'Kant's Conception of a Human Race,' in *The German Invention of Race*, edited by Sara Eigen and Mark Larrimore, Albany: SUNY Press, pp. 55-72

Shell, Susan Meld, 2009, *Kant and the Limits of Autonomy*, Cambridge, MA: Harvard University Press

Shell, Susan, and Richard Velkley, 2017, "Rousseau and Kant: Rousseau's Kantian Legacy," in *Thinking with Rousseau: From Machiavelli to Schmitt*, edited by Helena Rosenblatt and Paul Schweigert, Cambridge: Cambridge University Press, pp. 192– 210

Shyllon, Folarin, 1977, *James Ramsay: the unknown abolitionist*, Edinburgh: Canongate

Sidgwick, Henry, 1907, *The Methods of Ethics*, 7th edition, Indianapolis/Cambridge: Hackett

Siep, Ludwig, 2009, 'What Is the Purpose of a Metaphysics of Morals? Some Observations on the Preface to the *Groundwork of the Metaphysics of Morals*,' in *Kant's Moral and Legal Philosophy*, edited by Karl Ameriks, Otfried Höffe, translated by Nicholas Walker, Cambridge: Cambridge University Press, pp. 77-92

Singer, Marcus George, 1961, *Generalizations in Ethics: An Essay in the Logic of Ethics, with the Rudiments of a System of Moral Philosophy*, New York: Alfred Knopf

Singer, Peter, 1972, 'Moral Experts,' *Analysis* vol. 32, no. 4, pp. 115-117

Singer, Peter, 1974, 'Philosophers Are Back on the Job,' *New York Times Magazine*, July 7, pp. 6-7, 17-20

Sinnott-Armstrong, Walter, 2019, 'Moral Skepticism,' *Stanford Encyclopedia of Philosophy*, Summer 2019 Edition, edited by Edward N. Zalta

Smith, Adam, 2002, *The Theory of Moral Sentiments*, edited by Knud Haakonssen, Cambridge: Cambridge University Press

Smith, Michael, 1994, *The Moral Problem*, Oxford: Blackwell

Sosa, Ernest, 1991, *Knowledge in perspective: Selected essays in epistemology*, Cambridge: Cambridge University Press

Spellman, W. M., 1988, *John Locke and the Problem of Depravity*, Oxford: Clarendon Press

Spinoza, Benedict De, 2007, *Theological-Political Treatise*, edited by Jonathan Israel, translated by Michael Silverthorne and Jonathan Israel, Cambridge: Cambridge University Press

Sprengel, Matthias Christian, 1786, *Beiträge zur Völker- und Länderkunde. Fünfter Theil.* Leipzig: Weygandsche Buchhandlung

Stapleford, Scott, 2007, 'On the Contradiction in Conception Test of the Categorical Imperative,' *South African Journal of Philosophy* vol. 26, no. 3, pp 306-318

Stäudlin, Carl Friedrich, 1808, *Geschichte der christlichen Moral: seit dem Wiederaufleben der Wissenschaften*, Göttingen, J. F. Röwer

Stern, Robert, 2012, 'On Hegel's Critique of Kant's Ethics: Beyond the Empty Formalism Objection,' in *Hegel's Philosophy of Right: Essays on Ethics, Politics, and Law*, edited by Thom Brooks, London: Blackwell, pp. 73-99

Stern, Robert, 2013, 'Moral Skepticism, Constructivism, and the Value of Humanity,' in *Constructivism in Ethics*, edited by Carla Bagnoli, Cambridge: Cambridge University Press, pp. 22–41

Sticker, Martin, 2015, 'The Moral-Psychology of the Common Agent – A Reply to Ido Geiger,' *British Journal for the History of Philosophy*, vol. 23, no. 5, pp. 976-989

Sticker, Martin, 2016, 'Kant's Criticism of Common Moral Rational Cognition,' *European Journal of Philosophy* vol. 25, no. 1, pp. 85-108

Sticker, Martin, 2020, 'Kant und das fehlbare Gewissen,' in *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant,* edited by Sara Di Giulio and Albert Frigo, Berlin and Boston: de Gruyter, pp. 289-310

Sticker, Martin, 2021, *Rationalizing (Vernünfteln)*, Cambridge University Press

Suarez, Francisco, 2003, 'On Law and God the Lawgiver,' in *Moral Philosophy from Montagne to Kant*, anthology edited by J. B. Schneewind, Cambridge: Cambridge University Press, pp. 86-85

Sullivan, Roger J., 1989, *Immanuel Kant's moral theory*, Cambridge: Cambridge University Press

Sutter, Alex, 1989, 'Kant und die Wilden. Zum impliziten Rassismus in der Kantischen Geschichtsphilosophie,' *Prima Philosophia* vol. 2, no. 2, pp. 241-265

Terra, Ricardo, 2013, 'Hat die kantische Vernunft eine Hautfarbe?' in *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses. Band 1*, edited by Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca and Margit Ruffing, Berlin and Boston: de Gruyter, pp. 431-447

Theunissen, L. Nandi, 2013, 'Kant's Commitment to Metaphysics of Morals,' *European Journal of Philosophy* vol. 24, no. 1, pp. 103-128

Thorpe, Lucas, 2006, 'The Point of Studying Ethics According to Kant,' *Journal of Value Inquiry* vol. 40, no. 4, pp. 461-474

Timmermann, Jens, 2000, 'Kant's Puzzling Ethics of Maxims,' *Harvard Review of Philosophy*, vol. 8, pp. 39-52

Timmermann, Jens, 2007a, *Kant's Groundwork of the Metaphysics of Morals: A Commentary*, Cambridge: Cambridge University Press

Timmermann, Jens, 2007b, 'Simplicity and Authority: Reflections on Theory and Practice in Kant's Moral Philosophy' *Journal of Moral Philosophy*, vol. 4, no. 2, pp. 167-182

Timmermann, Jens, 2015, 'Metaphysik der Sitten,' in *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr and Stefano Bacin, Berlin: de Gruyter, pp. 1542-1543

Timmons, Mark, 2017, *Significance and System: Essays on Kant's Ethics*, Oxford: Oxford University Press

Tittel, Gottlob August, 1786, *Über Herrn Kants Moralreform*, Frankfurt und Leipzig: Gebrüder Pfähler

Tobin, James, 1785, *Cursory Remarks upon the Reverend Mr Ramsay's Essay on the Treatment and Conversion of African Slaves in the Sugar Colonies*, London: G. and T. Wilkie

Tutino, Stefania, 2018, *Uncertainty in Post-Reformation Catholicism. A History of Probabilism*, Oxford: Oxford University Press

Unna, Yvonne, 2003, 'Kant's Answers to the Casuistical Questions Concerning Self-Disembodiment,' *Kant-Studien*, vol. 94, no. 4, pp. 454–473

Varden, Helga, 2017, 'Kant and Women,' *Pacific Philosophical Quarterly* vol. 98, no. 4, pp. 653-694

Velkley, Richard L., 1989, *Freedom and the End of Reason: On the Moral Foundation of Kant's Critical Philosophy*, Chicago: University of Chicago Press

Velleman, J. David, 2012, 'Reading Kant's Groundwork,' in *Ethics: Essential Readings in Moral Theory*, edited by George Sher, Oxford: Routledge, pp. 343-359

von Freier, Friedrich, 1992, 'Kritik der Hegelschen Formalismusthese,' *Kant-Studien* vol. 83, no. 3, pp. 304-323

von Vattel, Emerich, 1760, *Völkerrecht; oder: gründliche Anweisung wie die Grundsäze des natürlichen Rechts auf das Betragen und auf die Angelegenheiten der Nationen und Souveräne angewendet werden müßen*, third part, translated from the French by Johann Philipp Schulin, Frankfurt and Leipzig

Vorländer, Karl, 1893, *Der Formalismus der Kantischen Ethik in seiner Notwendigkeit und Fruchtbarkeit*, Marburg: Universitäts-Buchdruckerei

Vorländer, Karl, 1959, 'Einleitung des Herausgebers,' in Kant, Immanuel, 1959, *Kritik der praktischen Vernunft*, edited by Karl Vorländer, Hamburg: Meiner, pp. xi-xlvii

Walschots, Michael, 2020, 'Merely a New Formula? G.A. Tittel on Kant's 'Reform' of Moral Science' *Studi Kantiani* vol. 33, special issue on "The Early Critique of Kant's Moral Philosophy," pp. 49-64

Walsh, W. H., 1969, *Hegelian Ethics*, London: Macmillan

Warda, Arthur, 1922, *Immanuel Kants Bücher*, Berlin: Breslauer

Watson, Alan, 1996a, 'Seventeenth-Century Jurists, Roman Law, and the Law of Slavery,' in *Slavery & the Law*, edited by Paul Finkelman, Madison, Wis.: Madison House, pp. 367-377

Watson, Alan, 1996b, 'Thinking Property at Rome,' in *Slavery & the Law*, edited by Paul Finkelman, Madison, Wis.: Madison House, pp. 419-435

Wehofsits, Anna, 2020, 'Passions: Kant's psychology of self-deception,' *Inquiry* vol. 66, no. 6, pp. 1184-1208

Westphal, Kenneth R., 1991, 'Hegel's Critique of Kant's Moral World View,' *Philosophical Topics* vol. 19, no. 2, pp. 133-176

Wildt, Andreas, 1982, *Autonomie und Anerkennung: Hegel's Moralitätskritik im Lichte seiner Fichte-Rezeption*, Stuttgart: Klett-Cotta

Willaschek, Marcus, 2009, 'Right and coercion: Can Kant's conception of right be derived from his moral theory?' *International Journal of Philosophical Studies* vol. 17, no. 1, pp. 49-70

Willaschek, Marcus, 2012, 'The Non-Derivability of Kantian Right from the Categorical Imperative: A Response to Nance,' *International Journal of Philosophical Studies* vol. 20, no. 4, pp. 557-564

Williams, Bernard, 1995, *Making Sense of Humanity, and other philosophical papers 1982-1993*, Cambridge: Cambridge University Press

Williams, Bernard, 2006, *Ethics and the Limits of Philosophy*, first published 1985, London and New York: Routledge

Williams, Howard, 2004, 'Colonialism in Kant's Political Philosophy,' *Diametros* vol. 39, pp. 154-181

Wolf, W. Clark, 2023, 'Kant's Formula of Universal Law as a Test of Causality,' *Philosophical Review* vol. 132, no. 3, pp. 459-490

Wolff, Christian, 1736, *Vernünftige Gedanken von der Menschen Thun Lassen, zu Beförderung ihrer Glückseeligkeit, den Liebhabern der Wahrheit mitgetheilet*, Frankfurt und Leipzig

Wolff, Christian, 2003, 'Reasonable Thoughts About the Actions of Men, for the Promotion of Their Happiness,' in *Moral Philosophy from Montagne to Kant*, anthology edited by J. B. Schneewind, Cambridge: Cambridge University Press, pp. 333-347

Wood, Allen W., 1970, *Kant's Moral Religion*, Ithaca and London: Cornell University Press

Wood, Allen W., 1990, *Hegel's Ethical Thought*, Cambridge: Cambridge University Press

Wood, Allen W., 1997, 'The Final Form of Kant's Practical Philosophy' *Southern Journal of Philosophy* vol. 36, pp. 1-20. Reprinted in *Kant's Metaphysics of Morals: Interpretive Essays*, edited by Mark Timmons, 2002, Oxford: Oxford University Press, pp. 1-21

Wood, Allen W., 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press

Wood, Allen W., 2007, *Kantian Ethics*, Cambridge: Cambridge University Press

Wood, Allen W., 2017, *Formulas of the Moral Law*, Cambridge: Cambridge University Press

Wuerth, Julian, 2011, 'Moving beyond Kant's account of agency in the *Grounding*,' in *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics*, edited by Lawrence Jost and Julian Wuerth, Cambridge: Cambridge University Press, pp. 147-163

Zammito, John H., 2002, *Kant, Herder, and the Birth of Anthropology*, Chicago: University of Chicago Press

Zammito, John H., 2006, 'Policing Polygeneticism in Germany, 1775 (Kames,) Kant, and Blumenbach,' in *The German Invention of Race*, edited by Sara Eigen and Mark Larrimore, Albany: SUNY Press, pp. 35-54

Zöller, Günter, 2013, 'Idee und Notwendigkeit einer Metaphysik der Sitten (MS 6:205–209, 214–218 und TL 6:375–378),' in in *Kant's Tugendlehre: A Comprehensive Commentary*, edited by Andreas Trampota, Oliver Sensen and Jens Timmermann, Berlin: de Gruyter, pp. 11-24

Zweig, Arnulf, 1999, 'Herbert, Maria von,' biographical sketch, in *Immanuel Kant: Correspondence*, translated and edited by Arnulf Zweig, Cambridge: Cambridge University Press, pp. 578-579