



KANT AND THE PROBLEM OF MORALITY

RETHINKING THE CONTEMPORARY WORLD

Edited by
Luigi Caranti and Alessandro Pinzani

KANT AND THE PROBLEM OF MORALITY

This book examines the significance of Kant's moral philosophy in contemporary philosophical debates. It argues that Kant's philosophy can still serve as a guide to navigate the turbulence of a globalized world in which we are faced by an imprescriptible social reality wherein moral values and ethical life models are becoming increasingly unstable. The volume draws on Kantian ethics to discuss various contemporary issues, including sustainable development, moral enhancement, sexism and racism. It also tackles general concepts of practical philosophy such as lying, the different kinds of moral duties and the kind of motivation one needs for doing what we consider the right thing.

Featuring readings by well-known Kant specialists and emerging scholars with unorthodox approaches to Kant's philosophy, the volume will be of great interest to scholars and researchers of philosophy, politics and ethics. It will also appeal to moral theorists, applied ethicists and environmental theorists.

Luigi Caranti is Professor of Political Philosophy at the Università di Catania. He focuses on Kant, human rights, peace studies and distributive justice. Principal investigator of numerous EU-funded research projects, he is currently coordinating the Marie Curie Rise project "Kant in South America". Among his recent publications are *The Kantian Federation* (2022), (ed. with D. Celentano) *Paradigms of Justice: Redistribution, Recognition and Beyond* (2021) and *Kant's Political Legacy: Human Rights, Peace, Progress* (2017).

Alessandro Pinzani is Professor of Ethics and Political Philosophy at the Federal University of Santa Catarina, Florianópolis (Brazil) and, since 2006, a fellow researcher of CNPq (Brazilian Research Council). His publications include *Jürgen Habermas* (2007), *An den Wurzeln moderner Demokratie* (2009) and *Money, Autonomy, and Citizenship* (with W. Leão Rego, 2018).



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

KANT AND THE PROBLEM OF MORALITY

Rethinking the Contemporary World

*Edited by Luigi Caranti and
Alessandro Pinzani*

First published 2023
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2023 selection and editorial matter, Luigi Caranti and Alessandro Pinzani; individual chapters, the contributors

The right of Luigi Caranti and Alessandro Pinzani to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record has been requested for this book

ISBN: 978-0-367-46125-6 (hbk)

ISBN: 978-1-032-28950-2 (pbk)

ISBN: 978-1-003-04312-6 (ebk)

DOI: 10.4324/9781003043126

Typeset in Sabon
by Apex CoVantage, LLC

CONTENTS

<i>List of contributors</i>	vii
<i>Acknowledgments</i>	ix
<i>List of sigla</i>	xi
An introduction	1
LUIGI CARANTI AND ALESSANDRO PINZANI	
1 On dealing with Kant's sexism and racism	6
PAULINE KLEINGELD	
2 Kant the naturalist	22
ROBERT B. LOUDEN	
3 Pleasure and displeasure as moral motivation	37
MARIA BORGES	
4 Inefficacy, despair, and difference-making: a secular application of Kant's moral argument	47
ANDREW CHIGNELL	
5 Lying, deception and dishonesty: Kant and the contemporary debate on the definition of lying	73
STEFANO BACIN	
6 The duty and the maxims: elements for a morality and culture of sustainable development	92
ANSELMO APORTONE	

CONTENTS

7	Kant as an <i>ante litteram</i> theorist and critic of the moral enhancement	126
	ALBERTO PIRNI	
8	<i>Spielraum</i> : narrow and wide duties and their consequences	138
	CLAUDIO LA ROCCA	
	<i>Index</i>	158

CONTRIBUTORS

Anselmo Aportone is Professor of Theoretical Philosophy at the University of Rome “Tor Vergata” and a board member of the Italian Society of Kantian Studies. The main line of his research is focused on Kant’s philosophy, to which he has devoted several publications, including the volumes *Erfahrung, Kategorien und Möglichkeit* (Roma 2012), *Gestalten der transzendentalen Einheit* (Berlin – New York 2009) and *Kant et le pouvoir réceptif. Recherches sur la conception kantienne de la sensibilité* (Paris 2014).

Stefano Bacin is Associate Professor in History of Philosophy at the University of Milan. He is the author of *Il senso dell’etica. Kant e la costruzione di una teoria morale* (2006), *Fichte in Schulpforta (1774–1780). Kontext und Dokumente* (2007) and *Kant e l’autonomia della volontà* (2021). He is also co-editor of a three-volume *Kant-Lexikon* (with Marcus Willaschek, Georg Mohr and Jürgen Stolzenberg; De Gruyter, 2015), *The Emergence of Autonomy in Kant’s Moral Philosophy* (with Oliver Sensen, 2019) and *Fichte’s “System of Ethics”: A Critical Guide* (with Owen Ware, 2021).

Maria Borges is Full Professor of Philosophy at the University of Santa Catarina (Brazil). She was Visiting Scholar at the University of Pennsylvania (USA/1999), Humboldt Universität (Germany/2006–7) and Columbia University (USA/2014). She is researcher of the CNPq/Brazil and President of the Brazilian Kant Society. She is the author of *Body and Justice* (2011), *Emotion, Reason, and Action in Kant* (2019) and many articles, including “What Can Kant Teach Us About Emotions” (*The Journal of Philosophy*) and “Physiology and the Controlling of Affects in Kant’s Philosophy” (*Kantian Review*). Her philosophical interests are Kant, ethics, theory of emotions and feminist philosophy.

Andrew Chignell is Laurence S. Rockefeller Professor in the University Center for Human Values at Princeton University, with appointments in the departments of philosophy and religion. His research focuses on

Kant and early modern philosophy, philosophy of religion, the ethics of belief and moral psychology. He is the director of the Princeton Project in Philosophy and Religion and currently serves as the President of the North American Kant Society.

Pauline Kleingeld is Professor of Philosophy at the University of Groningen, the Netherlands. Her academic interests are in ethics and political philosophy, with an emphasis on Kantianism and the work of Immanuel Kant. She is the author of *Kant and Cosmopolitanism: The Philosophical Ideal of World Citizenship* (2012) and *Fortschritt und Vernunft: Zur Geschichtsphilosophie Kants* (1995). She has also edited three volumes with translations of texts by Immanuel Kant and Gottfried Achenwall.

Claudio La Rocca is Professor of Theoretical Philosophy and Philosophy of Mind at the University of Genoa. His research interests focus on Italian philosophy of the 20th century, phenomenology, hermeneutics, philosophy of mind and in particular on the philosophy of Kant, on which he has published several volumes and numerous essays in Italian, English, French, Spanish and German. From 2006 to May 2016, he was President of the Società Italiana di Studi Kantiani. He is Editor-in-Chief of the journal *Studi kantiani*.

Robert B. Loudon is Distinguished Professor of Philosophy at the University of Southern Maine, USA. His publications include *Johann Bernhard Basedow and the Transformation of Modern Education* (2021), *Anthropology from a Kantian Point of View* (2021), *Kant's Human Being* (2011), *The World We Want* (2007), *Kant's Impure Ethics* (2000) and *Morality and Moral Theory* (1992). A former president of the North American Kant Society (NAKS), Loudon is also co-editor and translator of two volumes in *The Cambridge Edition of the Works of Immanuel Kant*.

Alberto Pirni is Associate Professor in Moral Philosophy at the Sant'Anna School of Advanced Studies (Pisa), where he coordinates the research area in public ethics. He is also Responsible Director of the international journal *Lexicon of Public Ethics*. His publications include *Il regno dei fini in Kant* (il melangolo, 2000); *Kant filosofo della comunità* (ETS 2006); (ed. by) *Kant y las declinaciones de la armonía*, "Con-Textos Kantianos", 2, 2015; "The Place of Sociality. Models of Intersubjectivity According to Kant", in A. Faggion, A. Pinzani, N. Sanchez Madrid (eds.); *Kant and Social Philosophy* (2016); and "Kant e la filosofia cosmopolitica", in B. Henry, A. Loretoni, A. Pirni, M. Solinas (eds.), *Filosofia politica* (2020).

ACKNOWLEDGMENTS

Does Kant still have something to say to us after more than 200 years from the publication of his major works (the three *Critiques*, the *Groundwork*, the *Metaphysics of Morals*, the political writings and those on the philosophy of history)? Can his thinking still serve as a guide to navigate the turbulence of a globalized world, of an imprescriptible social reality in which moral values and ethical life models have lost their stability, while sciences – even the most exact ones – are painfully aware of the precarity of their foundations, and the borders between domestic and international issues are increasingly blurred?

The editors of these volumes are convinced that these questions are to be answered positively and that Kant remains a source of inspiration for dealing with the latest developments in areas that include ethics, politics and the theory of knowledge. Guided by this belief, we have invited Kant specialists from different backgrounds to discuss contemporary epistemological, moral and political issues from a Kantian perspective. Some have chosen to stay close to Kant's texts, and others have used his work merely as a source of inspiration; all of them, we believe, have shown how his critical philosophy retains a capacity for interpreting reality in salient ways and for offering solutions to our problems – be they new ones, which Kant himself could not foresee, or old ones, which have haunted humankind from its beginning.

Editing these volumes was a journey that began organically some years ago, first through personal conversations between us, then through common projects. The decisive pivot occurred during the 8th Multilateral Kant Congress, held in Catania October 11–13, 2018, which was in turn made possible by a grant from the European Commission that enabled a four-year-long structured cooperation between Kant scholars from Europe and South America (Marie Curie RISE n.777786 “Kant in South America”). The present book constitutes the main scientific output of that project. Its articulation in three volumes, devoted separately to knowledge, ethics and politics, reflects the three scientific work packages into

ACKNOWLEDGMENTS

which that project was divided. Many of the authors who participated in this publication presented a first draft of their contributions in Catania, while others joined the project later. Our gratitude goes to all of them. Special thanks to Routledge and to our editor there, Aakash Chakrabarty, for believing in the validity of our idea and for being willing to publish the book despite its dimensions.

SIGLA

References to Kant's works follow the sigla reproduced in the following. The English translation is that of the Cambridge edition unless otherwise indicated by the authors.

Kant, Immanuel: Gesammelte Schriften Hrsg.: Bd. 1–22 Preussische Akademie der Wissenschaften, Bd. 23 Deutsche Akademie der Wissenschaften zu Berlin, ab Bd. 24 Akademie der Wissenschaften zu Göttingen. Berlin 1900ff.

AA	Akademie-Ausgabe
Anth	Anthropologie in pragmatischer Hinsicht (AA 07)
AP	Aufsätze, das Philanthropin betreffend (AA 02)
BDG	Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes (AA 02)
Br	Briefe (AA 10–13)
DfS	Die falsche Spitzfindigkeit der vier syllogistischen Figuren erwiesen (AA 02)
DI	Meditationum quarundam de igne succincta delineatio (AA 02)
EaD	Das Ende aller Dinge (AA 08)
EACG	Entwurf und Ankündigung eines Collegii der physischen Geographie (AA 02)
EEKU	Erste Einleitung in die Kritik der Urteilskraft (AA 20)
Ethica (1763)	Ethica Philosophia (AA 27)
FBZE	Fortgesetzte Betrachtung der seit einiger Zeit wahrgenommenen Erderschütterungen (AA 01)
FEV	Die Frage, ob die Erde veralte, physikalisch erwogen (AA 01)
FM	Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnizens und Wolff's Zeiten in Deutschland gemacht hat? (AA 20)
FM/Beylagen	FM: Beylagen (AA 20)
FM/Lose Blätter	FM: Lose Blätter (AA 20)

FRT	Fragment einer späteren Rationaltheologie (AA 28)
GAJFF	Gedanken bei dem frühzeitigen Ableben des Herrn Johann Friedrich von Funk (AA 02)
GMS	Grundlegung zur Metaphysik der Sitten (AA 04)
GNVE	Geschichte und Naturbeschreibung der merkwürdigsten Vorfälle des Erdbebens, welches an dem Ende des 1755sten Jahres einen großen Theil der Erde erschütteret hat (AA 01)
GSE	Beobachtungen über das Gefühl des Schönen und Erhabenen (AA 02)
GSK	Gedanken von der wahren Schätzung der lebendigen Kräfte (AA 01)
GUGR	Von dem ersten Grunde des Unterschiedes der Gegenden im Raume (AA 02)
HN	Handschriftlicher Nachlass (AA 14–23)
IaG	Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht (AA 08)
KpV	Kritik der praktischen Vernunft (AA 05)
KrV	Kritik der reinen Vernunft
KU	Kritik der Urteilskraft (AA 05)
Log	Logik (AA 09)
MAM	Mutmaßlicher Anfang der Menschheitsgeschichte (AA 08)
MAN	Metaphysische Anfangsgründe der Naturwissenschaft (AA 04)
MoPh	Metaphysicae cum geometria iunctae usus in philosophia naturali, cuius specimen I. continet monadologiam physicam (AA 01)
MpVT	Über das Mißlingen aller philosophischen Versuche in der Theodicee (AA 08)
MS	Die Metaphysik der Sitten (AA 06)
RL	Metaphysische Anfangsgründe der Rechtslehre (AA 06)
TL	Metaphysische Anfangsgründe der Tugendlehre (AA 06)
MSI	De mundi sensibilis atque intelligibilis forma et principiis (AA 02)
NEV	Nachricht von der Einrichtung seiner Vorlesungen in dem Winterhalbjahre von 1765–1766 (AA 02)
NG	Versuch, den Begriff der negativen Größen in die Weltweisheit einzuführen (AA 02)
NLBR	Neuer Lehrbegriff der Bewegung und Ruhe und der damit verknüpften Folgerungen in den ersten Gründen der Naturwissenschaft (AA 02)

NTH	Allgemeine Naturgeschichte und Theorie des Himmels (AA 01)
OP	Opus Postumum (AA 21 und 22)
Päd	Pädagogik (AA 09)
-PG	Physische Geographie (AA 09)
PhilEnz	Philosophische Enzyklopädie (AA 29)
PND	Principiorum primorum cognitionis metaphysicae nova dilucidatio (AA 01)
Prol	Prolegomena zu einer jeden künftigen Metaphysik (AA 04)
Refl	Reflexion (AA 14–19)
RezHerder	Recensionen von J. G. Herders Ideen zur Philosophie der Geschichte der Menschheit (AA 08)
RezHufeland	Recension von Gottlieb Hufeland's Versuch über den Grundsatz des Naturrechts (AA 08)
RezMoscatti	Recension von Moscatis Schrift: Von dem körperlichen wesentlichen Unterschiede zwischen der Structur der Thiere und Menschen (AA 02)
RezSchulz	Recension von Schulz's Versuch einer Anleitung zur Sittenlehre für alle Menschen (AA 08)
RezUlrich	Kraus' Recension von Ulrich's Eleutheriologie (AA 08)
RGV	Die Religion innerhalb der Grenzen der bloßen Vernunft (AA 06)
SF	Der Streit der Fakultäten (AA 07)
TG	Träume eines Geistersehers, erläutert durch die Träume der Metaphysik (AA 02)
TP	Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis (AA 08)
TW	Neue Anmerkungen zur Erläuterung der Theorie der Winde (AA 01)
UD	Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral (AA 02)
ÜE	Über eine Entdeckung, nach der alle neue Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll (AA 08)
ÜGTP	Über den Gebrauch teleologischer Principien in der Philosophie (AA 08)
UFE	Untersuchung der Frage, ob die Erde in ihrer Umdrehung um die Achse, wodurch sie die Abwechselung des Tages und der Nacht hervorbringt, einige Veränderung seit den ersten Zeiten ihres Ursprungs erlitten habe (AA 01)
VAEaD	Vorarbeit zu Das Ende aller Dinge (AA 23)
VAKpV	Vorarbeit zur Kritik der praktischen Vernunft (AA 23)

VAMS	Vorarbeit zur Metaphysik der Sitten (AA 23)
VAProl	Vorarbeit zu den Prolegomena zu einer jeden künftigen Metaphysik (AA 23)
VARGV	Vorarbeit zur Religion innerhalb der Grenzen der bloßen Vernunft (AA 23)
VARL	Vorarbeit zur Rechtslehre (AA 23)
VASF	Vorarbeit zum Streit der Fakultäten (AA 23)
VATL	Vorarbeit zur Tugendlehre (AA 23)
VATP	Vorarbeit zu Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis (AA 23)
VAÜGTP	Vorarbeit zu Über den Gebrauch teleologischer Prinzipien in der Philosophie (AA 23)
VAVT	Vorarbeit zu Von einem neuerdings erhobenen vornehmen Ton in der Philosophie (AA 23)
VAZeF	Vorarbeiten zu Zum ewigen Frieden (AA 23)
VBO	Versuch einiger Betrachtungen über den Optimismus (AA 02)
VKK	Versuch über die Krankheiten des Kopfes (AA 02)
VNAEF	Verkündigung des nahen Abschlusses eines Tractats zum ewigen Frieden in der Philosophie (AA 08)
V-Anth/Busolt	Vorlesungen Wintersemester 1788/1789 Busolt (AA 25)
V-Anth/Collins	Vorlesungen Wintersemester 1772/1773 Collins (AA 25)
V-Anth/Fried	Vorlesungen Wintersemester 1775/1776 Friedländer (AA 25)
V-Anth/Mensch	Vorlesungen Wintersemester 1781/1782 Menschenkunde, Petersburg (AA 25)
V-Anth/Mron	Vorlesungen Wintersemester 1784/1785 Mrongovius (AA 25)
V-Anth/Parow	Vorlesungen Wintersemester 1772/1773 Parow (AA 25)
V-Anth/Pillau	Vorlesungen Wintersemester 1777/1778 Pillau (AA 25)
V-Eth/Baumgarten	Baumgarten Ethica Philosophica (AA 27)
V-Lo/Blomberg	Logik Blomberg (AA 24)
V-Lo/Busolt	Logik Busolt (AA 24)
V-Lo/Dohna	Logik Dohna-Wundlacken (AA 24)
V-Lo/Herder	Logik Herder (AA 24)
V-Lo/Philippi	Logik Philippi (AA 24)
V-Lo/Pölitz	Logik Pölitz (AA 24)
V-Lo/Wiener	Wiener Logik (AA 24)
V-Mo/Collins	Moralphilosophie Collins (AA 27)

V-Mo/Kaehler(Stark)	Vorlesungen zur Moralphilosophie (hrsg. von Werner Stark, Berlin – New York, de Gruyter, 2004)
V-Mo/Mron	Moral Mrongovius (AA 27)
V-Mo/Mron II	Moral Mrongovius II (AA 29)
V-Met/Arnoldt	Metaphysik Arnoldt (K 3) (AA 29)
V-Met/Dohna	Metaphysik Dohna (AA 28)
V-Met/Heinze	Metaphysik L1 (Heinze) (AA 28)
V-Met/Herder	Metaphysik Herder (AA 28)
V-Met-K2/Heinze	Metaphysik K2 (Heinze, Schlapp) (AA 28)
V-Met-K3/Arnoldt	Metaphysik K3 (Arnoldt, Schlapp) (AA 28)
V-Met-K3E/Arnoldt	Ergänzungen Kant Metaphysik K3 (Arnoldt) (AA 29)
V-Met-L1/Pölit	Metaphysik L1 (Pölit) (AA 28)
V-Met-L2/Pölit	Metaphysik L2 (Pölit, Original) (AA 28)
V-Met/Mron	Metaphysik Mrongovius (AA 29)
V-Met-N/Herder	Nachträge Metaphysik Herder (AA 28)
V-Met/Schön	Metaphysik von Schön, Ontologie (AA 28)
V-Met/Volckmann	Metaphysik Volckmann (AA 28)
V-MS/Vigil	Die Metaphysik der Sitten Vigilantius (AA 27)
V-NR/Feyerabend	Naturrecht Feyerabend (AA 27)
V-PG	Vorlesungen über Physische Geographie (AA 26)
V-Phil -Th/Pölit	Philosophische Religionslehre nach Pölit (AA 28)
V-PP/Herder	Praktische Philosophie Herder (AA 27)
V-PP/Powalski	Praktische Philosophie Powalski (AA 27)
V-Th/Baumbach	Danziger Rationaltheologie nach Baumbach (AA 28)
V-Th/Pölit	Religionslehre Pölit (AA 28)
V-Th/Volckmann	Natürliche Theologie Volckmann nach Baumbach (AA 28)
VRML	Über ein vermeintes Recht, aus Menschenliebe zu lügen (AA 08)
VT	Von einem neuerdings erhobenen vornehmen Ton in der Philosophie (AA 08)
VUB	Von der Unrechtmäßigkeit des Büchernachdrucks (AA 08)
VUE	Von den Ursachen der Erderschütterungen bei Gelegenheit des Unglücks, welches die westliche Länder von Europa gegen das Ende des vorigen Jahres betroffen hat (AA 01)
VvRM	Von den verschiedenen Racen der Menschen (AA 02)
WA	Beantwortung der Frage: Was ist Aufklärung? (AA 08)
WDO	Was heißt sich im Denken orientiren? (AA 08)
ZeF	Zum ewigen Frieden (AA 08)



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

AN INTRODUCTION

Luigi Caranti and Alessandro Pinzani

Kant's moral philosophy has always been considered one of his major legacies to the history of Western thought. Already since the publication of the *Groundwork* or of the *Critique of Practical Reason*, Kantian ethics was met with enthusiasm by some and provoked strong rebuttal from others. Hegel's criticism, which accused it of abstraction and empty formalism (1991), has been rekindled and repeated with remarkable continuity to the present, despite all the attempts at rebuking it undertaken by Kantians. As always happens with classical thinkers, this debate appears fated to remain forever open. In the last decades, however, the accusation of formalism has lost much of its momentum due to the fact that an increasing number of Kantian interpreters have highlighted the fact that Kant's ethics cannot be reduced to the Categorical Imperative and that, beyond the *Groundwork* and the second *Critique*, in which Kant offers the foundation of his ethics, one should also consider the *Doctrine of Virtue*, in which Kant discusses the concrete application of metaphysical principle of ethics to reality (O'Neill 1989; Höffe 1994). Some interpreters have highlighted the role of the faculty of judgment when it comes to the application of the moral law (Herman 1993; Guyer 1993, 2000); some have pointed out that applying the moral law demands an effort of self-reflection on the part of the subject (Allison 1990, 2020); some have put Kant's ethics within a wider attempt of realizing rationality within the spheres of history and society (Korsgaard 1996); some have claimed that Kant's concepts of autonomy and of human dignity offer the basis for further arguments for justice and human rights (Pogge 1988; Caranti 2017); some have gone so far as to defend the necessity of including the *Anthropology* in the canon of Kant's ethical oeuvre due to the fact that in that work, he discusses relevant ethical questions such as moral motivation and character (Louden 2000; Munzel 1999; Borges 2019; see also the chapter by Louden and Borges).

This latest point is particularly relevant in our time, since in recent years, attempts at founding morality on a naturalistic basis have multiplied (paradigmatically; see Prinz 2007). While Kantian ethics has traditionally been considered a pristine example of deontological ethics as opposed to

consequentialist ethics (utilitarianism) and virtue ethics (neo-Aristotelianism), it is also to be seen as the major alternative to naturalistic ethics, which focus on moral emotions or feelings (a position defended by Hume, among others, and whose weakness Kant had already pointed out) or on some allegedly evolutionary mechanisms that determine our moral behavior or simply deny the existence of moral freedom on the basis of the dubious funding of some neuroscientists. Kant and Kantian ethicists insist that, when acting, we ought to follow the moral law, which tells us what to do because it is the morally right thing to do, independently from what we *feel* to be the right thing to do or from what is the evolutionary best strategy. Furthermore, they insist on the necessity of maintaining the assumption of free will if we do not want to renounce completely not only the notions of individual freedom and of responsibility for actions but the very idea that there is something like a moral good and, therefore, something like morally good actions. Those who deny the existence of free will and attribute our actions to unconscious processes taking place in our brain are holding a deterministic position that Kant had already deemed indemonstrable in his discussion of the third antinomy in the *Critique of Pure Reason*. Those who oppose this naturalistic version of determinism as well as any form of naturalistic ethics find in Kantian ethics a formidable ally (Robert B. Loudon's chapter offers a good example of how determinism can be avoided even when one accepts some form of naturalistic ethics, even one inspired by Kant's ethical theory). Kantian ethics can furthermore offer a basis for discussing present issues such as sustainable development or moral enhancement (see the chapters by Aportone and Pirni) but also sexism and racism (see the chapter by Kleingeld) or more general questions of practical philosophy like lying, the different kinds of moral duties or the kind of motivation we need for doing what we consider the right thing (see respectively the chapters by Bacin, La Rocca and Chignell).

Pauline Kleingeld discusses a topic that has become increasingly relevant in recent years, namely Kant's attitude toward race and gender issues. In doing so, she first provides a brief description of Kant's view on sexual and racial hierarchies and of the way they intersect. She then tackles the question of whether we should "remove and set aside" Kant's undeniable sexism and racism or "translate" his egalitarian principles into inegalitarian ones (Kleingeld advocates a third position). She expresses her perplexity towards the proposal to use inclusive language and female pronouns in discussing Kant's moral and political philosophy. Finally, she advances some preconditions that might allow for fruitfully using Kant's principles to criticize the very sexism and racism he defended.

In his chapter, Robert B. Loudon takes up again his concern with what he dubbed "an impure rather than a pure Kant, an a posteriori rather than an a priori Kant, an empirical rather than a rationalist Kant" and claims that this means also "a *naturalistic* rather than a transcendental Kant". Loudon therefore defends that Kant's ethical theory can also be considered

naturalistic, at least in part. He acknowledges that there are problems connected with this reading, but he makes a compelling case for the hypothesis that there are “two Kants”, the naturalist and the transcendental philosopher – both equally relevant and both capable of helping us solve our present ethical problems.

On a similar line, Maria Borges opposes Kant’s ethics and two forms of moral emotionism: metaphysical emotionism, epitomized by Jesse Prinz, and expressivism, as defended by Simon Blackburn. Borges starts from the claim that Kant would oppose both forms of emotivism. However, she defends that feelings can also play a role in moral action from a Kantian point of view. Borges stresses the relevance of moral feelings in Kant’s oeuvre and offers a fascinating new gaze on the issue.

Andrew Chignell discusses the relevant question of how to find the motivation for doing what we know to be the right thing to do. He refers to Kant’s concept of hope in order to build a possible argument. His starting point is the assessment of the fact that people who like certain products but believe it is wrong to consume them are often demoralized by the apparent inefficacy of their private choices. This is particularly true in contemporary industrial contexts, where many supply chains seem deeply insensitive to slight changes in demand. According to Chignell, “although not a consequentialist, Kant was sensitive to this ‘consequent-dependent’ side of our moral psychology”. Chignell aims to show that Kant’s idea of hope can provide “a model for contemporary, secular arguments regarding attitudes that can be morally justified by their ability to sustain our resolve and avoid demoralization”.

Stefano Bacin’s chapter picks up a moral question that, after having been neglected for quite some time, has again become an important focus of the philosophical debate, namely lying. One of the most famous (and infamous) writings by Kant is dedicated to this topic, and Kant’s position has been widely criticized and considered unsustainable – sometimes also by Kantians. Bacin starts his analysis by considering the very structure of lying, in particular with regard to the question of whether lying is to be understood as a form of deception. Bacin claims that “in Kant’s view, deception is not, and cannot be, the defining feature of lying”. Kant’s position allows for non-deceptive lies, for instance. Therefore, his view should be seen, according to Bacin, “as a specific variant of intrinsic anti-deceptionism”. After having discussed whether the characteristic feature of lying is dishonesty instead of deception, Bacin discusses some “methodological differences between Kant’s account and the contemporary discussion with regard to the role of intuitions and definitions”.

Anselmo Aportone claims that “sustainable development is a true and necessary end of humanity” and that Kant’s ethical theory allows reaching the level of universality and objectivity that moral reflections on this topic should reach in order to guide our practice. More specifically, Kant’s moral philosophy “can support the trust in the faculty of human beings to

overcome selfish motives” as well as the hope that, through the recourse to practical reason, “human beings can be free and pursue happiness in harmony with each other”. Once established that Kant’s ideal is the mutual compatibility of individual freedoms, we could ask the following question: “if sustainable development presupposes ethical actions, which ethics should we develop in view of sustainability?” According to Aportone, the answer could be offered by a Kantian ethics “that combines autonomy, rationality and moral sensitivity to specific contexts”. In order to reach a first verification of this hypothesis, the author analyses the goals formulated in the UN Agenda 2030 and translates them into Kantian imperatives.

Alberto Pirni poses the question of whether Kant could be seen as a precursor of those authors who advocate the admissibility and even necessity of moral enhancement. In order to answer this issue, Pirni follows a twofold path. On the one hand, he claims that pure practical reason itself implies the idea of “enhancement”. On the other hand, he evaluates the possibility that the moral agent can be enhanced “from outside”, namely through three forms of enhancements that are (even if partially) at our disposal nowadays. These are: pharmacological enhancements, digital moral evaluators and neuroprosthetic implants. Finally, the chapter addresses two problematic aspects connected to such forms of enhancement. Specific attention is here devoted to the Kantian idea of autonomy and to the actor’s capacity of being the conscious starting point of a chain of effects in the phenomenal world. Both concepts are undoubtedly challenged by the possibility of artificial improvement of our moral agency.

Claudio La Rocca discusses one of the most central questions of Kant’s ethical theory, namely the distinction between narrow and wide duties and the further distinction between perfect and imperfect duties, which, as the author observes, is considered “the most controversial of Kant’s divisions of moral principles”. The narrow/wide distinction becomes more relevant in the later phases of his thinking, particularly in the *Doctrine of Virtue*. The passages of this work dedicated to this topic, however, pose serious interpretative problems. La Rocca aims to shed some light on the issue, highlighting how Kant’s ethical thought is still an inexhaustible source of philosophical reflection.

References

- Allison, H.E. 1990. *Kant’s Theory of Freedom*. Cambridge: Cambridge University Press.
- Allison, H.E. 2020. *Kant’s Conception of Freedom. A Developmental and Critical Analysis*. Cambridge: Cambridge University Press.
- Borges, M.L. 2019. *Emotion, Reason, and Action in Kant*. London: Bloomsbury.
- Caranti, L. 2017. *Kant’s Political Legacy: Human Rights, Peace, Progress*. Cardiff: University of Wales Press.

AN INTRODUCTION

- Guyer, P. 1993. *Kant and the Experience of Freedom*. Cambridge: Cambridge University Press.
- Guyer, P. 2000. *Kant on Freedom, Law, and Happiness*. Cambridge: Cambridge University Press.
- Hegel, G.W.F. 1991. *Elements of the Philosophy of Right*. Edited by A. Wood. Cambridge: Cambridge University Press.
- Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- Höffe, O. 1994. *Immanuel Kant*. New York: SUNY Press.
- Korsgaard, C.M. 1996. *Creating the Kingdom of Ends*. Cambridge, MA: Harvard University Press.
- Louden, R.B. 2000. *Kant's Impure Ethics. From Rational Beings to Human Beings*. Oxford: Oxford University Press.
- Munzel, G.F. 1999. *Kant's Conception of Moral Character. The "Critical" Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press.
- O'Neill, O. 1989. *Constructions of Reason. Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
- Pogge, T. 1988. "Kant's Theory of Justice." *Kant-Studien* 79: 407–433.
- Prinz, J. 2007. *The Emotional Construction of Morality*. Oxford: Oxford University Press.

1

ON DEALING WITH KANT'S SEXISM AND RACISM

Pauline Kleingeld

Introduction

Immanuel Kant is known as an ardent defender of the moral equality and inviolable dignity of all humans. Yet he also contended that men are naturally superior to women and – for much of his life – that “whites” are naturally superior to other “races”. On these grounds, he defended the rule of men over women and – again for much of his life – the rule of whites over the rest of the world.

Kant is no exception in having held sexist and racist views, and we should not regard his views as a matter of merely contingent personal prejudice. Sexism and racism were endemic features of the Western philosophical discourse of his era and of the belief systems, social practices and political institutions that form the historical context of this discourse.

Kant's case is especially poignant, however. He is one of the greatest philosophers of all time, he was able to break with received opinions on many other issues and he formulated egalitarian moral principles that he claimed to be valid for all human beings – and indeed, more broadly still, for all rational beings. Yet he long defended European colonial rule over the rest of the world and the enslavement, by “whites”, of those he racialized as being “yellow”, “black”, “copper-red” and “mixed”-race. Late in life, around his 70th birthday, Kant dropped the thesis of racial hierarchy and began to criticize European colonialism, but he never made parallel revisions to his account of the status of women.

Many moral theorists have been inspired by Kant's conception of human dignity, equality and the duty of respect. Many also believe that the moral principles Kant articulated can be used precisely to show what is wrong with racism and sexism. But is it possible to do so when we know that Kant himself endorsed racist and sexist views during the very years in which he formulated his egalitarian moral principles? Can we separate the principles from the objectionable views and use Kant's principles to criticize his own biases? These are the questions at issue in this essay.¹

I first provide a brief description of Kant's view on sexual and racial hierarchies and of the way they intersect (Section 1). I then move to the question

whether we should “remove and set aside” Kant’s sexism and racism or “translate” his egalitarian principles into inegalitarian ones, and I advocate a third position (Section 2). In Section 3, I argue that the use of inclusive language and female pronouns, in discussions of Kant’s moral and political philosophy, carries significant risks. I conclude by proposing preconditions for fruitfully using Kant’s principles to criticize sexism and racism.

1. Kant on the sexes and the races

In the 1780s, the decade of the *Groundwork for the Metaphysics of Morals* (1785) and the *Critique of Practical Reason* (1788), Kant defended the view that there is a sexual and racial hierarchy that justifies the subjection of women to men and of non-whites to whites. In the subsequent decade, he gave up his commitment to the racial hierarchy but not to the sexual hierarchy. I will present his views only briefly here, since my interest in this chapter lies in the follow-up questions they raise.²

1.1. Sexual difference and sexual hierarchy

From his early pre-critical writings to his last publications, Kant described women as having very different characteristics than men – characteristics that bear directly on moral agency. In a long chapter on the “contrast” between the sexes, in the early *Observations on the Feeling of the Beautiful and the Sublime* (1764), Kant writes:

The virtue of the woman is a *beautiful virtue*. That of the male sex ought to be a *noble virtue*. Women will avoid evil not because it is unjust but because it is ugly, and for them virtuous actions mean those that are morally beautiful [*sittlich schön*]. Nothing of ought, nothing of must, nothing of obligation. . . . They do something only because they love to, and the art lies in making sure that they love only what is good. I hardly believe that the fair sex is capable of principles, and I hope not to give offense by this, for these are also extremely rare among the male sex.

(GSE 2:231–232)³

Of course, Kant’s gallant ending in this passage does not diminish the gravity of his characterization of women as unreceptive to moral obligation and that of men as having to master the art of directing women toward the good. Nor does Kant’s claim that these sexual differences have been arranged wisely by “Nature” or “Providence” make this sound any better (GSE 2:228–243).

In *Anthropology from a Pragmatic Point of View* (1798), one of Kant’s last publications, he continues to distinguish between “feminine” and

“masculine” virtue, asserting that each has a different “incentive”, that women have “their own vocation” and that this is all part of a grand providential scheme (Anth 7:303–311).

When two people unite, Kant writes with reference to marriage, one must be subordinate to the other. Nature has made men superior to women in strength and courage, whereas women are naturally fearful, and this gives men the right to command. Women, by contrast, are superior to men in being able to conquer the inclination of the other sex toward them. As if this was not already damning with faint praise, Kant adds that men gladly submit to their wives’ regimes so as to be able to go about their own business (Anth 7:303–304).

In his legal and political philosophy, Kant never criticizes the legal tutelage of women; indeed, he justifies it explicitly by reference to male superiority. In the *Metaphysics of Morals*, Kant asserts that the only “human right” is the “innate right to freedom” that “belongs to every human being by virtue of his humanity”. He further explicates this as a right to freedom, equality and independence (including the right to be *one’s own master* [*Herr*], MSRL 6:237). Yet Kant also argues that the “natural superiority” of men gives a husband the right to command over his wife as her master (*Herr*) (MSRL 6:279; cp. Anth 7:209). Further, he classifies “all women” as “passive citizens”, that is, as lacking civil independence and the right to vote. Dependent men (such as domestic servants) are also passive citizens, but Kant explicitly states that they should always have the option of working their way up to active citizenship (MSRL 6:314–315). Nowhere does Kant condemn the subordinate legal and political status of women or call for their emancipation.

Kant shows some awareness of the tensions in his own account. He feels the need to declare that the characteristics of women and their subordinate status do not run counter to the fundamental equality of men and women (MSRL 6:279), but his comments hardly move beyond a reaffirmation of natural male superiority. Moreover, he admits that the very notion of “passive citizenship” “seems to contradict the concept of a citizen as such” (MSRL 6:314). But this does not motivate him to apply his own republican principles to the internal organization of the family or the legal status of women. His claim in the *Anthropology* that when two people unite, one must be subordinated to the other (see above) contradicts his account of the freedom and equality of the citizens who are united in the republic (MSRL 6:314).

In his moral theory, the characteristics he assigns to men, such as courage, appear as the virtues of *human beings*. These are qualities that – he there claims – all human beings ought to strive to realize fully and in a morally appropriate way. The female characteristics do not appear to mark potential *human* excellences, however, and what Kant calls “feminine virtue” is not moral virtue in the strict sense of his ethics.

Kant repeatedly acknowledges that there are women whose conduct does not fit his characterization, such as women scientists. Rather than celebrating

their exceptional accomplishments and calling for their civil and political emancipation, however, he describes them as aberrations (GSE 2: 229–230; V-Anth/Parow, 25: 355; Anth 7:307).⁴ He says that he “would rather not deal with such women” and that, as a rule, “nature has put something into the man for which one will look in vain in a woman” (V-Anth/Parow 25: 355). The women he does praise are “upright women who, in connection with their household, laudably maintained a character suitable to their vocation” (Anth 7:308).⁵ He praises *womanly* women, women who do their *womanly* duties.

1.2. *Racial difference with and without racial hierarchy*

Whereas Kant attributes to women characteristics that *contrast* with those of men, while also asserting their equality, until the mid 1790s, he explicitly describes the “yellow”, “Negro” and “copper-red” races as having increasingly serious *deficits* compared to “whites” and as lacking the capacity to govern themselves. On this basis, Kant defends white colonial rule over the rest of humankind, including the exploitation of non-white slaves. (It is worth noting here that Kant does not restrict the original region of “whites” to Europe but includes Africa north of the Sahara and large parts of Asia; see BBM 8:92.)

Kant portrays “whites” as occupying the highest rung of the racial ladder and as entitled to give laws to all other parts of the world. In his 1782 lectures on physical geography, Kant claims that the peoples of India would be much happier under European rule (V-PG 26: 1058). In drafts of his anthropology lectures, he notes that “Americans and Negroes cannot govern themselves. Thus, [they] serve only as slaves” (Refl 15:878). In the lectures, he is reported as having said that [Native] Americans are the lowest of the four races because they are weak and incapable of being educated. He places “Negroes” above them because they can be trained to be slaves (but are incapable of other forms of education), and he remarks that although the inhabitants of India can be educated, this does not extend to the use of abstract concepts (V-Anth/Mensch 25:1187), and hence they are incapable of being magistrates (Refl 15: 877). Kant also refers to this hierarchy in his published works, such as the 1788 essay “On the Use of Teleological Principles in Philosophy”, which appeared just months after the *Critique of Practical Reason* (ÜGTP 8:176).

Kant’s discussions of chattel slavery until the mid-1790s are strikingly matter of fact.⁶ He reports on the types of slaves needed for various types of labor (VRM 2:438n.), endorses an anti-abolitionist tract (ÜGTP 8:174n.) and remarks that “Negroes” “seem to be made to serve others” (V-Anth/Kowalewski: 363) and “were created for” the harsh labor conditions on the so-called [Caribbean] “Sugar Islands” (V-PG Dohna: 421). The 1780s lecture transcripts include passages such as the following:

The Mandinka are the very most desirable among all Negroes up to the Gambia river, because they are the most hardworking

ones. These are the ones that one prefers to seek for slaves, because these can tolerate labor in the greatest heat that no human being [*Mensch*] can endure. Each year 20,000 of this Negro nation have to be bought to replace their decline in America, where they are used to work on the spice trees. . . . One gets the Negroes by having them catch each other, and one has to seize them with force.

(V-PG 26: 1080)

Note in this passage the implicit contrast between “slave” and “human being” and Kant’s adoption of the perspective of the slave owner when explaining to his students which kinds of slaves “one prefers” and which “have to be bought”.

In the middle of the 1790s, however, not long before the publication of *Toward Perpetual Peace*, Kant *abandoned* the thesis of racial hierarchy and white superiority. In contrast to his earlier characterization of Native Americans as weak, for example, he now calls them courageous, on a par with medieval European knights (ZeF 8:365). Whereas he had previously described conditions on the “Sugar Islands” without any hint of criticism, merely educating his students on the use of these territories for European profit, he shifts to being a vocal critic of colonialism and slavery. In *Toward Perpetual Peace*, Kant writes:

The worst of this (or, considered from the standpoint of a moral judge, the best) is that they [viz., the European states] do not even profit from this violence; that all these trading companies are on the verge of collapse; that the Sugar Islands, this place of the cruelest and most calculated slavery, yield no true profit.

(ZeF 8:359)

Importantly, not only does Kant begin to criticize colonialism and slavery, but he simultaneously adds a new category of public right to his legal and political theory. This is the category of “cosmopolitan right”. Cosmopolitan right grants full and equal juridical status to all humans – to all “citizens of the earth” (*Erdbürger*, MSRL 6:353). It covers relations between states and foreign individuals or groups, including non-state peoples. Among other things, cosmopolitan right prohibits states from imperialist intrusion. No one has a right to settle land used by others, except when expressly permitted through a treaty (ZeF 8:358–359). Kant appeals to this new type of right when he condemns European colonialism and slavery. He calls the “trade in Negroes” a grave violation of their cosmopolitan right (Refl. 23:173–174). He sharply criticizes the fact that the inhabitants of “America, the Negro countries, the Spice Islands, the Cape, etc.”, were treated as ownerless things and “displaced or enslaved” by the Europeans (ZeF 8:358; R 23:173–174). He now strongly condemns the founding of colonies by

annexation, mentioning territories of “American Indians, the Hottentots, and the inhabitants of New Holland” as examples (MSRL 6:266). Instead, he now expresses the hope that “remote parts of the world can establish relations peacefully with one another, relations which ultimately become regulated by public laws and can thus finally bring the human species ever closer to a cosmopolitan constitution” (ZeF 8:358).

The fact that Kant abandons the racial hierarchy of intellectual and agential capacities does not, however, mean that he also drops the notion of race as a physiological concept. He retains this as a biological notion, but he no longer argues that the *physiological* differences between the races are associated with differences in their capacities for *thought and action*. Accordingly, he emphasizes that racial differences are irrelevant to his project in *Anthropology from a Pragmatic Point of View* (1798) since they have no bearing on action (Anth 7:120; cp. Anth 7:320). And, indeed, the *Anthropology* no longer offers – as his earlier lectures on the subject had – an account of the different “characters” of the races (Anth 7:320–321).

1.3. *Where hierarchies intersect*

The two previous sections discussed Kant's sexism and racism separately, and it is important to do so, if only to show the difference between Kant's characterization of women in terms of a *contrast* with men and his characterization of the races in terms of the *deficits* of non-whites (“non-white” being an appropriate expression indeed in this context). With this said, separate discussion of racism and sexism presents only part of the picture. They *intersect*: they exist simultaneously, and this is evident in Kant's description of each.⁷ “*The woman*” described in the *Anthropology* is clearly not a “yellow, black, or red” slave living in a European colony; she has her “own household”, and this is not a corner in the slaves' quarters. “*The family*” described in the *Metaphysics of Morals* is a single household with one married heterosexual couple, their offspring and their male or female dependent servants (MSRL 6:282–283), not the extended family common in many regions outside of Europe.

Conversely, when Kant defends the racial hierarchy, he describes the deficiencies of the “yellow, black, and red” races in terms of their lack of qualities that he attributes to white *men*. As we saw, he claims that other races cannot govern themselves, that they lack courage and that some lack physical strength – these are all weaknesses that he also attributes to all *women*. Similarly, after Kant discards the idea of a racial hierarchy, he begins to attribute “courage” to Native Americans (ZeF 8:365), a characteristic that he then still claims *women* are lacking. In other words, “the” Native Americans he describes are *men*.

Kant does not thematize the different ways in which the intersection of various hierarchies and forms of subordination impact those involved. As a result, important questions remain unaddressed. Let me mention just one

example. Kant writes: “When I make a contract with a servant, he must also be an end . . . and not merely a means. He must also want it” (V-NR/Feyera-bend 27:1319). Kant further argues that a servant ought to obey the head of the household but should have the right to cancel the contract (MSRL 6:283). Although he explicitly includes both “male and female servants” in the household (*ibid.*), he does not discuss whether a *female* servant has a right to enter or cancel the contract on her own or whether this must be done by – and hence with the approval of – her male “*Herr*”, guardian and representative (e.g. her husband, her father). Presumably, Kant’s position is the latter, since he defends women’s civil dependence and writes that the contract is between the (male) head of the household (*Hausherr*) and “free persons” (*ibid.*). But nowhere does he discuss the normative principles that govern the guardian’s conduct in such cases, that is, the conditions under which he *ought* to give approval to a female charge who wishes to terminate her job. Nor does Kant thematize the compounded dependence of female servants in the household – let alone that of female servants of color.

In sum, by examining Kant’s account of the races in light of his sexism, and by examining his account of the sexes in light of his racism, we can expose implicit assumptions in each that might otherwise go unnoticed.

2. Inconsistency, inegalitarianism, and race-neutral and gender-neutral language

With regard to his racism and sexism, there is debate over the question of whether Kant is best seen as an inconsistent egalitarian or a consistent inegalitarian. The motivation behind this debate is not so much to determine whether it is possible to “save” our dear Kant from inconsistency but rather to determine whether it is possible to use Kant’s principles to criticize his biases. After all, if Kant’s prejudice contradicts his principles, it seems much easier to shed the first and retain the second than if they form one coherent set of beliefs.

Some authors argue that Kant was a consistent inegalitarian. Charles Mills, for one, suggests that Kant’s sexism and racism are clear indications that his moral theory is meant to apply only to white men, despite the seemingly inclusive terminology in which it is articulated. Mills argues that Kant saw only white men as “humans” in the *full* sense and that he meant the Categorical Imperative to apply only to white men. All others were, to Kant, inferior beings, *Untermenschen* or sub-persons: biologically human but below the threshold of full personhood on a par with white men (Mills 2005). In this view, it is impossible to use Kant’s principles against his biases since Kant’s principles themselves *carry* the bias. As Mills puts it: “the racist texts are *part* of his theory, not contradictions to it”, and “race in a racist sense is central to his thought” (Mills 2019, 31–32; see also Eze 1994; Bernasconi 2001). Mills argues that we should “translate” Kant’s principles:

In my opinion, there is no “tension” here, and putatively universalist Kantian egalitarian proclamations really need to be translated as restricted in their scope to the white male minority.

(Mills 2019, 34)

Others argue that Kant should be considered an inconsistent egalitarian. They maintain that his hierarchical views on the sexes and races *contradict* his moral and political principles. As a consequence, they argue, we can and should focus on the principles and leave Kant's racist and sexist comments to the side. In this vein, David McCabe has recently argued that Kant's views on race are “not worthy of our serious attention” and that “our eyes should be on Kant's moral theory” (McCabe 2019, 7; cf. Loudon 2000, 105). He writes:

[I]t is not clear why we should be interested in someone's views except where they seem likely to be *philosophically* significant and fruitful, and Kant's views on race are certainly not that.

(McCabe 2019, 7)

Similar views have been defended with regard to Kant's sexism. Mari Mikkola has argued that in those cases where Kant's views on women are inconsistent with his main claims about the use of reason, they should “be bracketed off” or “put to one side” (Mikkola 2011, 105, 107).⁸

McCabe argues that the “logic” of Kant's moral theory is “at odds with other views he had endorsed” and that “Kant's developed moral theory is unambiguous in asserting equal status for all rational beings” (McCabe 2019, 7). In other words, McCabe justifies his claim that Kant's racism contradicts his egalitarianism by appealing to the fact that his moral principles are articulated race neutrally.

In light of Mills's challenge, however, the view that Kant is an inconsistent egalitarian requires more argumentation than a mere reference to the race-neutral and gender-neutral terms in which his principles are formulated. Mills asserts that Kant assumes that only white men are “humans” in the *full* sense and hence that Kant restricts the applicability of the Categorical Imperative to white men. It does not suffice to respond that the Categorical Imperative *must* apply to all women and to all men of color too (“yellow, black, and red” and “mixed”-race) because Kant writes that it is valid for all humans. After all, what is required for showing that Kant grants equal status to women and men of all races is proof that he attributes to women and non-white men those qualities that he claims are required for counting as human *in the full sense*.

Now one might believe that establishing this is easier in Kant's case than in the case of philosophers who wrote in English. Kant wrote in German, and unlike English, where “man” means both “human being” and “male individual”, German has a separate word for each. “Mensch” means “human

being”, and “Mann” means “male human being”. Hence it might seem that whenever Kant makes claims about “Menschen”, we can safely assume that they apply to both sexes and all races distinguished by Kant. Helga Varden has recently defended this view, adding that it seems “somewhat unfair” to “accuse Kant . . . of saying one thing while meaning something else entirely” (Varden 2017, 683–684).

The relevant texts make clear, however, that things are more complicated. As far as race goes, the passage quoted previously in which Kant discusses the Mandinka is a case in point. Here Kant asserts that “no human being” (*Mensch*) can allegedly stand the heat *but that the Mandinka can*. Here, “Mensch” clearly does *not* refer to the Mandinka, although there is no doubt that Kant regarded them as belonging to the human species. Apparently, however, not everything that Kant predicates of “humans” also holds for “Negroes” (just as not everything Kant predicates of the races applies to their female members; see Section 1).⁹ Whether general terms are indeed used in this way can be determined only on the basis of their context.

The same point applies in the case of sex and gender. There are many passages in which Kant moves back and forth between “Mensch” and “Mann”. Consider this claim from the *Critique of Practical Reason*: “There are cases in which human beings [*Menschen*] show from childhood . . . early wickedness and progress in it . . . continuously into their manhood” [*Mannesjahre*] (KpV 5: 99–100). Similarly, the “children” [*Kinder*] of a household eventually become “their own masters” [*Herren*] (MSRL 6:282). Thus, we cannot validly infer from Kant’s use of the general term “Mensch” that he includes all humans in its scope.

This phenomenon is of course by no means peculiar to Kant. In the years after the 1789 declaration of the “rights of man and citizen”, Olympe de Gouges in vain claimed “human” rights for women, and she died under the guillotine. Some particularly interesting examples are found in high-profile legal cases. Nineteenth-century Dutch laws concerning citizenship and voting rights were formulated in gender-neutral language, in terms of “Netherlanders” having to fulfill certain requirements (such as paying a certain amount of taxes). In 1883, Aletta Jacobs, the first Dutch woman to satisfy the requirements, was nevertheless denied the right to vote, even by the Dutch High Council (the supreme court of the Netherlands). The court’s reasoning was that the terms “Netherlander” and “subject” here had to be understood as referring to men only, “because were this not the case . . . it would undoubtedly have been stated clearly and unequivocally”.¹⁰ In other words, the gender-neutral terms *should not be assumed* to apply to women, for if they did apply to women, this would have been explicitly mentioned. In a similar case, the Canadian Supreme Court came to the same conclusion.¹¹

These supreme courts used as their principle the exact *opposite* of the idea that gender-neutral terms such as “Mensch” and “Netherlander” *should be assumed to include* both men and women. This is important to keep in mind

when we 21st-century readers approach historical texts. In many contexts, the general term *should be assumed to exclude* women unless indicated otherwise. This does not mean, of course, that women are always excluded – Aletta Jacobs did have to pay her taxes. But *whether* women are included or excluded can be determined only by reference to the background assumptions and wider context in which the general terms are used.

Nevertheless, the gender-neutral and race-neutral terminology creates a *gap* or *tension* between the “neutral” wording of principles and their application as restricted by unstated assumptions. Precisely this tension made it possible for Olympe de Gouges and Aletta Jacobs to appeal to the general terminology in claiming women’s rights. Kant produces a similar tension between the general wording of his highest principles and the tacitly assumed restrictions that become visible only in their application. At the same time, given the general terminology in which he articulates these principles, their scope as stated extends well beyond his own restrictive construal of their application.

Because of this built-in tension in Kant’s principles (and in the theory he develops on their basis), we can use these principles to criticize Kant himself and say that Kant “violates his own principles”. We could not do this were we to follow Mills’s suggestion and “translate” Kant’s formulations into the language of white male superiority. Nor, I should add, could Mills develop his own version of black radical Kantianism on the basis of the resulting white male supremacist theory. Mills, too, makes use of the critical potential implicit in Kant’s general formulations (Mills 2019).

The built-in tension in Kant’s theory also makes it possible to argue that Kant is *inconsistent* at least in important respects. If we take the principles as stated, *abstracting* from the unstated assumptions that restrict their application, then we can show, for example, that Kant’s Formula of Humanity contradicts his condoning non-white slavery. The argument can go as follows. Per definition, all human races share the essential characteristics common to all humans as such, and Kant explicitly says as much in the context of his 1785 essay on race:

Properties that belong to the species itself in its essence, and which are hence common to all human beings as such, are inevitably hereditary; but because human beings do not differ with regard to these properties, these will be kept out of the discussion of the subdivision of the races.

(BBM 8:99)

Furthermore, Kant argues that the essential characteristics common to all humans as such include their rational nature. He also justifies the prohibition against using others merely as means in terms of humans’ rational nature or even simply in terms of their being human (GMS 4: 429–430). People who use other human beings as slaves use them as mere means.

Hence Kant's condoning of non-white slavery *contradicts* his own prohibition against using other humans merely as means.

The fact that such contradictions can be pointed out, however, by no means implies that we can "remove and set aside" Kant's racism and sexism. The contradiction is generated by *abstracting* from the racist and sexist assumptions guiding their application. Kant does not apply the principles in an egalitarian way, and we should not lose sight of that fact when we focus on his race-neutral and gender-neutral formulations. It would be very odd to assume, say, that Kant would seriously defend *genuine racism* during his 1787 Physical Geography and Anthropology classes but would seriously defend *genuine egalitarianism* before and after class while working on his *Critique of Practical Reason*. We would have to imagine him switching back and forth between opposing positions on the same day, each day. Furthermore, simply setting aside his sexism and racism ignores how they influence his ethics and political theory more broadly, as I will argue in the next section.

What we need, therefore, is a middle position: we should acknowledge the tension between the general phrasing of Kant's principles and the unstated restrictions in their application. If we translate the Categorical Imperative into a principle for white males, we lose one side of this tension; if we remove and set aside Kant's racism and sexism, we lose the other.

3. How to avoid distortions and replications of Kant's sexism and racism

McCabe and many others assume that *if* Kant can be shown to be *inconsistent* in the sense that his racist and sexist biases *contradict* his philosophical principles, we can turn our philosophical focus solely to his egalitarian theory as the only thing that is philosophically significant and fruitful. I would now like to show, on the basis of the results of the previous sections, that this assumption can become highly counterproductive.

3.1. *The dangers of inclusive language and female pronouns*

If we focus on Kant's egalitarian-worded principles while disregarding his sexual and racial hierarchies, there will be a strong temptation to use inclusive language in our discussions. After all, if Kant's *principles* are egalitarian (and contradicted by things he says about gender and race), then why not use inclusive language in our discussion of them?

The use of explicitly gender-inclusive language, and even the exclusive use of female pronouns in discussions of Kant's ethics and political philosophy, has indeed become customary in recent years. The intentions behind this change are laudable, but the practice is nevertheless problematic.

The first problem is that of *misrepresentation*. If we focus on Kant's principles while abstracting from his objectionable views on racial and sexual hierarchy,

we risk depicting Kant as defending claims that he did not defend or even as defending claims he actively argued against. Let me develop one example.

Consider the “innate right to freedom”, as formulated in the *Metaphysics of Morals*, in some more detail:

Freedom (independence from being compelled by the arbitrary will of another), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right belonging to every human being [*jedem Menschen*] by virtue of his humanity [*Menschheit*].

(MSRL 6:237)

This innate right underlies much of Kant's political philosophy. In order to realize and secure this right to freedom, he argues, citizens ought to unite and collectively give themselves the laws they ought to obey (MSRL 6: 313–314; 340–341), for if the laws are *their own* laws, then they are independent from being compelled by the will of *another*. The contrast case is despotism.

As formulated, this innate right is not restricted to males; indeed, Kant states explicitly that “every human being” has this right by virtue of his humanity. Yet Kant relegates women to civil dependence on their husbands or male guardians and to perpetual passive citizenship, meaning that *they* cannot take any part in legislating the laws they ought to obey. Assume, first, that our philosophical interest is in *Kant's principles*: our interest is in the innate right to freedom, and we wish to abstract from Kant's obsolete claims about the subordinate status of women, claims that are arguably in tension with his assertion of this “human” right. Then what? Then, when discussing the innate right, we should most certainly *not* write sentences such as the following: “Kant regards every citizen as entitled to cast *his or her* vote”, or “Kant argues that every servant has the right to cancel *his or her* contract”. He did not regard all women as entitled to the right to vote and to cancel contracts on their own. For the same reason, we should *not* say of Kant's citizen that *she* has a right to independence or of the status of a servant that *she* has the right to cancel her contract. Using inclusive language and female pronouns makes these sentences downright false.

Now assume, alternatively, that our interest is in *Kant's sexism*: our interest is in the way Kant justifies the subjection of women and their perpetual dependence, and we are aware of the gender-neutral language Kant uses when formulating the innate right to freedom. Then what? Then, discussing the “innate right to freedom”, we should *not* write that this is a right Kant argued to be *innate to* (economically independent) *males alone*. The reason is that describing the right in this way prevents readers from seeing the tension between Kant's sexism on the one hand and the egalitarian wording of the innate right to freedom on the other. For the same reason, we should not follow Mills's recommendation that we “translate” the Categorical Imperative.

We should *not* write, for example: “According to the Categorical Imperative, white males ought to act only on maxims they can simultaneously will as universal laws for white males”. There are tensions between Kant’s egalitarian principles and his claims about gender and race, and we should portray Kant as neither more nor less of an (in)egalitarian than he was.

3.2. *Ignorance and the danger of inadvertent replication*

A second danger, closely connected with the first, is that of self-incurred *philosophical naïveté* and the unintentional replication of elements of Kant’s theories that were developed under the influence of his racism and sexism. The use of inclusive language or exclusively female pronouns not only risks misrepresenting Kant’s position but also mutes signals that might otherwise encourage further critical philosophical scrutiny. By making Kant sound like a decent egalitarian, it makes it less likely that readers will recognize that his racist and sexist prejudices had a profound impact on the shape of his moral and political theory as a whole.

Consider the issue of domestic labor. Kant assumed that within marriage, wives or servants would take care of such tasks. He did not problematize the gendered division of labor as a live philosophical issue that deserves discussion. If we now wish to use Kant’s political theory for our own philosophical purposes, we should do more than merely remove Kant’s essentialist descriptions of women’s character. We also need to dig up and problematize associated assumptions and consider the traces they have left. We need to *rethink*, for instance, Kant’s way of distinguishing between the private and the public spheres, his gendered conceptualization of what counts as labor and how various tasks should be divided among members of society, the status of servants as dependents in the household, Kant’s identification of human virtues and “masculine virtues” (and the omission of “feminine virtues”) and so on. If we do not do this, we risk reproducing elements that Kant originally introduced on the basis of sexist and racist assumptions, many of which remain implicit much of the time.

Theorists who want to *use* Kant’s theory for present purposes must engage with Kant’s own racism and sexism, if only to avoid replicating their effects in their own work. These elements of Kant’s thought should not be regarded as mere items in the history of philosophy museum of bigotry. Dale Spencer and Susan Moller Okin have warned against the tendency in recent political theory to just “add women and stir”. It is equally important to guard against the related tendency to just “remove sexism and racism and set aside”.

For an example of the type of theoretical remodeling that may be required, we can point to Kant’s own writings. Kant reworked his international political theory when he dropped the idea of white superiority. In *Toward Perpetual Peace* and the *Metaphysics of Morals*, he introduced a novel principle of public right, namely the “cosmopolitan right” mentioned previously. This

right specifically rules out colonial conquest and explicitly attributes full juridical status to humans on all continents (although, again, he still simultaneously defended the subordination of women). In other words, Kant not only removed any talk of a racial hierarchy but also added something new to his theory of public right as a necessary adjustment.

This is not to say that Kant went far enough in making adjustments to overcome his earlier racism. As Peter Niesen (2014) has argued, Kant fails to consider the issue of restorative justice. Charles Mills (2019) has developed Kantianism further by incorporating key elements of Afro-modern thought and experience. Elvira Basevich (2020) has expanded Kant's model of public reason to develop a model of interracial civic fellowship. These are a few examples of constructive transformations of Kant's approach that go further than Kant did himself.¹²

Furthermore, we should make visible, in Kant's work as well as our own use of it, the compounded effects of multiple forms of subordination. Kant discusses the "character" of the races and the sexes separately, but on closer inspection, it turns out that his discussion of the races is in fact a discussion of the allegedly different characteristics of their *male* members alone. In order to recognize this pattern, we need to go beyond Kant's discussions of race and examine his discussion of the sexes in *relation* to his discussion of race. Conversely, his characterization of women in the early anthropology lectures does not mention the alleged deficits of "yellow, black, red" or "mixed"-race women, but this does not mean that it is race neutral.

Thus, we should not take the structure of Kant's theory as a given when using it for our current philosophical purposes, believing that we can simply set aside Kant's objectionable views. We need to ask ourselves what further adjustments must be made. To be able to answer that question, we need to know exactly how and where his racism and sexism influenced his moral and political theory.

In sum, if we want to use Kant's work for current philosophical purposes without unwittingly reproducing some of his biases, we need to *research* his views on sexual and racial difference, *reveal* the unstated assumptions that guide his application of egalitarian principles, *reconstruct* the influence of these assumptions on the shape of his philosophical theories and *remodel* these theories where necessary.

Notes

- 1 In this chapter, I focus only on racism and sexism in Kant's work because these are the focus of the special issue of *SGIR Review* in which this chapter was first published (*SGIR Review* 2 [2019], 3–22). Many of the arguments developed in this chapter can be extended to other aspects of Kant's work and to the work of other figures in the history of philosophy. I thank Gerad Gentry for inviting me to write this chapter, in which I combine, further elaborate and extend several arguments originally introduced in separate contexts. I thank him for organizing

the panel at the 2019 Eastern APA where this chapter was first presented, and I am also grateful to the audience at that session, as well as to Elvira Basevich, Carolyn Benson, Michael Gregory, Suzanne Jacobi, Marijana Vujosevic and Lieuwe Zijlstra for helpful comments. I thank Gerad Gentry for permission to have this chapter republished in the present volume.

- 2 I have discussed Kant's views on race and gender in more detail in Kleingeld 1993, 2007, 2014, and I draw on these papers in this section.
- 3 References to Kant are to *Kants gesammelte Schriften*, edited by the Prussian (later German) Academy of Sciences (Berlin: Georg Reimer, subsequently Walter de Gruyter, 1900–). Kant's writings are cited by the abbreviated title, followed by the Academy volume and page number. Some (parts of) lectures on physical geography are excluded from volume 26 of the Academy Edition. I cite these here by abbreviated title, name of the transcript, and page number. These lectures are available on <http://kant.bbaw.de/base.htm>. Translations are my own, but I have made use of the Cambridge Edition of the Writings of Immanuel Kant.
- 4 I thank Kate Moran for the reference to the Parow lectures.
- 5 Mari Mikkola (2011, 102) claims that in this sentence, Kant gives an example of women acting on *moral principles*. But given the reference to “their household” and to “a character *suitable to their vocation*” rather than character simpliciter, this does not seem to be the case. Moreover, immediately following this passage, Kant gives several examples of men (Milton, Socrates) who displayed “masculine virtue” when faced with unlawful or dishonorable requests from their wives and who did so “without diminishing the merit of feminine virtue” (Anth 7:308; where Kant also presents Milton's rather condescending response to his wife).
- 6 By “slavery”, I mean chattel slavery. Kant distinguishes this from slavery imposed as criminal punishment. In the Doctrine of Right, he argues that this punishment is permissible within (very wide) limits and as long as it is imposed only on the person who committed the crime (MSRL 6:329–330). Presumably he refers to the use of prisoners for forced labor, but in this chapter, I leave this issue to the side.
- 7 Black feminist theorists have long pressed this point. See Crenshaw 1989; Davis 1982; hooks 1984; Collins 2000. For more recent discussions, see Alcoff 2006; Collins and Bilge 2016.
- 8 She also argues that Kant's views on women are not as dim as they are usually made out to be, Mikkola 2011; see also Varden 2017.
- 9 Whether Kant expressed this sentence exactly in these words when he gave his lecture does not matter for the purpose of my argument. My point here is general and applies both to the person who transcribed the lecture and to Kant.
- 10 *Weekblad van het regt* 4917, 7 August, 1883, p. 1.
- 11 As mentioned by Richard Rorty 1990, 5, n.6.
- 12 Of course, one can also use Kant to move beyond Kant without focusing on the structure of his own theory but rather by using it to tell us something about the nature of racism (Allais 2016) or by using Kant's notion of self-respect for purposes of feminist theory (Hay 2013).

References

- Alcoff, L.M. 2006. *Visible Identities: Race, Gender and the Self*. Oxford: Oxford University Press.
- Allais, L. 2016. “Kant's Racism.” *Philosophical Papers* 45: 1–36.
- Basevich, E. 2020. “Reckoning with Kant on Race.” *Philosophical Forum* 51: 221–245.

- Bernasconi, R. 2001. "Who Invented the Concept of Race? Kant's Role in the Enlightenment Construction of Race." In R. Bernasconi (ed.), *Race*. Oxford: Blackwell, 11–36.
- Collins, P. Hill. 2000. *Black Feminist Thought: Knowledge, Consciousness and the Politics of Empowerment*. New York: Routledge.
- Collins, P. Hill and Bilge, S. (eds.). 2016. *Intersectionality*. Cambridge: Polity Press.
- Crenshaw, K. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum* 1: 139–168.
- Davis, A. 1982. *Women, Race, and Class*. London: The Women's Press.
- Eze, E.C. 1994. "The Color of Reason: The Idea of 'Race' in Kant's Anthropology." In K.M. Faulstich (ed.), *Anthropology and the German Enlightenment*. Lewisburg: Bucknell University Press, 200–241.
- Hay, C. 2013. *Kantianism, Liberalism, and Feminism: Resisting Oppression*. New York: Palgrave Macmillan.
- hooks, b. 1984. *Feminist Theory: From Margin to Center*. Cambridge: South End.
- Kleingeld, P. 1993. "The Problematic Status of Gender-Neutral Language in the History of Philosophy: The Case of Kant." *The Philosophical Forum* 25: 134–150.
- Kleingeld, P. 2007. "Kant's Second Thoughts on Race." *The Philosophical Quarterly* 57: 573–592.
- Kleingeld, P. 2014. "Kant's Second Thoughts on Colonialism." In K. Flikschuh and L. Ypi (eds.), *Kant and Colonialism: Historical and Critical Perspectives*. Oxford: Oxford University Press, 43–67.
- Louden, R. 2000. *Kant's Impure Ethics*. Oxford: Oxford University Press.
- McCabe, D. 2019. "Kant Was a Racist: Now What?" *APA Newsletter on Teaching Philosophy* 18: 2–9.
- Mikkola, M. 2011. "Kant on Moral Agency and Women's Nature." *Kantian Review* 16: 89–111.
- Mills, C. 2005. "Kant's *Untermenschen*." In *Race and Racism in Modern Philosophy*. Ithaca: Cornell University Press, 169–193.
- Mills, C. 2019. "Black Radical Kantianism." *SGIR Review* 2(2): 23–64.
- Niesen, P. 2014. "Restorative Justice in International and Cosmopolitan Law." In K. Flikschuh and L. Ypi (eds.), *Kant and Colonialism*. Oxford: Oxford University Press, 170–196.
- Rorty, R. 1990. "Feminism and Pragmatism." *The Tanner Lectures on Human Values, Delivered at the University of Michigan*, December 7.
- Varden, H. 2017. "Kant and Women." *Pacific Philosophical Quarterly* 98: 653–694.

KANT THE NATURALIST

Robert B. Loudon

Kant is widely admired (and sometimes also widely despised) as the founding father of transcendental philosophy. But in much of my own writing on Kant, I have been concerned with a very different Kant: an impure rather than a pure Kant, an a posteriori rather than an a priori Kant, an empirical rather than a rationalist Kant, and, yes, a *naturalistic* rather than a transcendental Kant. This other Kant has often been overlooked, particularly by professional philosophers who (to borrow from Hume) when they think of Kant, “can scarcely forbear reflecting on” (2005, 24) the huge wave of transcendental philosophy which starts with him and follows in his wake. And in those relatively rare cases where the other Kant is not overlooked, he is often belittled and accused of being “shallow and unoriginal”, a “minor scribbler”¹ in comparison with what Moses Mendelssohn called the “all-destroying [*alleszermalmende*]” (1971–98: 3.2: 3) philosopher of the first *Critique*.

In what follows, I hope to convince you not only that Kant the naturalist does indeed exist but that he is also a philosophical force to be reckoned with. While there are admittedly new problems that arise once one sets out on this journey (e.g. how do these two Kants relate to each other? How can one be both a naturalist and a transcendental philosopher?), my own view is that Kantian studies become much richer and more substantive when the other Kant is brought into the picture. Two Kants are better than one.

1. Naturalism in the philosophy of education²

Let’s start with a few texts. Perhaps the clearest example of Kant’s naturalism is to be found in his philosophy of education. In the second of his two short essays on Johann Bernhard Basedow’s (1724–1790) famous experimental school in Dessau, the Philanthropin, founded in 1774, Kant praises the school in part because its “educational method . . . is wisely derived from nature itself [*aus der Natur selbst gezogen*] and not slavishly copied from old habit and unexperienced ages” (AP 2: 449).³ All other schools, Kant asserts, “were spoiled at the outset, . . . because everything in them works against nature [*der Natur entgegen arbeitet*]” (AP 2: 449; cf. 447). Small wonder,

then, that Kant elsewhere (viz., in the *Friedländer* anthropology lecture) singles out Basedow's school as "the greatest phenomenon that has appeared in this century for the improvement of the perfection of humanity" (V-Anth/Fried 25: 722; see also V-Mo/Collins 27: 471, Päd 9: 451).

Although Kant never went west to Dessau to observe Basedow's new school firsthand (he was not a traveler and spent his entire life in the vicinity of his hometown of Königsberg), he was well acquainted with at least some of Basedow's voluminous writings.⁴ For instance, he used one of Basedow's most famous works – *Das Methodenbuch für Väter und Mütter der Familien und Völker* (1770, 2nd ed. 1771, 3rd ed. 1773) [The Method-Book for Fathers and Mothers of Families and Nations] – as his text for his first course on practical pedagogy at the University of Königsberg in the winter semester of 1776–77. And Basedow also repeatedly emphasizes his naturalistic teaching method in many of his own publications. For instance, the second chapter of his *Magister* thesis for the University of Kiel, written in Latin and defended on June 7, 1752, is entitled: "The Unused and Natural Method of All Scholastic Studies, Chiefly Latin".⁵ And in the Preface of a 1774 publication intended to increase public awareness of his experimental school (viz., *Das in Dessau errichtete Philanthropinum, eine Schule der Menschenfreundschaft und guter Erkenntnis* – The Philanthropinum Established in Dessau, a School of Human Friendship and Good Knowledge), he writes:

Nature! [*Natur*] *School!* *Life!* Friendship is under these three; [and] so will the human being, what he should be, and cannot be right away; happy in childhood, cheerful and curious in adolescence, peaceful and useful as an adult. But when nature is whipped out of school, and school is mocked in the life of man, then in the end the human being grows into a deformity.

(Basedow 1774, XIII)

Similarly, in another of his most influential publications, the four-volume *Elementarwerk* (1774, 2nd ed. 1785), which was accompanied by nearly 100 copper engravings by Polish artist Daniel Chodowiecki (1726–1801), Basedow expresses the hope that his method of "natural education and instruction [*natürliche Erziehung und Unterweisung*] . . . will be introduced in public schools" (reprinted in Basedow 1965, 197). Many of the curricular innovations that Basedow introduced at the Philanthropin and for which it is justly famous – for instance, the conversational method of teaching foreign languages, the doctrine of learning through play and the introduction of both physical education and sex education into the curriculum – have their roots in his naturalism.⁶

In Enlightenment pedagogical theory, this emphasis on a naturalistic educational method is most commonly associated with Rousseau's *Émile* (1762). As a result, Basedow is often dismissed as a "relatively mediocre thinker" (Parker 1912, 216), an unoriginal German Rousseau. For instance, Robert Quick, in

his *Essays on Educational Reformers*, writes: “the root-ideas of Basedow put forth in his ‘Book of Method’ [viz., the *Methodenbuch*], and other writings, are those of Rousseau” (1896, 279). But while Rousseau in his *Émile* does advocate a general principle “of letting nature alone in everything”, and while he continually criticizes humans for wanting “nothing as nature made it, not even man” (1979, 131, 37; cf. 107), it should be noted that *Émile* was not published until 1762, that is, ten years after Basedow began advocating his “natural method” of education in his 1752 *Magister* thesis at the University of Kiel. And the material in Basedow’s 1752 thesis is based on his earlier personal experience as a private tutor in the von Qualen family home in Borghorst from 1749–52. So it is clearly not the case that Basedow’s naturalistic pedagogical theory is borrowed from Rousseau. As Joseph Landschoof rightly remarks: “While the name of Rousseau, to whom undeserved credit long has been given for any influence he is said to have had upon Basedow, was still unknown and unheralded; . . . Basedow, as a young and inexperienced teacher, was laying the foundations of [his own] method” (1933, 53f.).

And, in fact, this popular Enlightenment appeal to nature as a positive norm for human conduct extends far beyond the field of pedagogy. It is a key part of the Enlightenment’s break with pre-modern attitudes. For instance, Leopold Mozart (a.k.a. the father of Wolfgang Amadeus), in his famous treatise on the principles of violin playing, continually advises his readers to follow “nature herself [*von der Natur selbst*]” in determining how best to play their instrument,⁷ and Mozart’s book was published in 1756, also well before the appearance of Rousseau’s *Émile*.

So Kant, in his own appeals to nature as a norm for educational practice, is clearly part of a much bigger cultural phenomenon than either Basedow or Rousseau. For instance, in the 1765 *Announcement of the Program of His Lectures*, he declares his intention “to make public education more adapted to nature [*nach der Natur mehr zu bequemen*]” (NEV 2: 305). However, this remark predates neither Rousseau’s *Émile* nor the intense admiration for Rousseau that Kant expresses in his Remarks in the *Observations on the Feeling of the Beautiful and the Sublime* (1764).⁸ To complicate matters further, in Herder’s notes on Kant’s moral philosophy lectures from 1763–64, Kant does briefly mention Basedow.⁹ So it’s not clear to me whether Kant arrived at his own naturalist convictions in education independently of his familiarity with either Rousseau or Basedow. I would like to believe that he did, but I am not aware of any Kantian texts that can settle the matter. Again, though, the Enlightenment appeal to nature as a positive norm for human conduct is clearly something much bigger and more fundamental than Rousseau, Basedow or even Kant.

2. Humans-only norms

As Hume rightly remarks, “nature” is a term “than which there is none more ambiguous and equivocal”,¹⁰ and as a result, care must be exercised

in calling Kant a naturalist. There are many different kinds of naturalism. In applying this term to Kant, I do not mean to imply that he necessarily rejects all supernatural entities, that he thinks philosophical methods should not differ from the methods of the natural sciences or that he believes natural science offers all the knowledge that is humanly possible. Rather, I mean primarily that in his discussions of human practices, he frequently argues that many natural processes are good and departures from them are bad and that human practices should therefore try to follow nature rather than resist it or overcome it in those cases where nature is judged to be good. In earlier writings, I have called this “a species of *weak* naturalism” (Louden 2000, 145). Much that we find in nature is good, and when nature is good, humans should emulate it rather than reject it. This weaker variety of naturalism is not quite as gung-ho as some of the more reductive naturalisms that have been popular in recent years, but that may not be a bad thing. At the same time, however, I do believe that Kant’s qualified naturalism – a weaker or more liberal version of naturalism – is arguably more in tune with those contemporary accounts of naturalism that, while viewing naturalism as a positive force in discussions of human nature and morality, do not see it as a panacea that can magically answer every philosophical question (for related discussion, see De Caro and Macarthur 2010). To put it in more Kantian terms, Kantian naturalism is a *critical* naturalism, and “our age is, in especial degree, the age of criticism, to which everything must submit” (KrV A xii).

But calling nature “good” instantly raises the specter of naturalist doctrines regarding normativity. Am I claiming that Kant holds that the concept “good” can and should be *derived* from nature? To make such a claim would seem to fly in the face of some of Kant’s strongest-held commitments regarding the nature of moral norms. For instance, in the first *Critique*, he asserts that the moral ought expresses a species of necessity and a connection with grounds which does not occur anywhere in the whole of nature. In nature, the understanding can cognize only **what exists**, or has been, or will be. It is impossible that something in it **ought to be** other than what it in fact is; indeed, the **ought**, if one has merely the course of nature before one’s eyes, has no significance whatever (KrV A 547/B 575).¹¹

In Kant’s official view, genuine moral principles cannot be derived from empirical facts, regardless of how firmly established the latter may be. Any attempt to derive moral principles from empirical facts results in something much too parochial and flimsy. You cannot get to the categorical imperative merely from empirical facts. As he writes in the *Groundwork*:

Empirical principles are not fit to be the foundation of moral laws at all. For the universality with which they are to hold for all rational beings regardless of difference [*ohne Unterschied*] – the unconditional practical necessity that is thereby imposed upon them – vanishes if

their ground is taken from the *particular arrangement of human nature*, or the contingent circumstances in which it is placed.
(GMS 4: 442)

But in the remainder of this section, I wish to take a second look at Kant's position on moral normativity. Does he in fact – as the previous citations from the *Critique of Pure Reason* and the *Groundwork of the Metaphysics of Morals* seem to imply – completely reject any and all naturalistic strategies to ground moral norms in nature? Or is his considered position actually a bit more complicated than this? In what follows, I shall argue that Kant, contrary to received doctrine, frequently does rely on a type of norm that I call “humans-only norms”.¹² Humans-only norms are impure, a posteriori and empirical – they are based on general facts about human nature and the world they live in. And because they are based on general facts about human nature and world in which they live, these norms are also naturalist norms in the sense indicated earlier: they are norms based on natural processes which Kant believes are good. Furthermore, some of these humans-only norms, or so I shall argue, can also legitimately be called a type of *moral* norm, even though they are not categorical imperatives that “must hold not merely for human beings but for *all rational beings in general* [*alle vernünftige Wesen überhaupt*]” (GMS 4: 408; see also 410 n., 412, 426, 431) and even though they have not been “completely cleansed of everything that may be only empirical and belongs to anthropology” (GMS 4: 389).

For the most part, Kant's humans-only norms are located in what he calls “the second part” of morality, “moral anthropology, to which the empirical principles belong” (V-Mo/Mron II 29: 599). The second part of morality remains to this day under-explored territory in Kant scholarship, because most philosophers, as Richard Rorty observed, are intent on “keeping philosophy pure” (1982), and so they tend to focus exclusively on what Kant calls “the metaphysics of morals, or *metaphysica pura*”. But *metaphysica pura*, Kant also notes, “is only the first part of morals” – the second part, again, is “moral anthropology, to which the empirical principles belong. . . . Moral anthropology is morals that are applied to human beings” (V-Mo/Mron II 29: 599).

Some examples of humans-only norms (or what Kant might call “*Menschenpflichten*” – see KpV 5: 8) taken from his own texts include the following:

- 1 *Politeness*. Humans ought to practice politeness because politeness promotes virtue. The practice of politeness “enables us to deceive the deceiver in ourselves, the inclinations” (Anth 7: 152). If we are able to practice politeness successfully, we may be able to trick our inclinations into following the demands of practical reason. And if this trick is performed successfully, we are on the road to moral virtue. As Kant states in his *Anthropology*: “In order to save virtue, or at least lead the human being to it, nature has wisely

implanted in him the tendency to allow himself willingly to be deceived” (Anth 7: 152; cf. MS 6: 473–4). Part of his point here is that humans are built in this specific way, while other types of rational agents might not be. It is (or so he claims) a contingent, empirical fact about normal members of the human species that they are able to trick their inclinations into doing reason’s bidding by acting politely. Other rational beings might be built differently. And part of the Kantian anthropologist’s job is to alert readers to this general feature of human nature.¹³

- 2 *Education.* In the opening sentence of his *Lectures on Pedagogy*, Kant proclaims: “the human being is the only creature that must be educated [*das einzige Geschöpf, das erzogen werden muß*]” (Päd 9: 441). Granted, the claim may be false – perhaps other creatures, terrestrial or otherwise, must also be educated. Not all of the empirical generalizations made by anthropologists (Kantian or otherwise) turn out to be true. But that’s another story. Kant’s main point is that it is the anthropologist’s job to inform us about our shared nature¹⁴ and to do so with an eye toward helping us to better meet morality’s demands. As he states in the *Metaphysics of Morals*:

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 *carrying out* [*Ausführung*] the laws of a metaphysics of morals. It would deal with the development, spreading, and strengthening of moral principles (in education, in schools, and in popular instruction) and with similar other teachings and precepts based on experience [*auf Erfahrung gründende Lehre und Vorschriften*].

(MS 6: 217)

In other words, what is it about this particular species that makes it difficult for them to act on moral principle (= “hindrances”)? Additionally, what empirical features in our nature might make morality easier for us than it is for other types of rational beings (= “helps”)?

A second humans-only norm (or rather, an *allegedly* humans-only norm: Kantian anthropologists, like other humans, are fallible creatures, and they do not always have their facts straight) articulated in the *Lectures on Pedagogy* is Kant’s claim that “the human species ought [*soll*] to bring out, little by little, humanity’s entire natural predisposition by means of its own effort” (Päd 9: 441). The underlying claim is that this gradual unfolding of predispositions over the course of many generations does not happen with other species. With us it is a collective achievement known as “culture”,¹⁵ something we do not see in other terrestrial creatures. As he remarks in the *Anthropology*:

With all other animals left to themselves, each individual reaches its complete destiny [*seine ganze Bestimmung erreicht*]; however with the human being only the species, at best, reaches it; so that the human race can work its way up to its destiny only through progress in a series of innumerable many generations.

(Anth 7: 324; cf. V-Anth/Mensch 25: 1196,
V-Anth/Mron 25: 1417)

- 3 *Aesthetics*. In the third *Critique* as well as in several anthropology transcriptions, Kant proclaims that beauty – the central concepts of aesthetics – “is valid only for human beings [*nur für Menschen*], i.e., animal but also rational beings” (KU 5: 210; cf. V-Anth/Collins 25: 175, V-Anth/Mensch 25: 1108, V-Anth/Busolt 25: 1513). He also argues that aesthetic experience helps humans to develop their capacity for moral judgment. “The culture of taste is a preparatory exercise [*Vorübung*] for morality”¹⁶ (Ref 993, 15: 438). Beauty is – for humans, but perhaps not for other creatures who are built differently than us – a symbol of morality, for the human experience of “the beautiful prepares us to love something, even nature, without interest” (KU 5: 267). “Without interest” – that is, for its own sake rather than as a means to something else. Aesthetic experience teaches us how to freely love something for its own sake, and this is a crucial aspect of moral judgments. The proper appreciation of works of art, at least for humans, is thus a means to morality.

The previous examples of humans-only norms taken from Kant’s texts are all based on empirical generalizations about our distinct nature. And because they contain norms about how we ought to behave that are based on our nature, they are a type of naturalist norm. The ought comes from our nature, or rather, we are calling an aspect of our nature good and then saying that we ought to follow it rather than work against it. (Note that I am making a distinction here between *calling* nature good and *deriving* good from nature. The Kantian weak naturalist prefers the former route, the strong naturalist the latter.)

But what makes humans-only norms *moral* norms? Not all norms are moral norms. There are also legal norms, social norms, norms of etiquette, linguistic norms and so on. (As Hilary Putnam supposedly says somewhere, “normativity is ubiquitous in our thought and talk”.)¹⁷ However, I think it is clear that some of Kant’s humans-only norms, including each of the examples discussed previously, are in fact moral norms – even though they are impure rather than pure, a posteriori rather than a priori. Why is this so? Those humans-only norms that are necessary means to obligatory ends are also *moral* norms because, as Kant states in the *Groundwork*, “whoever wills the end also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are in his control” (GMS 4: 417). Granted, “skill in the choice of means” (GMS 4: 416) is

often associated merely with prudence¹⁸ rather than morality and thus with hypothetical rather than categorical imperatives. But in Kant's view, not all ends are optional – we also have “ends that are also duties” (MS 6: 385) – viz., “*one's own perfection and the happiness of others*” (MS 6: 385). And in cases where we are pursuing nonoptional, morally obligatory ends, we are also obligated to pursue the necessary means toward these ends – it is irrational not to do so; a sign of “volitional inconsistency” (O'Neill 1989, 91).¹⁹ And in each of the examples discussed earlier, we are dealing with a norm that tells humans what they must do to promote a necessary moral end. Politeness is (for humans) a means to virtue, for it enables us to fool our inclinations. Aesthetic experience is (for humans) a means toward moral judgment, for it teaches us how to love something for its own sake. And the two examples of humans-only norms in the sphere of education are means toward a moral end because humans must be educated into morality. We are not born as autonomous moral agents. Rather, we develop our moral reasoning capacities slowly over a number of years, through a complex and extensive process of moral education.

So while not all humans-only norms are moral norms (some might be norms restricted to other spheres of human life such as language or etiquette), the ones described previously are. But these norms are not quite categorical, for they apply only to humans and not to rational agents in general, and Kant is adamant that all genuine moral laws apply to rational beings throughout the universe. “Everyone must admit [*Jedermann muß eingestehen*] that a law, if it is to hold morally, . . . does not just hold for human beings, as if other rational beings did not have to heed it; and so with all remaining genuine moral laws [*alle übrige eigentliche Sittengesetze*]” (GMS 4: 389). However, they are also not quite hypothetical. Hypothetical imperatives are desire-based commands (“if you want X, then you must do Y”). But in the examples discussed previously, we are dealing with norms that have the following structure: “if you're a human being, then you must do Y”. Because the antecedent does not describe a subjective desire, one cannot evade the consequent simply by changing one's desires. (“I no longer desire X. Therefore, I'm not obligated to do Y”.) Those humans-only norms that are also necessary means to moral ends are – for humans but not for other types of rational being – inescapable duties. We can't escape from them for the simple reason that we can't escape from our humanity.²⁰ This is who we are.

3. The eye of true philosophy²¹

Although my remarks have for the most part focused on the contested issue of normativity and naturalism rather than on naturalism *überhaupt*, it should be clear by now that Kant the naturalist does indeed exist. Kant is not just a transcendental idealist. He wears at least two hats, and a great

deal of his teaching and research was in fact devoted to empirical work. His empirical work deals not just with nonhuman nature but also with human nature, and when writing on the latter topic, he frequently appeals to a kind of naturalist norm that applies only to humans and in some cases is also a *moral* norm – his occasional pronouncements that there are “no oughts in nature” notwithstanding (see KrV a 547/B 575).

But what about the earlier-mentioned charge²² of “shallowness and unoriginality” – that is, the acknowledgement that while Kant the naturalist does exist, his work in this area is the product of a “minor scribbler” who is philosophically uninteresting and unimportant and who pales in comparison with the transcendental Kant of the three Critiques? In this final section of my chapter, I shall challenge this popular assumption. Kant the naturalist does not just exist. He is also a philosophical force to be reckoned with, an important and original thinker whose insights add value to the significance of his overall system.

In a much-discussed *Reflexion* that has recently become the subject of an entire book (see Tommasi 2018), Kant criticizes an increasingly familiar kind of scholar who lacks humanity and as a result “misjudges himself and trusts his own powers too much” (Refl 903, 15: 395). Max Weber calls such a person a “specialist without spirit” (1958, 182), but Kant says: “I call such a person a Cyclops” (Refl 903, 15: 395). This one-eyed giant (Kant is referring to a famous passage from Homer’s *Odyssey*),²³ he adds, “needs another eye, so that he can consider his object from the point of view of other human beings” (Refl 903, 15: 395). This necessary second eye, which is precisely what the scholarly Cyclops lacks, is what “grounds the humanity of the sciences; that is, gives them the affability of judgment [*die Leutseligkeit des Urteils*], through which one submits to the judgment of others” (Refl 903, 15: 395).

The one-eyed scholar, a.k.a. the Weberian specialist without spirit, thus needs to cultivate some humanity by cultivating what Kant elsewhere calls “the broadened way of thinking”, that is, “thinking from the position of everyone else [*an der Stelle jedes andern denken*]” (KU 5: 294; cf. Log 9: 57, Anth 7: 200). This broadened way of thinking, he adds, though by no means easy to achieve, is nevertheless one of the three most fundamental “maxims of common understanding [*Maximen des gemeinen Menschenverstandes*]” (KU 5: 294).

Kant also invokes his Cyclops and second eye metaphors in several other texts, but here his advice to the one-eyed scholar is slightly different. For instance, in his discussion of genius in the *Anthropology*, he writes:

There is also *gigantic* erudition, which is nevertheless often *cyclopean*, that is to say, missing one eye: namely the eye of true philosophy [*das der wahren Philosophie*], by means of which human

reason appropriately [*zweckmäßig*] uses this mass of historical knowledge, the load of a hundred camels.

(Anth 7: 227)

Similarly, in the Jäsche logic lecture (1802), he states:

Mere polyhistory is a *cyclopic* learnedness, which lacks one eye, the eye of philosophy [*das Auge der Philosophie*], and a Cyclops among mathematicians, historians, natural historians, philologists, and linguists is a scholar who is great in all these matters, but who for all that holds philosophy to be dispensable.

(Log 9: 45)

The underlying message in these passages is that one-eyed scholars (who are unfortunately all too well represented in academic philosophy and elsewhere at present) need to acquire a broader, more humanistic way of thinking, one which Kant elsewhere calls “philosophy in the *cosmopolitan* sense [*in sensu cosmopolitico*]” and “the science of the ultimate ends of human reason” (V-Met-L2/Pölitz 28: 532; cf. Log 9: 23). This “high concept” of philosophy in the cosmopolitan sense, he adds “gives philosophy dignity [*Würde*], that is, an absolute worth” (Log 9: 23; see also GMS 4: 434–5), and it also “gives worth to all other sciences” (V-Met-L2/Pölitz 28: 532; cf. Log 9: 24).

But philosophy in this essential cosmopolitan and humanistic sense is also necessarily empirical and historical – and, in the qualified sense in which I have been using the term in the present chapter, *naturalistic*, for it includes an awareness of those humans-only norms which help us to achieve our essential ends. In other words, it is what Kant himself calls the “worldly concept [*Weltbegriffe*] of philosophy” (Log 9: 23) – a conception of philosophy informed by moral anthropology, through which we learn about our species’ destiny (*Bestimmung*) and the ultimate ends of human reason. And in order to avoid the fate of the Cyclops (who, in Homer’s account, had his eye burned out when Odysseus twisted a burning stick in it – but, like Kant, I am only speaking metaphorically at present!), it is necessary that one-eyed scholars in all disciplines (philosophy included) cultivate this empirically informed eye of true philosophy. Only then will their science have dignity, that is, an “inner worth” (GMS 4: 435) beyond all price.

A more mundane way of putting some of these points is to remind ourselves of Kant’s remark in the Preface to the *Groundwork* that “natural as well as moral philosophy each have their empirical part” (GMS 4: 387). Since Kant’s day, philosophers have increasingly abandoned the empirical parts of natural and moral philosophy to their younger colleagues in the natural and human sciences. However, this was not Kant’s wish, even though he himself (due to the growing interest in the transcendental side of

his philosophical system) is partly the cause of this unfortunate development.²⁴ Throughout his career, Kant displayed a strong interest in empirical work, particularly empirical work on human nature.²⁵ And, as I have shown, he worried that philosophy and science would be taken over by one-eyed scholars (or what William Blake calls “single vision”²⁶ theorists) unless its practitioners cultivate “the eye of true philosophy”. This eye of true philosophy is an empirical eye that gives its possessors a “broadened way of thinking”, one which includes a knowledge of human nature and of humanity’s essential ends.

So, in conclusion, Kant is both a transcendental philosopher and a naturalist. Granted, he was adamant that pure philosophy “must come first” (GMS 4: 390), and because of this, pride of place should always be given to the transcendental side of his project. But it is time to give the naturalist side its due. How these two different sides of Kant’s philosophical system fit together is a thorny question that lies beyond the scope of the present chapter, but in closing, perhaps a few very brief words are nevertheless in order. The general strategy for reconciling the two Kants – and I think Kant himself follows this strategy – is simply to insist that the transcendental and naturalist sides not encroach on each other’s territory. Each side needs to respect the property rights of the other. Kant’s transcendental philosophy deals with concepts and principles that are allegedly necessary conditions for the possibility of experience, while his naturalism, in the manner presented here, deals with the more modest task of analyzing and evaluating nature, with the goal of determining which aspects of it should be judged good and embraced as models for human living. In principle, these are two very different tasks that need not conflict with one another. Indeed, there is reason to believe that they can be good neighbors. At any rate, they need to at least believe that it is possible for them to live together amicably, for since true philosophy has both a transcendental and a naturalist side, it is necessary and important to sustain both.²⁷

Notes

- 1 Fleischacker (2018). Fleischacker’s comments occur in his remarks about my contribution to the volume that he is reviewing, that is, *Kant and the Scottish Enlightenment* – see Loudon (2017b).
- 2 In this section, I borrow a few points from Loudon (2017a, 724f., 2011, 144f.).
- 3 Kant’s works are cited in the body of the text according to volume and page number in the Academy Edition of his writings, using the standard German abbreviations for individual titles. Translations are my own. Citations from Kant’s *Critique of Pure Reason* (KrV) refer to the A (1st) and B (2nd) editions of this work. For further discussion of Kant’s debts to Basedow, see Loudon 2021c.
- 4 Basedow published well over 100 works during his life. See the bibliography in Basedow (1965, 267–274).
- 5 Basedow (1752), cf. Basedow (1965, 267).

- 6 Commentators also frequently draw attention to the naturalist orientation of Basedow's school. Friedrich Paulsen, for instance, in *German Education: Past and Present*, notes that Basedow's principal aim "was to make room . . . for an education that was in touch with real life and in accordance with human nature. The life of the pupils was arranged with a view to conformity with Nature" (1908, 134). Similarly, Frank Pierrepont Graves writes: "The unifying principle of the school was 'everything according to nature'. The natural instincts and interests of children were only to be directed and not altogether suppressed. They were to be trained as children and not as adults, and the methods of learning were to be adapted to their stage of mentality" (1912, 116f.).
- 7 Mozart (1756, 193; cf. 238 – "*aus der Natur selbst*").
- 8 In a famous note that Kant wrote in the margins of his own copy of *Observations on the Feeling of the Beautiful and the Sublime* (1764), he states: "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).
- 9 "*Syncretism*: here one tries to develop contradictory doctrines as if they were in agreement. It seldom catches on; is usually futile and often damaging – Basedow is syncretic" (V-PP/Herder 27: 88).
- 10 Hume (1978, 474). For related discussion, see Fink (2006).
- 11 A related concern about discussing normativity from a naturalist perspective, which, due to space limitations, I mention only in passing, is the assumption that naturalism implies causal determinism. Kant, in *Religion within the Boundaries of Mere Reason* (1793), alludes to this issue when he writes: "But lest anyone be immediately scandalized by the expression *nature*, which would stand in direct contradiction to the predicated *morally* good or *morally* evil if taken to mean (as it usually does) the opposite of the grounds of actions from *freedom*, let it be noted that by 'the nature of the human being' we only understand here the subjective ground . . . of the exercise of the human being's freedom in general" (RGV 6: 21). Needless to say, in the present discussion of Kant's naturalism, I am not using the term "naturalism" in a manner that implies determinism or rules out responsibility for voluntary behavior.
- 12 Some of the following material was first presented in Loudon 2018b, 2021a.
- 13 For related discussion, see Loudon 2021b.
- 14 Kantian anthropology differs from much contemporary cultural anthropology in its focus on universal rather than culturally relative traits. In Kant's view, a proper or correct anthropology "is not a local, but a general anthropology. In it one comes to know the nature of humanity, not the condition of human beings. . . . Anthropology is not a description of human beings, but of human nature" (V-Anth/Fried 25: 471; cf. V-Anth/Pillau 25: 734).
- 15 For more on Kant's concept of culture, see Loudon Forthcoming.
- 16 Paul Guyer, who heard a version of this chapter presented at the University of Catania, Sicily in October 2018, remarked during discussion afterward that not all humans who are exposed to art become morally good agents. I agree. My point is simply that Kant believes that this art-morality connection holds. But his belief may of course be false.

- 17 See De Caro and Macarthur (2010, 9). (They attribute the remark to Putnam without indicating its source). For discussion of different types of norms, see Brennan et al. (2013).
- 18 As Kant himself remarks in the *Mrongovius* anthropology lecture, “prudence [*Klugheit*] is a proficiency or knowledge in reaching one’s aims” (V-Anth/Mron 25: 1210).
- 19 For related discussion, see Louden (2003, 85–90).
- 20 Or at least we can’t quite escape from our humanity at present. If the dreams of posthumanism come true, this claim might become falsifiable in the future. For discussion, see Wolfe (2010).
- 21 In this section I borrow a few points from Louden (2018a, 29–31).
- 22 See n. 1, previously.
- 23 In the *Odyssey*, Odysseus escapes from the Cyclops by seizing a “fire-point-hardened timber” and twirling it the giant’s eye socket until “the roots of his eye crackle” (Homer 1965, 147).
- 24 The good news is that there are a few signs of resistance within present-day philosophy to this development – for example, in the experimental ethics movement. The bad news is that these philosophers view themselves as anti-Kantian. For discussion, see Lütge et al. (2014).
- 25 For discussion and documentation of this claim, see Louden (2000).
- 26 Now I a fourfold vision see,
 And a fourfold vision is given to me;
 ’Tis fourfold in my supreme delight
 And threefold in soft Beulah’s light
 And twofold Always. May God us keep
 From single vision & Newton’s sleep!
 (Blake 1946, 209f.)
- 27 *Grazie* to Andrea Staiti for his invitation to present an earlier version of this chapter at the conference on “Transcendental Philosophy and Naturalism”, held in beautiful Parma, Italy, in May 2018. Later versions of the chapter were also presented as invited lectures at Roma Tre University in May 2018, at the 8th Multilateral Kant Colloquium, University of Catania, Sicily in October 2018, and at the University of Hawaii, Manoa in February 2019. I would also like to thank the *Journal of Transcendental Philosophy*’s two anonymous reviewers for their helpful suggestions and comments on an earlier version of the written text.

References

- Basedow, J.B. 1965. *Ausgewählte pädagogische Schriften*. Edited by A. Reble. Paderborn: Ferdinand Schöningh.
- Basedow, J.B. 1752. *Inusitata et optima honestioris iuventutis erudiendae methodus*. Kiel: publisher not identified.
- Basedow, J.B. 1774. *Das in Dessau errichtete Philanthropinum, eine Schule der Menschenfreundschaft un guter Erkenntnis für Lernende und junge Lehrer, Arme und Reiche*. Leipzig: Siegfried Lebrecht Crusius.
- Blake, W. 1946. *The Portable Blake*. Edited by Alfred Kazin. New York: The Viking Press.
- Brennan, G., Eriksson, L., Goodin, R.E., and Southwood, N. 2013. *Explaining Norms*. Oxford: Oxford University Press.

- De Caro, M., and Macarthur, D. (eds.). 2010. *Naturalism and Normativity*. New York: Columbia University Press.
- Fink, H. 2006. "Three Sorts of Naturalism." *European Journal of Philosophy* 14: 202–221.
- Fleischacker, S. 2018. "Review of E. Robinson and C. Suprenant (eds)." *Kant and the Scottish Enlightenment. Notre Dame Philosophical Reviews*, February 14. <https://ndpr.nd.edu/news/kant-and-the-scottish-enlightenment> (accessed on March 4, 2019).
- Graves, F.P. 1912. *Great Educators of Three Centuries: Their Work and Its Influence on Modern Education*. New York: Macmillan.
- Homer. 1965. *The Odyssey*. Trans. R. Lattimore. New York: Harper & Row.
- Hume, D. 1978. *A Treatise of Human Nature*. Edited by L.A. Selby-Bigge. Oxford: Clarendon Press.
- Hume, D. 2005. "An Enquiry Concerning Human Understanding." In L.A. Selby-Bigge (ed.), *Enquiries Concerning Human Understanding and the Principles of Morals*, 3rd ed., Oxford: Clarendon Press.
- Landschoof, J.A. 1933. "The Life and Work of Johann Bernhard Basedow." PhD diss. New York University.
- Louden, R. B. 2000. *Kant's Impure Ethics: From Rational Beings to Human Beings*. New York: Oxford University Press.
- Louden, R.B. 2003. "The Second Part of Morals." In B. Jacobs and P. Kain (eds.), *Essays on Kant's Anthropology*. Cambridge: Cambridge University Press, 60–84.
- Louden, R.B. 2011. *Kant's Human Being: Essays on His Theory of Human Nature*. New York: Oxford University Press.
- Louden, R.B. 2017a. "Becoming Human: Kant's Philosophy of Education and Human Nature." In M.C. Altman (ed.), *The Palgrave Kant Handbook*. London: Palgrave Macmillan, 705–727.
- Louden, R.B. 2017b. "A Writer More Excellent than Cicero: Hume's Influence on Kant's Anthropology." In E. Robinson and C. Suprenant (eds.), *Kant and the Scottish Enlightenment*. New York: Routledge, 164–180.
- Louden, R.B. 2018a. "Kant's Anthropology: (Mostly) Empirical Not Transcendental." In F.V. Tommasi (ed.), *Der Zyklus in der Wissenschaft: Kant und die anthropologia transcendentalis*. Hamburg: Felix Meiner, 19–33.
- Louden, R.B. 2018b. "The Moral Dimensions of Kant's Anthropology." In G. Lorini and R.B. Louden (eds.), *Knowledge, Morals, and Practice in Kant's Anthropology*. London: Palgrave Macmillan, 101–116.
- Louden, R.B. 2021a. "Humans-Only Norms: An Unexpected Kantian Story." In A. Lyssy and C.L. Yeomans (eds.), *Kant on Morality, Legality, and Humanity: Dimensions of Normativity*. London: Palgrave Macmillan, 131–147.
- Louden, R.B. 2021b. "'An Illusion of Affability that Inspires Love': Kant on the Value and Disvalue of Politeness." In X. Chaoqun (ed.), *The Philosophy of (Im) politeness*. Berlin: Springer, 213–224.
- Louden, R.B. 2021c. *Johann Bernhard Basedow and the Transformation of Modern Education: Educational Reform in the German Enlightenment*. London: Bloomsbury.
- Louden, R.B. Forthcoming. "'The Perfection of Humanity': Kant on Culture and Education." In T. Morawski (ed.), *Kant's Idea of Culture*. Rome: Sapienza Università Editrice.

- Lütge, C., Rusch, H., and Uhl, M. (eds.). 2014. *Experimental Ethics: Toward an Empirical Moral Philosophy*. London: Palgrave Macmillan.
- Mendelssohn, M. 1771–98. “Morgenstunden oder Vorlesungen über das Dasein Gottes.” In F. Bamberger et al. (eds.), *Gesammelte Schriften: Jubiläumsausgabe*, vol. 3, part 2, Stuttgart-Bad Canstatt: Frommann-Holzboog.
- Mozart, L. 1756. *Versuch einer gründlichen Violine-Schule*. Augsburg: published by the author.
- O’Neill, O. 1989. *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. Cambridge: Cambridge University Press.
- Parker, C. 1912. “Experimental Schools in Germany in the Eighteenth-Century.” *The Elementary School Teacher* 12(5): 215–224.
- Paulsen, F. 1908. *German Education: Past and Present*. Trans. T. Lorenz. New York: Charles Scribner’s Sons.
- Quick, R.H. 1896. *Essays on Educational Reformers*. New York: D. Appleton and Company.
- Rorty, R. 1982. “Keeping Philosophy Pure: An Essay on Wittgenstein.” In R. Rorty (ed.), *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press, 19–36.
- Rousseau, J. 1779. *Émile, or On Education*. Trans. A. Bloom. New York: Basic Books.
- Tommasi, F.V. (ed.). 2018. *Der Zyklop in der Wissenschaft: Kant und die anthropologia transcendentalis*. Hamburg: Felix Meiner.
- Weber, M. 1958. *The Protestant Ethic and the Spirit of Capitalism*. Trans. T. Parsons. New York: Charles Scribner’s Sons.
- Wolfe, C. 2010. *What Is Posthumanism?* Minneapolis: University of Minnesota Press.

PLEASURE AND DISPLEASURE AS MORAL MOTIVATION

Maria Borges

1. Are emotions and feelings necessary for moral action?

In the book *The Emotional Construction of Morals*, Jesse Prinz claims the most fundamental division in moral philosophy is between the philosophers who think that feelings and emotions are essential to morality and those who think they are not. Prinz stands among the first, and Kant is considered the most expressive member of the second group. In fact, Prinz supports a theory he calls *emotionism*, according to which morality is based on emotions. There are two versions of emotionism: metaphysical emotionism (Prinz 2007, 14), according to which moral properties are essentially related to emotions, and epistemic emotionism, in which moral concepts and judgments are essentially related to emotions. Emotivists will reject metaphysical emotionism, because they do not believe there is anything such as moral properties. However, they will accept epistemic emotionism, because they believe right and wrong are nothing more than expressions of emotions. For this reason, some emotivists are called expressivists.

A well-known expressivist, Simon Blackburn, in the book *Ruling Passions*, claims that ethics is about what we feel:

Ethics is a practical subject, manifested in our reactions to things and the motivation we feel. . . . If there is such a thing as ethical knowledge, it is matter of knowing how to act, when to withdraw, whom to admire, more than knowing that anything is the case.
(1998, 1)

Moreover, he claims our abilities for action would be damaged if we didn't have any emotion.¹

A Kantian would reject both forms of emotionism. Moral properties are not related to emotions but to the possibility of being universalized in a coherent way, without contradiction. Kant is also an opponent of epistemic emotionism. For him, moral judgments are not related to emotions but to a concept of the right action that could not be dependent on emotional states

of the subject. But could a Kantian really swear that feelings do not play any role in moral action?

2. The fairy tale of a motivation without inclinations

Kant is well known as a moral philosopher who claims that morality should be based on reason alone and not need the help of passions. He is usually understood as the opposite of the sentimentalists who base morality on passions or on moral sensibility.

However, many Kantian philosophers still feel the need to answer the Humean question: how can you explain moral action without any sensible incentive at all? Many are the Kant scholars who are still haunted by what I would call “the Humean ghost”. Iain Morison, in the introduction to the book *Kant and the Role of Pleasure in Moral Action*, also considers this problem in the following terms: “how can Kant account for moral motivation while divorcing the basis of morality from the pathological, and therefore motivational side of human agents?” (2008, 1). In the article “Kant and Motivational Externalism”, Karl Ameriks insists that

if a Kantian still wishes to reach a truly cosmopolitan audience, it makes sense to step back beyond the perspective of Kantian scholarship alone, and to reflect on the basic features that contemporary philosophers would insist that any acceptable theory treat with sensitivity.
(2006, 3)

But perhaps the most striking case of capitulation to the Humean ghost is an article by Paul Guyer, “A Passion for Reason”, where we can read in last paragraph: “I personally find Kant’s theory that we have noumenally free will that always allows us to do the right thing no matter what our inclinations a fairy tale” (\ 2016, 15).

The “fairy tale” Guyer refers to is called an *arbitrium sensitivum liberum* and is defined in the Critique of Pure Reason:

The human power of choice is indeed an *arbitrium sensitivum*, yet not *brutum* but *liberum*, because sensibility does not render its action necessary, but in the human being, there is a faculty of determining oneself from oneself, independent of necessitation by sensible impulses.

(KrV, A 534/B562)

But in order to act morally, should we always act in the absence of sensible motivation? In the *Groundwork*, the answer seems to be positive. Sympathy for other people’s circumstances, as a feeling that leads to beneficence, is analyzed in the well-known example of the *Groundwork*. When explaining

the difference between acting from duty and according to duty, Kant presents the example of two philanthropists, distinguishing the one who possess a strong pleasure in spreading joy to his fellow human beings from the one who helps other people out of duty. The philanthropist who acts from duty is the one who is not incited by any inclination (GMS, 4: 398). Kant asks if we should consider his action would have a higher worth if nature had put little sympathy in his heart, and the answer is negative: “By all means! It is just then that the worth of character comes out, which is moral and incomparably the highest, namely, that he is beneficent not from inclination but from duty” (GMS 4: 399).

The difference between the two philanthropists is that the incentive of the compassionate one is sympathy, which is a sensible inclination, while the action of second philanthropist is performed from respect for the moral law. Kant considers that, if compassion for other people is the incentive of an action, then this action does not have a true moral value.

However, according to many contemporary commentators, the mere presence of some feelings, such as sympathy, does not make an action morally unworthy if respect for moral law is a sufficient incentive for the accomplishment of the action. In this sense, they follow Henson (1979), one of the first who claimed that a morally worthy action does not need the absence of inclinations. In what he calls a *fitness report model* interpretation, there is no need to banish all other inclinations toward an action if respect would have been a sufficient reason to cause the action. Then, the moral action does not require an absence of inclination, provided that respect for duty is present and would suffice to produce the dutiful action.

But even if, in the last years, there has been a tendency to accept the broad interpretation of Henson, if you go back to the texts of the *Groundwork* and *Critique of Practical Reason*, it is easy to see that the necessity to support morality in a non-material practical principle leads to the refusal to ascribe the role of moral incentives to feelings such as love, benevolence and affection. The *Groundwork*, as well as the *Critique of Practical Reason*, have the aim of demonstrating that reason can determine the will without the help of empirical incentives. In this context, Kant refuses to ascribe the role of an incentive to benevolent feelings, since these would be empirical and contingent, not being able to be taken as a ground for the determination of the will. In these texts, Kant states clearly that a feeling is a subjective incentive, inappropriate for the establishment of morality and its foundation upon reason.

3. Respect and sensibility

Even though sympathy should not be considered an incentive for moral actions, there is one feeling that is always present in the performance of a moral action: the feeling of respect. Is Kant presenting us with the solution for the sensible motivation for moral action, offering a sensible incentive to it?

Is respect a feeling? Or is it only the awareness of moral law? Kant is ambiguous about that. In the *Groundwork*, he states that “the determination of the will by the law and the consciousness of this determination is respect” (GMS 4: 401 note), apparently encouraging the interpretation that respect is not a feeling but only the recognition of duty. However, in the same text, he also states that respect is a feeling (*Gefühl*), although not created by an external object:

But though respect is a feeling, it is not one received by means of influence; it is, instead, a feeling self-wrought by mean of a rational concept and therefore specifically different from all feelings of the first kind, which can be reduced to inclination or fear.

(GMS 4: 401n)

What exactly is this feeling that is different from all feelings? Is it a feeling of pleasure and pain? Or it is only the consciousness of moral law? If it is a feeling, may it motivate a moral action? Iain Morrison has argued that despite respect being a feeling, it can motivate moral action:

it is possible to read Kant as saying that respect is a complex feeling – made up of feelings of pleasure and pain – that is somehow produced or caused by (and therefore, distinct from) the moral law. On this view, respect is not identified with the moral law. Instead, it is a feeling caused by the moral law, and it motivates insofar as it is made up of a combination of feelings of pleasure and pain.

(2008, 5)

Although I recognize, as Morrison does, that respect is a complex feeling and not only the conscience of moral law, I argue that this feeling is not what motivates us, because no feeling should motivate us in a pure moral action. And that is what made Kant different from his sentimental predecessors. Morrison considers that only the cognitive dimension of respect motivates moral action. For me, what Morrison calls the cognitive dimension of respect is nothing more than the thought of moral law and not the feeling of respect itself. Then, I am inclined to maintain that respect is a feeling caused by moral law and is not the incentive to act morally, but it is only a sensible effect of the acceptance of moral law.

Frierson, in the book *Kant’s Empirical Psychology*, calls attention to the debate between “intellectualists” and “affectionists”:

Intellectualists (Allison, Guyer,² Reath) claim that morally good action is motivated solely by cognition or consciousness of moral law, with a feeling (of respect) generally seen as an effect of moral motivation rather than its cause. Affectionists (McCarthy, Singleton,

Hererra, Morrisson, Nauckhoff) argue that the feeling is the immediate cause of moral motivation, the means by which an otherwise inert cognition of the moral law can give rise to an action.

(2014, 117)

According to Frierson, the intellectualist-affectionist debate presents textual evidence for both sides. Janelle DeWitt, in the article “Respect for Moral Law: The Emotional Side of Reason”, claims that respect has a duality and seems to embody a contradiction, since it is both a moral motive and a feeling, and these are thought to be mutually exclusive. She also shows that there are two dominant views in regard to which of the two states they identify as respect: “Andrews Reath is one who argues that respect is consciousness of moral law and so the moral motive, whereas Paul Guyer identifies respect with the pathological aftermath, and so views it as a feeling” (De Witt 2014, 32).

We could consider that there are three positions: the intellectualists, the affectionists and the mixed position (Iain Morrison, Janelle De Witt). De Witt, for instance, states that emotions take the form of evaluative judgments and respect is a purely rational emotion, mixing the sensible and rational realms of motivation.

I maintain that the stronger evidence is on the side of the intellectualists. In the *Groundwork*, Kant claims that “immediate determination of the will by means of the law and consciousness of this is called respect, so that this is regarded as the *effect* of the law on the subject and not as the *cause* of the law” (GMS 4: 402). Also, in chapter III of *the Critique of Practical Reason*, Kant claims that:

If the determination of the will takes place *conformably* with the moral law but only by means of a feeling, of whatever kind, that has to be presupposed in order for the law to become a sufficient determining ground of the will, so that the action is not done *for the sake of the law*, then the action will contain *legality* indeed, but not *morality*.

(KpV 5: 71)

In this sense, I agree with some commentators, such as Reath, for whom the feeling of respect cannot be a motivation for the moral law. To claim that respect is a feeling does not entail that this feeling should have the role of an incentive or motive for a moral action. Also, the fact that respect is a feeling self-wrought by a rational concept does not imply that respect lacks sensible properties. In the *Anthropology*, Kant claims that feelings of pleasure and displeasure can be an effect of an idea or concept (Anth 7:230). Then, an intellectual feeling, as he calls it, does not mean that a feeling is itself a concept, lacking sensible properties, or that emotions take the form of evaluative judgments, as is claimed by DeWitt. It means that it is a feeling (*Gefuhl*) and has sensible properties but was caused by an idea. A misunderstanding

of the sensible property of something that can be caused by an idea is the source of some Kant scholars' mistakes.

Kelly Sorenson, in his article "Kant's Taxonomy of Emotions" (2002), explains the nature of respect when he analyses the relation between desires and feelings;

Desire necessarily involves feelings, but desire is of two sorts, depending on whether the pleasure associated with it is the *cause* of the desire or instead its *effect* (KU, 5: 221–2; MS 6: 212). Kant calls *pleasure-caused desire* "desire [*Begierde*] in the narrow sense" (MS 6:212). In this case, an agent seeks to bring about the existence of some object or state of affairs because of some antecedent pleasure. When these desires are habitual, Kant gives them their own term: inclinations. In the case of the other sort of desire, pleasure is the effect of the desire. Here it is *reason* that causes the desire, which in turn results in pleasure. As early as the *Groundwork*, Kant recognizes the existence of at least one such *reason-caused desire*: he calls the feeling necessarily connected with it "respect" or "moral feeling".

(2002, 114)

The difference between object-caused desire and reason-caused desire can explain how pleasure and displeasure could be an important part of morality without being the motive or incentive of moral action.

4. Moral feeling

According to Sorenson, what is said about respect also refers to moral feeling. In the *Critique of Judgment* (KU 5: 289), Kant claims that when an *a priori* principle determines the will, there is pleasure associated with that, the moral feeling, that is the consequence of that determination. This is not in contradiction with what is stated in the *Groundwork*, where respect is said to be "a feeling self-wrought by means of a rational concept" (GMS 4:401n).

In the introduction to *Doctrine of Virtue* (MSTL 6:399), Kant explains that there are aesthetic pre-notions of the susceptibility of the mind (*Gemüth*) relative to respect (*Ästhetische Vorbegriffe der Empfänglichkeit des Gemüts Achtung*). These are moral feeling, conscience, love of human beings and self-respect. They regard not the obtainment *a priori* of moral law or its conditions but the analysis of the moral agent, the human being and his natural predispositions, which facilitates the effective reception of the moral imperative. Moral feeling is defined as "the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent or contrary to the law of duty". (MSTL 6: 399), stressing the importance of sensibility for the conscience of duty. It is this feeling which gives us conscience of obligation, because it makes us aware of the coercion present in

the mere thought of duty. We do not have any duty to feel such an emotion, since it is in us as moral beings. Someone completely deprived of this feeling would be morally dead:

No human being is entirely without moral feeling, for was he completely lacking in receptivity to it he would be morally dead; and if (to speak in medical terms) the moral vital force could no longer excite this feeling, then humanity would dissolve (by chemical laws, as it were) into mere animality and be mixed irretrievably with the mass of other natural beings.

(MSTL 6:400)

This ambiguous feeling can be pathological or moral: pathological if it “precedes the representation of the law”, moral if it “can only follow upon it”. (MSTL 6:399). Moral feeling seems to be a product of the representation of moral law; consequently, it is not an incentive to act morally. However, Kant is ambiguous when he claims that moral feeling is not a sense for the morally good but a susceptibility on the part of free choice to be moved by pure practical reason. When he asserts that someone who lacks moral feeling would be morally dead, it is implicit that one cannot act morally without it, although it does not imply that moral feelings necessarily precede the moral action and act as an incentive.

Moral feeling is not, however, moral sense (*Sinn*) understood as a feeling that, by itself, indicates what is correct and what is not. In other words, moral feeling does not have any theoretical ability to directly perceive the correctness of an action, since this must be given to us by reason. It simply refers to the internal perception of the coercion duty exerts, provoking a feeling of pleasure when our actions are according to the law and of displeasure when they are against it. Moral feeling, however, while a feeling (*Gefühl*), is not an affect, since the latter is uncontrolled and hinders reflection. In this sense, Kant clarifies that the apathy he proclaims is not the absence of all feelings, since this would be moral indifference and thereby a weakness. Apathy sought as an ideal is absence of affect (*Affekt*) and, consequently, a beneficial strength in the practice of virtue.

The clear separation proclaimed in the *Groundwork* and the *Critique of Practical Reason* between the realms of morality and sensibility – understood as the feeling of pleasure and displeasure – is questioned in some passages of *Critique of Judgment* and *Metaphysics of Morals*, which talk about the possibility of, and even the need for, moral feeling. One of the greatest differences between these two approaches comes from the necessity of feeling pleasure and displeasure to receive the concept of duty. Without a doubt, moral feeling goes beyond the feeling of respect. This becomes clear when moral feeling is considered necessary for the mind to be affected by the feeling of respect. And the possibility of the mind being empirically affected is decisive for the

effective realization of moral action. For the action to take place, it is necessary that the mind be affected, which is done through moral feeling.

The idea that moral feeling is important for moral action can be also found in the *Anthropology Friedlander*: “Although, as it is, the concepts of good and evil are not object of feeling, yet they can still serve to rouse feeling to act in accordance with these concepts; then one acts in accordance with principles and maxims” (V-Anth/Fried 25: 650).

According to these early lectures from the years 1775–1776, the concept of good will awakens a feeling to act in accordance with this concept, and this feeling is what Kant refers to as moral feeling:

Thus, concepts are to become incentives in us, they are to rouse feeling, and to motivate us to act in accordance with such concepts, and thus according to principles. Human beings who do not have such a feeling which can be roused through concepts have no moral feeling.
(V-Anth/Fried 25: 650)

This position is close to the claim of the *Metaphysics of Morals* that without moral feeling, we would be morally dead. However, in the *Anthropology Mrongovius*, moral feeling is said to be misleading: “In our actions we must not orient ourselves according to sensations, for these give no determinate rule and always mislead us. So too, with moral feeling” (V-Anth/Mron 25: 1321).

The lectures of the winter semester of 1784–85, based on the transcription of Mrongovius, are compatible with the denial of feelings as incentives for actions in the *Groundwork*, which was published the same year these lectures were given (1785). In the *Critique of Practical Reason* (1798), Kant also claims that moral sensibility is not a good guiding light for a moral action, although the determination of the will by the moral law may produce feeling of pleasure:

I certainly do not deny that frequent practice in conformity with this determining ground can finally produce subjectively a feeling of satisfaction with oneself; on the contrary, to establish and to cultivate this feeling, which alone deserves to be called moral feeling strictly speaking, itself belongs to duty; but the concept of duty cannot be derived from it – otherwise we should have to think of a feeling of a law as such and make what can only be thought by reason an object of sensation.

(KpV 5: 38)

Here Kant rejects all that can lead to a “mechanical play of refined inclinations” (KpV 5: 39), as he calls this idea of a moral life based on moral emotions. If one compares the *Anthropology* lectures of 1772–73 with the lectures of 1784–85, there is a difference between the moral relevance of

moral feeling. In the first, it is considered an important feeling that prompts us to act according to the concept of good, while in the second, it is a sensible feeling that can mislead us. This idea that moral feelings are motives for moral action disappears at the time of the *Groundwork* (1785), and sensibility works as a deceiving and misleading guide. In the *Critique of Practical Reason* (1788), though there is a refusal of sensibility as the source of duty, it is considered that the cultivation of moral feeling belongs to duty, as a pleasure that goes with the determination of the will by the law. In the *Doctrine of Virtue* (1797), moral feeling is necessary for a moral life, because, without it, we will be morally dead.

The discussion of moral feeling seems to be much more complex than the feeling of respect. One of the reasons is that moral feeling is a regard for pleasure that is absent regarding respect. As Guyer highlights (1997, 145), the feeling of respect in *Groundwork* is a feeling of displeasure proportional to the coercion of the law or, at most, a feeling of self-approval, which radically differs from the feeling of pleasure in the third *Critique* and the *Doctrine of Virtue*. A result of the determination of will, which opts for the moral action, such a feeling seems to be one of reward.

That morality can give us a sort of satisfaction and, moreover, that were we to lack this capacity – *Doctrine of Virtue* tells us – we would be morally dead, seems to be very far from respect as a feeling of displeasure. Though some commentators, such as Morrison, think respect is the incentive to moral action, my position, as explained earlier, is that respect is a consequence of the determination of the will by the moral law, not a motive itself. Could one claim the same applies to moral feeling, that it is not a motive or incentive but only a consequence of the moral determination of the will?

I assume that the scenario is much more complicated here. Moral feeling seems to work as a moral sensibility for the reception of the concept of duty. On the one hand, it is not properly an incentive for moral action; on the other hand, morality cannot exist without it. In the *Doctrine of Virtue*, moral feeling is defined as “a susceptibility on the part of free will to be moved by our practical reason (and its law)” (MSTL 6:400). It is a necessary sensible condition, even if it is not an incentive.

Perhaps it is the sensible element that prevents Kantian morality from being only a fairy tale, without falling into a mechanical play of refined inclinations as expressivists would do.

Notes

- 1 Blackburn, in the book *Ruling Passions*, refers to the work of Antônio Damasio, who has allegedly shown that people with injuries in the emotional part of the brain will not be able to act properly and choose a course of action. See Blackburn 1998, 125f.
- 2 I consider Guyer an affectionist and that Frierson is mistaken in taking him as an intellectualist.

References

I use the translations of the Cambridge Edition of The Works of Immanuel Kant.

- Ameriks, K. 2006. "Kant and Motivation Externalism." In H. Klemme, M. Kuhn, and B.D. Schoneker (eds.), *Moralische Motivation Kant und die Alternativen*. Kant-Forschungen 16. Hamburg: F. Meiner.
- Blackburn, S. 1998. *Ruling Passions*. New York: Oxford University Press.
- De Witt, J. 2014. "Respect for Moral Law: The Emotional Side of Reason." *Philosophy* 89: 31–62.
- Frierson, P. 2014. *Kant's Empirical Psychology*. Cambridge: Cambridge University Press.
- Guyer, P. 1997. *Kant and the Claims of Taste*. Cambridge: Cambridge University Press.
- Guyer, P. 2016. "A Passion for Reason." In P. Guyer (ed.), *Virtues of Freedom*. Oxford: Oxford University Press.
- Henson, R. 1979. "What Kant Might Have Said: Moral Worth and the Overdetermination of a Dutiful Action." *Philosophical Review* 88: 39–54.
- Morrison, I. 2008. *Kant and the Role of Pleasure in Moral Action*. Athens: Ohio University Press.
- Prinz, J. 2007. *The Emotional Construction of Morals*. New York: Oxford University Press.
- Sorensen, K.D. 2002. "Kant's Taxonomy of Emotions." *Kantian Review* 6: 109–128.

INEFFICACY, DESPAIR, AND DIFFERENCE-MAKING

A secular application of Kant's moral
argument

Andrew Chignell

It does not matter how many doubts from history may be raised against my hopes, which, if these doubts were *proved*, could move me to desist from a task so apparently futile; *as long as these doubts cannot be made quite certain* I cannot exchange the duty for the rule of prudence not to attempt the undoable.
(TP 8: 309)

1. Inefficacy and opportunism

Let's start by assuming that there is a sound argument for the following:

Don't Produce: It is morally wrong to participate voluntarily in Production Activity A.

"Production Activity A" is just a placeholder; the reader is invited to fill in the blank. By "participate voluntarily", I mean something like "direct or engage in, when not under compulsion or duress, typically with the aim of benefiting thereby". An example of an A from food ethics might be "the raising and harvesting of animals for the purpose of selling their bodies, secretions, and eggs".¹ Other A's might be some of the production activities that occur in sweatshops, cobalt mines, coffee farms, chocolate farms, avocado orchards, cellphone factories, tanneries, diamond mines and so on.²

Now consider an average consumer of middling means – let's call him Oppy. Oppy likes meat, cheese, chocolate and sweatpants and has a special weakness for leather shoes. He likes to save money by buying cheap sweaters and non-fair trade coffee. He also finds iPhones irresistible, especially if they come with a diamond-encrusted case (again: fill in the blank). So Oppy

would dearly like to purchase these products of our contemporary global supply chains. But he also believes that some of the activities involved in producing these products are morally objectionable; in other words, he believes that *Don't Produce* holds with respect to them. And he just assumes, like many of us do, that *Don't Produce* supports

Don't Purchase: It is morally wrong to purchase the products of Production Activity A if alternatives are readily available.³

And so, up to now, Oppy has resisted the urge to purchase.

Recently, however, Oppy has confronted some sobering empirical facts that seem to challenge the move from *Don't Produce* to *Don't Purchase* with respect to many production activities in the contemporary global context. The sobering empirical facts, roughly put, are that the sheer size and complexity of the industrial system that brings 99% of products to our refrigerators, pantries and closets make it *exceedingly* unlikely that an individual's choice not to purchase those products will have any effect at all on the system. This is particularly true of "lumpy" supply chains like those that supply us with industrial animal products. These systems are lumpy because the relevant goods (e.g. chicken products) are ordered in large lots all the way down the supply chain – from the restaurant to the various suppliers to the farmer to the incubator – rather than one by one. So even if a chicken sandwich from a fast food restaurant contains (let's say) 1/20th of a chicken, the sobering facts about the supply system give Oppy a *far* lower than 1-in-20 (i.e. 5%) chance that his order at a fast food counter will make a difference to any particular chicken. If we take account of seasonal variables, buffers, waste and other supply-chain "noise",⁴ then the chance that he will buy a "threshold" chicken sandwich becomes, in Mark Budolfson's words, "infinitesimal" (2016, 208).

Put another way: calculating the expected effect for each customer who orders a chicken sandwich would be difficult, given all these factors, but it is clearly *nowhere near* as high as 1/20th of a chicken.⁵ So the sobering empirical facts suggest that the standard expected utility response in these contexts – that is, divide through and assign responsibility for the *average* effect to each individual consumer – is flawed.⁶ Or that's what Oppy concludes, in any case. After reflecting on all of this, he comes to believe that his consumption patterns with respect to many industrial products almost certainly won't make a difference. And because he really likes his sausages and sweaters and cell phones, he now feels free to go ahead and purchase the next time he is at the drive-through. Note that this is not a case of weakness of will: Oppy remains fully committed to *Don't Produce*. He has simply given up the assumption that *Don't Produce* supports *Don't Purchase*. Oppy has become an opportunist.⁷

A number of philosophers (Martin 1976; Frey 1985; Budolfson 2018; Michaelson 2016a; Kingston 2020) have shown how such concerns about

inefficacy are pressing for act consequentialist advocates of the move from *Don't Produce* to *Don't Purchase*. This is because act consequentialism (hereafter “consequentialism”) endorses the following principle:

Causal Consequences Principle (CCP): Action A is morally wrong only if the causal consequences of A are worse than they would have been if some readily available alternative to A had been performed.⁸

“Worse” here is typically construed in a broadly *welfarist* way, where what matters is the well-being of sentient animals.⁹ So, given CCP, if a private¹⁰ purchase of a chicken sandwich that is wrongfully produced almost certainly *doesn't* cause any additional loss in welfare, and certainly *does* lead to the goods of nourishment and gustatory pleasure, then it cannot be morally wrong. It might even be morally required. This then sets up a trilemma for the consequentialist: give up CCP, give up *welfarism*, or give up *Don't Purchase* and get comfortable with opportunism. The whole problematic is now referred to as the *inefficacy* (or *futility* or *causal impotence*) objection to consequentialism.

Clearly it is an empirical question whether the sobering empirical facts described previously really obtain. Real-world industrial supply chains for non-perishables, for instance, are more likely to be vertically integrated than those for perishables.¹¹ In order to set aside this question, we can put the problem conditionally: *if* there are such sobering empirical facts about some of our industrial supply chains such that an individual purchase of their products is extremely unlikely to make a *causal* difference to overall welfare, then the intuitive-seeming shift from *Don't Produce* to *Don't Purchase* is blocked.¹²

A few philosophers (Michaelson 2016b; Budolfson draft) have argued that these concerns about futility also afflict some *non-consequentialist* efforts to move from *Don't Produce* to *Don't Purchase*. For example, they make things difficult for views on which *only* acts that involve or lead to the infringement or violation of rights count as morally wrong. That's because the sobering empirical facts indicate that an individual's choice to purchase the products of a gigantic, insensitive system is extremely unlikely to infringe or violate any rights. And most non-consequentialists of this variety will allow that it is permissible to perform actions that are extremely unlikely to infringe or violate rights.

Other varieties of non-consequentialism (certain forms of virtue theory, symbolic or expressive value theory, divine command theory and even Parfit's neo-Kantian theory)¹³ are not vulnerable to this form of the inefficacy objection, however. The decision to purchase the products of a morally wrongful activity can still be vicious, symbolically unacceptable, divinely proscribed or categorically irrational – and thus impermissible – even if it has no negative consequences for welfare.¹⁴

Oppy, for his part, finds the latter sorts of theories mysterious and/or implausible: he's a committed consequentialist. Thus, despite his unwavering

commitment to *Don't Produce*, he now feels comfortable being an opportunist. The principle underwriting Oppy's reasoning is something like this:

Badness as a Reason to Abstain: The badness of a production activity A is a reason to abstain from purchasing A's products *only if* abstaining has a non-negligible chance of making a positive difference with respect to the badness of A.

The more general idea here is:

Badness as a Reason to Act: The badness of some situation is a reason to perform an action *only if* that action has a non-negligible chance of making a positive difference with respect to the badness of that situation.

Oppy is also assuming, in keeping with CCP, that what it is to "make a difference" should be construed in a *causal* way:

Difference-Making, Causal: An action A makes a difference with respect to situation S if and only if A causes a change in S.

It would be worth thinking more about whether consequentialist theories can rebut the conceptual version of the inefficacy objection (and rule out opportunistic purchasing) without giving up one or more of their core principles. But I don't propose to go further into that debate here.¹⁵ In this chapter, I will just assume without further argument that Oppy is right: the inference from *Don't Produce* to *Don't Purchase* is hard to motivate within a consequentialist framework, given the sobering empirical facts.

In what follows, I propose to turn our attention away from Oppy in order to focus on another character: Hope. As we will see, Hope likes industrial products just as much as Oppy does. But Hope takes herself to have strong *independent* reasons for thinking that *Don't Purchase* is true. These reasons may include an appeal to one of the non-consequentialist theories just mentioned and in any case do *not* hang on whether her actions make a causal difference to welfare outcomes. All the same, Hope often – and quite understandably – finds her efforts to adhere to *Don't Purchase* threatened by her awareness of the sobering empirical facts that converted Oppy to opportunism. And she's not alone in this: for many of us, the recognition that our abstention almost certainly makes no causal difference *vis-à-vis* the morally objectionable system we're boycotting poses a serious *psychological* threat to our ongoing moral resolve. This is so even if we reject *Badness as a Reason to Abstain* – that is, even if our reasons for abstaining do not depend on our actions making a causal difference.

Here is another way to illustrate the difference between Oppy and Hope. The usual version of the inefficacy objection is *moral-conceptual*:

it challenges the intuitive-seeming assumption that *Don't Purchase* follows from *Don't Produce* by showing that consequentialist theories imply the opposite in circumstances where individual choices are almost certainly inefficacious. The version of the problem I consider here, by contrast, is *moral-psychological*: it starts with the conviction that *Don't Purchase* is true for non-consequentialist reasons. It then points out that many people (like Hope) who have this view but are still inclined to purchase will often be psychologically demoralized under conditions of *perceived* inefficacy such that they are unable to resist.¹⁶ So Hope's case, unlike Oppy's, does involve a kind of weakness of will.

Although he lived well before the advent of factory farms and gigantic supply chains, Immanuel Kant had something illuminating to say about Hope's situation. Or at least that's what I will suggest in the following. In Section 2, I set up the problem in more detail before going on in Section 3 to reconstruct one version of Kant's famous "moral proof" of God's existence and the future life of the soul. This is the version he developed in the 1790s towards the end of his career, just as his hopes regarding the French Revolution were giving way to despair. I call this the "moral-psychological" version of the proof in order to highlight the fact that – unlike the better-known version in the second *Critique* – it relies on empirical premises about our tendencies to hope and despair, and a putative need to trust in God or some other supersensible mechanism to shore up our moral-psychological resolve. As we will see, the conclusion of the moral-psychological proof is not that God and the future life exist but rather that we are morally justified in having Belief (the German term is "*Glaube*")¹⁷ that God and the future life exist. After making this detour through Kant's theistic moral psychology, my ultimate goal (in Section 4) is to see whether there is a secular analogue of the proof that applies to people in contemporary industrial contexts who, like Hope, are demoralized by our apparent inefficacy. Is there something analogous to trust in God that can shore up moral resolve in such circumstances, but without involving a supersensible commitment?

2. Demoralization and despair

We stipulated previously that Hope already takes herself to have an argument that allows her to bridge the gap from *Don't Produce* to *Don't Purchase*, without appealing to causal difference-making. One family of such arguments invokes the symbolic value of "standing with the good" and avoiding "moral taint" by not consuming the products of a morally objectionable practice (see Hill 1979; Appiah 1986; Cuneo 2016; Boey 2016). Another invokes an account of "complicity" or a "don't benefit from wrongness" principle that prohibits purchasing even if it makes no causal difference (Martin 2016; McPherson 2016). Another family appeals to virtue-theoretic or psychological considerations that don't require any

causal difference-making (Nobis 2002; Halteman and Halteman Zwart 2016; Bramble 2015). Yet another emphasizes that an individual can be a *joint cause* of bad outcomes even without making a causal difference and that the badness of being a joint cause can be a reason to abstain (Parfit 1984; Harman 2016; Albertzart 2019). Less common among applied ethicists than in the real world is the family of religious arguments: some people endorse *Don't Purchase* because they think it is against God's will or out of keeping with God's plan for us to purchase the products of morally wrongful systems.

As we saw previously, however, even if Hope accepts *Don't Purchase* on one of these other bases, or simply as an independent moral intuition, the sobering empirical facts remain: her food choices with respect to the industrial system are unlikely to make a significant difference. This has a profound psychological effect on Hope's moral resolve. In other words, her awareness of the sobering facts that led Oppy to opportunism leads Hope to a very natural kind of *demoralization*:

(D1) General discouragement and dejection in the face of the very long odds of making a significant positive difference with respect to the outcomes that one is morally concerned about.

This in turn produces another and more serious kind of *de-moralization*:

(D2) Loss of stable psychological resolve required to do what the agent *still* takes herself to have moral reason to do.

Call the state of being doubly demoralized in this way *despair*. When Hope falls into despair, she is liable to give up her boycott and revert to purchasing whatever is most convenient, tasty and affordable – despite her ongoing moral qualms.¹⁸

Kant was acutely aware of our psychological vulnerability to despair in the face of large-scale evil and injustice. In a sermonic passage in third *Critique*, he considers the psychology of an atheistic but “righteous man (like Spinoza)” who

does not demand any advantage for himself from his conformity to the moral law, whether in this world or another; rather, he would simply and unselfishly bring about the good to which that holy law directs all his powers.

(KU, AA 05:452)

Such a person's resolve, Kant says, is still liable to be sapped by the perception that his own moral strivings, as well as those of others, are inefficacious against large-scale structural evil and injustice:

But his strivings (*Bestreben*) have limits. . . . Deceit, violence, and envy always surround him, even though he is himself honest, peaceable, and benevolent. The other righteous people that he encounters at times will, in spite of all their worthiness to be happy, nevertheless be subject by nature, which pays no respect to that, to all the evils (*Übeln*) of poverty, illnesses, and untimely death, just like all the other animals on earth. It will always remain so until one wide grave engulfs them all together (whether honest or dishonest, here it makes no difference) and hurls them, the very ones who were capable of believing that they were the final purpose (*Endzweck*) of all creation, back into the abyss of the purposeless chaos of matter (*Schlund des zwecklosen Chaos der Materie*) from which they all were drawn.

(*ibid.*)

In the end, when faced with the abyss in this way, the righteous Spinoza has two options: either he will “certainly have to give up his end [i.e. of being righteous] as impossible” and collapse in despair, or “he will have to accept (*annehmen*) the existence of a moral author of the world (*Welturheber*), i.e. of God, from a practical point of view”. “Acceptance” is a synonym for “Belief”, and I read “impossible” here psychologically: Kant is saying that even someone as righteous as Spinoza simply cannot maintain moral resolve without some ability to trust that, ultimately, justice will prevail.¹⁹

These Kantian reflections on demoralization suggest that, in addition to finding out the sobering empirical facts about industrial systems, there is another empirical question that we need to consider in connection with inefficacy. It is an empirical-psychological question about which sorts of background commitments or states of trust do better by way of helping ordinary agents avoid despair in the presence of perceived inefficacy. Kant thought that someone who has Belief in the existence of a benevolent deity, and who trusts that this being is working behind the scenes to bend the arc of history towards justice (perhaps also recording and rewarding our efforts), will often have more resources to sustain their moral resolve, despite short-term setbacks and perceived inefficacy and futility. In the next section, I propose to look more closely at the structure of this moral-psychological version of Kant’s proof. After that, I will consider whether analogous forms of reasoning, and analogous kinds of trust, can help with the psychological version of the inefficacy objection in secular consumer contexts.

3. A moral argument for trust in God, or mere wishful thinking?

Kant’s project in *Religion Within the Bounds of Reason Alone* (1793) is to discover the rational essence of religion – the doctrines and practices that

can survive examination by the tribunal of universal rationality. The result is “moral religion”, “rational religion” or “rational Belief” (*Vernunftglaube*) – terms he uses to distinguish it from creedal or enthusiastic forms based in special revelation or alleged mystical experience. At one point, Kant offers the following as the rational essence of the doctrine of Providence:

Each must, on the contrary, so conduct himself as if everything depended on him. Only on this condition *may he hope* that a higher wisdom will provide the fulfilment of his well-intentioned effort.

(RGV, AA 06:101, my emphasis)

The passage displays what might be called a *consequence-dependent moral psychology*. Kant is famously not a consequentialist in ethics generally: “The fulfillment of duty consists in the form of the earnest will, not in the mediating causes of success” (KU, AA 05:451). Nor is he a divine command theorist: “Morality . . . needs neither the idea of another being above him in order to recognize his duty nor as an incentive anything other than the law itself in order to observe it” (RGV, AA 06:3). But Kant also recognizes that *after* we have willed to act on the moral law, it is often psychologically crucial to be able to *hope* that our ethical and political efforts will be fulfilled.

So Kant’s “rational” doctrine of Providence is sensitive, on the one hand, to the demand that we continue to be engaged, rather than waiting for God to do everything (“so conduct himself as if everything depended on him”). But it is also sensitive, on the other, to the empirical psychological fact that it is hard for us to keep soldiering on in the moral life without the hope that our efforts will have an effect (“the fulfillment of his well-intentioned effort”). This hope, in turn, motivates Belief and trust in a supersensible being (“a higher wisdom”). Here is another key passage:

Belief (Glaube) (simply so-called) is trust in the attainment of an aim the promotion of which is a duty but the possibility of the realization of which is not possible for us to have insight into. . . . The Belief, therefore, which is related to particular objects that are not objects of possible knowledge or opinion (in which case, especially if historical, it would have to be called credulity (Leichtgläubigkeit) and not Belief) is entirely moral. It is a free affirmation, not one for which dogmatic proofs for the theoretically determining power of judgment are to be found, nor one to which we hold ourselves to be obligated, but one which we assume for the sake of an aim in accordance with the laws of freedom; yet not like an opinion, without a sufficient ground, but as adequately grounded in reason (although only in regard to its practical use) for that aim; for without [this Belief] the moral way of thinking has no way to persevere in its collision with theoretical reason’s demand for a proof (of the

possibility of the object of morality), but vacillates between practical commands and theoretical doubts.

(KU, AA 05:472, original bold, my italics)

Kant recognizes, in other words, that prolonged experience of injustice – of a world-history in which the wicked prosper and virtue seems to make no difference – can chip away at our moral resolve. He also thinks that this fact can, in conditions of epistemic ambiguity, morally justify Belief and trust in Providence.²⁰

A concrete example: suppose that practical reason demands that we try to help the disadvantaged by donating money to various causes; for Kant, this will be true *regardless* of actual outcomes, and the fact that the moral law requires it is a sufficient rational motive. Kant sees that we (or many of us) *also* often have a “need” for our altruism to be effective:²¹ it helps us stabilize *not* our rational motive (which is based in the moral law) but our *psychological resolve* if we can believe or at least reasonably hope that the needy will genuinely benefit from our charitable efforts. Even in a Kantian context, then, a morally good person might well care a lot about the goodness of the *consequences* of her willing and not just about the goodness of her will.

This is a complicated piece of reasoning and a controversial reading of Kant’s argument. It is worth looking at it in an exhaustive step-by-step fashion:²²

- 1 I ought to do what is morally right. [Independent argument]
- 2 For me, it would be *demoralizing* in the (D1) sense (i.e. it would lead to discouragement and dejection) not to be able to have substantial hope that there is a moral order by which a just arrangement (i.e. a “moral world”) will come about, for then I would have to regard it as *certain* that the entire history of the world will not be good on the whole, no matter what I do. [Empirical premise, concept of “hope”]
- 3 Such demoralization has an enervating effect on my resolve, and is thus *de-moralizing* in the second, (D2) sense: I will almost certainly no longer perform actions that I take to be morally good or required. [Empirical premise]
- 4 Despair of this sort is seriously morally undesirable. [From (1)–(3)]
- 5 Therefore, there is serious moral advantage for me in being able to have substantial hope that there is a moral world order. [From (2)–(4)]
- 6 Substantial hope that *p* requires the positive belief or Belief that *p* is really practically possible. [Theoretical premise]
- 7 Therefore, there is serious moral advantage for me in being able to positively believe or Believe that a moral world order is really practically possible. [from (5)–(6)]

- 8 If a being or state of affairs is really practically possible, then there must be something in the actual world that can account for that fact. [Concept of “real practical possibility”]
- 9 The *actual* existence of God provides the only adequate account of the real practical possibility of a moral world order, because God is the only being we could trust to bring such an order about. [Theoretical premise]
- 10 Therefore, there is serious moral advantage, for me at least, in being able to believe or Believe that God exists. [from (7)–(9)]
- 11 There can be no sufficient epistemic reasons either for or against the existence of God. [Results of examination of natural theology in Transcendental Dialectic]
- 12 Rational belief requires sufficient epistemic reasons. [Kant’s evidentialism about belief]²³
- 13 Therefore, belief in God’s existence or non-existence is irrational. [from (11)–(12)]
- 14 Rational *Belief* (*Vernunftglaube*) does not require sufficient epistemic reasons; it can instead be based on sufficient moral or pragmatic reasons. [Conceptual truth]²⁴
- 15 Therefore, Belief (though not belief) that God exists is *prima facie* morally (though not epistemically) justified, for me at least. [From (10), (13), (14)]

The “for me at least” sounds worrisome, but Kant is explicit: “I must not even say ‘It is morally certain that there is a God’, etc., but rather ‘I am morally certain’ etc.” (KrV, A829/B857, original emphasis). Still, he thinks the proof does *not* count as an exercise in mere wishful thinking or self-deception, provided we acknowledge that

- the rational motive for action is not the production of the moral world order but simply the fact that the action is morally required
- the rational justification is moral rather than epistemic
- the result is not the attitude that we would call “belief” and Kant would call logical “conviction” (*Überzeugung*). The result, rather, is *Belief* (*Glaube*)
- the Belief in question is about a state of affairs that is evidentially ambiguous (in this case theoretically undecidable)
- the justification is defeasible – if Belief (*Glaube*) in God produces bad moral results, then it must be given up

Let’s take a closer look at a couple of the key premises. Regarding (6): everyone agrees that there is some sort of modal constraint on hope, but it is clear that superficial hopes often involve merely taking possibility

for granted (see Chignell 2014, 2022). Substantial, life-structuring hopes that p , on the other hand, arguably do require moderately clear conceptions of how, given the things and powers that exist in the world, p could turn out to be true.²⁵ It would be extremely hard to hope in a serious way that there will be perpetual peace – to really structure one’s life around this hope, for example – without having at least some positive conception of how we denizens of the actual world could really, practically get there from here. This is what I mean by “really, practically possible”.²⁶

Note, too, however, that by starting with the conditions on hope instead of on something stronger like Belief, we end up with a more acceptable moral-psychological pattern here. With respect to something like the Highest Good or “the moral world order”, it’s much easier (psychologically and rationally) to adopt hope that it will obtain than it is to form full-on belief or Belief that it will. But reflection on the conditions of having such hope then lead to the recognition in (6) that we are committed to its real, practical possibility, as well as anything else that is required for the latter.

(9) is obviously another lynchpin; a friend of the argument would have to rule out other accounts of the real, practical possibility of a moral world: dialectical historical processes, political revolutions, a karmic system that ensures that justice will be done, liberal democratic institutions and so on (Adams 1979). She would also have to explain why the full-blown classical deity, with all its omni-properties, is required to be the relevant object of trust. Kant himself makes efforts in this direction, suggesting that only a necessary, perfect being would be able and willing to arrange such an order. But many readers remain unconvinced (see Michalson 1999).

(11)–(14) are also crucial: the existence-claim here must be *evidentially ambiguous* in order for the Belief to be rational. Large swaths of the *Critique of Pure Reason* are dedicated to undermining traditional efforts to *prove* or even render *probable* the existence of God on either demonstrative or empirical grounds. Kant also rejects all atheistic arguments, including the empirical argument from evil. So, in the famous phrase, *all* knowledge-claims about God’s existence – theistic and atheistic – are “denied” in order to “make room” for the kind of moral Belief that overcomes demoralization (KrV, Bxxx).

In the passages from the *Religion* quoted at the beginning of this section, Kant is not discussing a moral world order generally but rather hope for *specific* outcomes of *specific* actions. But he clearly thinks that such specific hope, too, can underwrite Belief and trust in a supersensible mechanism: a providential “higher wisdom” that makes the “fulfilment of our well-intentioned efforts” possible. Returning to the contemporary inefficacy objection: it seems clear that people who *already* have such a supersensible commitment can just orient their hopes around that. They can focus, in other words, on the possibility that this mechanism arranges things such that their well-intentioned efforts *matter* – that they do indeed contribute to

the good, even if they seem in the moment to be futile. Kant rightly thinks that this would be important to them, even if they also trust that God will ultimately resolve things for the best in the end, no matter what. The hope is that *their* efforts will be included in the divine plan (so Kant's doctrine of Providence is anti-quietist). That hope in turn sustains and stabilizes their psychological resolve to keep doing what they regard as required.

I now want to turn to consider a version of this argument against demoralization that is available in secular contexts as well. Here the Belief and trust will not be in God but in fellow finite agents.

4. A secular moral-psychological argument in contemporary consumer contexts

Given that the morally objectionable A-system is deeply insensitive to slight changes in demand, any attempt on Hope's part to inculcate full-blown belief that *my abstention is going to make a significant positive difference* would be a miserable exercise in wishful thinking and self-deception. *Belief* in that proposition also looks more like a Kierkegaardian leap than sweet Kantian reason. As we have seen, Kant anticipates James and others in the moderate pragmatist tradition in rejecting the idea that a moral argument can support a conclusion whose *negation* we have strong reason to believe.

With this in the background, let's consider a moral argument that is based, like Kant's own moral-psychological proof, in a claim about what's required to maintain our moral resolve:

- 1* *Don't Purchase*: It is morally wrong to purchase the products of morally wrongful activity if alternatives are readily available. [From an independent argument or moral intuition]
- 2* It would be *demoralizing* in the (D1) sense (i.e. it would lead to discouragement and dejection) for me not to be able to have substantial hope that my abstention from purchasing A-products over time will make a significant positive difference with respect to A. [Empirical premise]
- 3* Such demoralization has an enervating effect on my resolve, and is thus *de-moralizing* in the second (D2) sense: I will almost certainly fail in my efforts to abstain from purchasing A-products over time. [Empirical premise]
- 4* Despair of this sort is seriously morally undesirable. [from (1*)–(3*)]
- 5* Therefore, there is serious moral advantage for me in being able to have substantial hope that my abstention from purchasing A-products will make a significant positive difference with respect to A. [From (2*)–(4*)]

So far, so good: this looks like a simple argument for the practical rationality of *hope* that by abstaining I will somehow make a significant positive difference with respect to the system in question, where “significant positive difference” involves an improvement in the welfare conditions of the animals and workers involved in A. This is much easier than directly generating *Belief* that I will make a significant positive difference; it’s also much less problematic from a rational point of view, given that my difference-making is deeply unlikely given the sobering empirical facts. Note that “significant” is left vague precisely because the amount of difference-making required to avoid despair will clearly differ from person to person.

What are the conditions on having this sort of resolve-sustaining hope? We saw earlier that

- 6* Substantial hope that p requires the positive belief or Belief that p is really, practically possible. [Theoretical premise]

And from (5*) and (6*), we can infer:

- 7* Therefore, there is serious moral advantage for me in being able to believe or have Belief that it is really, practically possible for my abstention from purchasing A-products to make a significant positive difference with respect to A.

Finally, as we saw previously,

- 8* If a being or state of affairs is really practically possible, then there must be something in the actual world that can account for that fact. [Concept of “real practical possibility”]

People of great willpower may find it psychologically possible to leave things there. Even if they dearly miss the taste of deep-fried spicy mesquite chicken or the cozy feel of a sweat-shopped hoodie, these moral pillars will abstain from purchasing in the steadfast hope that their actions, over time, will *somehow* make a significant positive welfare difference with respect to the industrial chicken system. In other words, despite the incredibly long odds (e.g. globally, the industry slaughters 2000 birds per second!), moral pillars can maintain their belief or Belief that it is really, practically possible for their abstention to make a significant difference – and that is enough to sustain their resolve.

But Kant correctly saw that such pillars are rare: remember that even someone as righteous as Spinoza is liable to despair in the absence of *some* reassurance that justice will prevail. For the rest of us, clinging in hope to the brute possibility that our choice will somehow make a significant difference is extremely hard, especially when the products in question are so convenient, tasty and affordable.

Others might try to shore up their resolve by focusing on the fact that their own pivotality in the supply chain is at least possible. It is just possible, in other words, that this is the month in which everyone else in the region will purchase exactly the right number of sandwiches – all the way up to the threshold point where another purchase will cause (via a series of effects down the supply chain) a new batch of (say) 5000 chickens to be incubated. Being pivotal like this is just barely possible, and if it obtains, then the choice to abstain will make a significant positive difference indeed.

Facts about individual empirical psychology again play a key role here. Someone like Hope who is threatened by demoralization can try to focus on the “infinitesimal” possibility of her own pivotality and see if that will sustain her. That is, she can try to give that possibility a prominent place in her mind every time her colleagues drag her to the bar on Shotz-n-Wingz Nite, and in this way allow the thought of all the chicken suffering she *might* be preventing to sustain her resolve.²⁷ But that’s not going to be easy: given the sobering empirical facts, the chance that she is *ever* at a threshold point – much less right there in the bar – is vanishingly small. It will thus be hard for Hope to keep from sliding into despair and doing what she takes to be morally wrong (especially after a few shotz).

According to the account of hope that I favor, the key difference between hope and despair has to do with where the subject is disposed to place her mental attention or “focus”. I might strongly desire something and regard it as possible, but if I keep focusing on it as massively improbable, I am in despair. If instead I am disposed to focus on it “under the aspect of its possibility” – that is, as having at least a chance of occurring – then I am hoping (Chignell 2022).²⁸ But other things equal, the more unlikely a desired outcome is, the harder it is to stay focused on it *as possible* rather than fixating on the overwhelming odds against. So while hoping for pivotality is an option for staving off despair, most of us will find it a hard psychological row to hoe.

Given the Kantian proof we considered in the last section, another way to go suggests itself:

- 9* The existence of God provides the only adequate account of the real, practical possibility that my abstention from purchasing A-products will make a significant positive difference with respect to A. [Theoretical premise]

If this were correct, then the rest of the argument could run like the previous one and conclude with full-blown moral theism. Only such Belief, and the trust in God that accompanies it, makes it psychologically possible to sustain the hope that my abstinence over time will make a significant positive difference to a morally objectionable system. As a result, that Belief and trust are *prima facie* morally justified.

But whereas (9) was debatable, (9*) is patently false. In other words, there was some plausibility in the idea that the existence of a supreme being is the *only* adequate account of the real, practical possibility of a moral world order (the Highest Good). But with respect to the more specific hope that Hope's individual actions will make a significant positive difference with respect to the A-system, (9*) itself is hopeless. After all, we just saw that there are empirical ways to account for the bare possibility here, even in a massive and lumpy supply system. It's at least possible, for instance, that Hope's choice is pivotal by way of being made on a threshold. So there can be nothing like a *proof* of moral Belief and trust that God, the "universe", karma, or fate is arranging for individual efficacy. Of course if someone already has a supersensible commitment that can do the job, then it makes sense for *him* to fix his hopes by appealing to a premise like

9** The existence of *supersensible mechanism X* provides an adequate account of the real possibility that my abstention from purchasing A-products will make a significant positive difference with respect to A.

But for someone (like Hope) who does not already have the supersensible commitment, there is no moral-psychological pressure to adopt one.

Given the difficulty Hope has in focusing on the slim possibility of her own pivotality, however, it would be useful to find some other naturalistic scenarios whereby boycotting an industrial system like the one that brings us our spicy mesquite chicken sandwiches might make a significant welfare difference. A candidate that comes to mind is this: Hope's decision might somehow be connected to the decisions of numerous *other* people such that, if she abstains, then a significant number of other people will or will be likely to abstain (where a "significant" number is precisely what's required for there to be a "significant" difference – keeping in mind that what counts as "significant" might be different across different subjects). This scenario divides into two: one causal and one evidential. I'll discuss each in turn.

The causal version of the scenario would invoke this premise:

9*** The existence of a *causal connection between my decision to abstain from A-products and the decisions of a significant number of other people to do the same* provides an adequate account of the real, practical possibility that my abstention from purchasing A-products will make a significant positive difference with respect to A.

This seems true: such a causal connection would provide an adequate account. Following the logic of the moral-psychological argument, we would then arrive at:

10* Therefore, there is serious moral advantage, for me at least, in being able to believe or have Belief that such an interpersonal causal connection exists. [from (7)–(9***)]

But could such belief or Belief be rational? That is, could it ever be rational to believe or have Belief in the following?

Interpersonal Causal Connection: If I choose to abstain from A-products, my action will **cause** a significant number of other people to abstain, and if I choose to purchase A-products, my action will cause a significant number of other people to purchase.

Interpersonal Causal Connection looks like a non-starter for people who aren't celebrities, dictators or top-level "influencers". There is no reason to think that there is any such connection between one's person's actions and the actions of a significant number of others. A quick look at Hope's number of Twitter followers provides decisive reason to think that there is *not* such a connection in her case.²⁹ But then both belief and Belief are ruled out on rational grounds (again, Kant is the forerunner of James rather than Kierkegaard here – an object of *rational* Belief has to be evidentially ambiguous).

There is a weaker version of the scenario, however, that looks more promising. Consider:

9**** The existence of a *strong evidential connection between my decision to abstain from A-products and the decisions of a significant number of other people to do the same* provides an adequate account of the real, practical possibility that my abstention from purchasing A-products will make a significant positive difference with respect to A.

9**** seems true, and so by the logic of Kant's moral-psychological argument, it would seem to give us *prima facie* moral justification for either belief or Belief that:

Interpersonal Evidential Connection: If I choose to abstain from A-products, that is strong **evidence that** a significant number of other people are likely to abstain from A-products, and if I choose to purchase A-products, that is strong evidence that a significant number of other people are likely to purchase A-products.

When applied to the case of Hope, is this a principle for which we have any warrant one way or the other? Here I think things are less obvious than they were with *Interpersonal Causal Connection*. It seems reasonable for Hope to think that a significant number of people *could* be motivated by the same

reasons against purchasing A-products that she is (in fact, such a thought might be a part of what it is to take them to be *reasons*). But that's different from thinking that if she chooses to abstain, a significant number of other people *will* be so motivated or are *likely to be* so motivated. Given that most of these decisions will be made in distant places and times (recall that Hope is not an "influencer"), it is *simply unclear* whether there is sufficient evidence for or against *Interpersonal Evidential Connection*. Indeed, even if over the short- to medium-term she acquires evidence that other people around her are *not* following her lead, her own choice might still count in favor of the idea that *over the long haul* a significant number of people (here or elsewhere) will see the power of her reasons.³⁰ This is particularly true if Hope is not a moral pillar but rather a fairly weak-of-will person – if *she* can do it, then certainly lots of other people can, too. The upshot here is that, in most contexts, anyway, forming a belief one way or other on the matter would be out of epistemic bounds. So *Interpersonal Evidential Connection* does not have sufficient epistemic grounds, and belief is not justified.

This is where Kant's famous thought that we can sometimes deny belief in order to make room for Belief and trust once again comes to the rescue. When there is no clear evidence one way or the other, we might still have *moral justification* for adopting a principle in this other, non-doxastic way. So if super-sensible mechanisms aren't a live epistemic option for Hope, and if the chance that she will be causally pivotal is too slim to sustain her focus, she may be defeasibly morally justified in taking *Interpersonal Evidential Connection* on *Belief*. That is: she might be defeasibly morally justified in holding, as an article of Belief, that if she acts on her boycotting reasons, then that is strong evidence that a significant number of other people (somewhere, sometime) are likely to abstain too. Thus, if she chooses to abstain now, she gains strong evidence that there will be a significant improvement in animal welfare.

Belief in a principle like that, and the trust in other people (somewhere, sometime) that it involves, could presumably stave off despair, even for those of us who are not moral pillars. However, if Hope thinks like this up to the moment of choice but then defects and purchases (perhaps assuming that all those other people will still act on the reasons that she has), she thereby *loses her evidence* that others will abstain. It is crucial to note that *Interpersonal Evidential Connection* says that she only gets the relevant evidence if she *actually* abstains and not just if she appreciates the reasons for abstaining. Morally justifying one's trust in others still involves taking on a certain amount of risk.³¹

There is obviously more to be said about this last proposal, but here I simply want to note two intriguing features. First, the reasoning is broadly Kantian in a manner that goes beyond the use of moral arguments. Hope relies on the idea that (other things equal) moral reasons are universalizable – a theme we typically associate with Kant. The "universal law" formulation of the categorical imperative says that I have a good reason to act on a certain maxim

only if it is also good reason for everyone else to do likewise in saliently similar circumstances. Here the universalization goes the other way, too: if it is *in fact* a good reason for *me* to act in such-and-such a way, then it is a good reason for everyone else in similar circumstances to do likewise. So my acting provides strong evidence that others (if they are rational) are likely to follow. When Hope “thinks for herself” about these matters, she also takes herself to be “thinking from the point of view of others” – a maxim that Kant thinks of as common sense (KU, AA 05:294). And so her moral Belief in *Interpersonal Evidential Connection* invokes a broadly Kantian idea about the universalizability of reasons.

Second, and even more intriguingly, if Belief in *Interpersonal Evidential Connection* is what sustains the hope to make a significant positive difference, then we have arrived in a very roundabout way at a key principle of *evidential decision theory* (EDT). According to EDT, an agent ought to perform actions that are such that, if she performs them, then the chances of the desired outcome are maximized, *regardless* of whether there is any causal connection between her actions and the outcome. This is sometimes characterized as the view that we should be guided by “auspiciousness over efficacy”. It also assumes an alternative view about difference-making:

Difference-Making, Evidential: An action A makes a difference with respect to situation S if performing A *provides evidence* that a change in S is likely to occur.

According to this principle, a causal connection between A and S is not required for A to make a difference with respect to S – the mere “auspiciousness” that A represents with respect to S’s changing for the better is sufficient.

(Side note: This is why EDT is thought to support taking just one box in the Newcomb problem.³² Even if there is no causal connection between your choice and the action of a near-perfect predictor of your choice [since the prediction has already been made], there is an *evidential* connection: what you choose provides strong evidence regarding what the predictor predicted. So on the evidential conception of difference-making, what you choose does make a difference and so you should take just the one box [which will thus contain a million dollars]. In the causal conception of difference-making, by contrast, your action can’t make a difference – the predictor has already done its work – and so you should be safe and take both boxes [thereby acquiring a thousand dollars]).³³

In the scenario we have just been considering, Hope abstains for independent moral reasons – they are what *rationally* motivate her. But she also has a psychological need to hope that her action makes a difference, especially over time. Sustaining that hope is of serious moral advantage, and this justifies her in accepting *Interpersonal Evidential Connection* as an item of moral Belief, with *Difference-Making, Evidential* in the background. So

Hope can have justified moral Belief that her purchasing and boycotting actions make a difference even while recognizing that the sobering empirical facts indicate that her action does not make a *causal* difference.³⁴

5. Conclusion

Discussions of the inefficacy or futility objection tend to focus on the conceptual issue – that is, on how inefficacy poses a challenge to traditional consequentialism. In this paper, I focused instead on the psychological side of the problem – that is, on how the *perception* of our inefficacy as individual consumers in gigantic industrial contexts can threaten not our rational motivation but our psychological resolve. I suggested that an analogue of Kant’s moral-psychological argument against despair can be used to justify various resolve-stabilizing strategies. These include focusing in hope on the possibility that

- 1 At least one of my abstaining actions occurs at a “threshold” point in the system such that my choice makes a significant positive difference to welfare outcomes.

If hope for such pivotality is psychologically out of reach, alternative strategies involve adopting defeasible moral Belief in one (or more) of the following claims:

- 2 A supersensible mechanism causally connects my abstinence to some significant positive difference in welfare outcomes.
- 3 My decision to abstain is strong evidence that a significant positive difference in welfare outcomes is likely, even if there is no causal connection between the two.

Although the argument has a structure that is analogous to Kant’s moral-psychological theistic proof, only one of the options here (i.e. 2) involves Belief or trust in something supersensible.

A final Kant-scholarly point: the structure of the moral-psychological argument I’ve been considering here reveals an underappreciated way in which “the practical has primacy” for Kant. The argument says that if we morally ought to act a certain way, and we are threatened by resolve-sapping despair, then we are defeasibly morally justified in implementing stabilizing strategies to sustain our hope and thus our resolve with respect to the ought in question. But although the moral commitment and the accompanying hope have *primacy*, Kant is not an advocate of irrational leaps. Theoretical reason does kick in at some point and require an account of how the hoped-for scenario could really, practically come about. That in turn makes us *prima facie* morally justified in having Belief in something that can ground that real, practical possibility. The Belief in question can even involve a

“theoretical” issue that is evidentially ambiguous: an existence-claim, for instance, or a principle in decision theory.³⁵

Notes

- 1 This chapter is a revised and abridged version of my contribution to a volume that is explicitly concerned with animal ethics (Allais and Callanan 2020), which is itself an expansion of some sections of a contribution to a handbook volume on food ethics (Barnhill et al. 2016). My thanks to the editors of these volumes, and of the present one, for allowing me to extract and revise material from those papers for use here. What is provided here is the essence of the argument in those other pieces, revised and updated in accordance with some more recent feedback.
- 2 Here consult the Netflix series *Rotten* (2019), which goes episode by episode through various popular food items and leaves the viewer thinking that no product in the industrial food system will ever be cruelty free.
- 3 I leave “alternatives” and “readily” unanalyzed for present purposes.
- 4 For instance: cultural trends can make the demand for certain *cuts* of poultry go up (turkey breast at Thanksgiving, chicken wings during years when the local football team does well, etc.). But animals come in wholes, and so if wings are in high demand this year, a decision not to eat a chicken sandwich (which is not made of wings) is even less likely to have any effect on how many chickens are produced and processed, since retailers are demanding as many wings as possible. Thanks to conversations with food scientist and poultry expert Joe Regenstein here; see also Parcell and Pierce 2000.
- 5 Halteman and McMullen (2018) acknowledge these facts about buffers in the real-world poultry supply, but then note that the consumer will “know little to nothing about the distance to a threshold at their particular retailer, and even less further down the supply chain”. This is true, but in the noisy, buffered market situation described here, a consumer’s ignorance of precisely where he is relative to the next threshold seems moot. He merely needs to be able to recognize that the chance of ever being on *any* threshold is extremely small in order for the futility concerns to arise. Oppy can easily come to know that the most likely outcome – given the buffers, the noise and the trend – is that his decision to abstain will simply increase that month’s wasted oversupply by precisely one patty.
- 6 For detailed versions of this argument, see Budolfson 2016, 2018. For various versions of the standard response, see Singer 1980; Kagan 2011; Norcross 2004; Halteman and McMullen 2018.
- 7 For the term “opportunistic carnivore”, see Almeida and Bernstein 2000. For more discussion of Oppy and opportunism, see Chignell 2016.
- 8 Almeida/Bernstein (2000, 206) state this as a biconditional, but the point I’m making doesn’t require something that strong.
- 9 Some welfarist views restrict the morally relevant outcomes to human animals only. Joseph Raz calls this “the humanistic principle”: “the explanation and justification of the goodness or badness of anything derives ultimately from its contribution, actual or possible, to human life and its quality” (1986, 194, see also Raz 1980). Here I am working with a welfarist picture that is not anthropocentric (and speciesist) in this way.
- 10 The privacy condition is meant to forestall the objection that his action might have an influence on the behavior of others around him.
- 11 See Halteman and Mullen 2018.

- 12 Note that the sobering empirical facts also suggest that the choice to purchase small, family-farmed products is *much* more likely to make a difference. So if someone like Oppy maintains that *Don't Produce* holds across the board, he should only opportunistically indulge with products from the *industrial* system.
- 13 See Parfit's reformulations of the Categorical Imperative in response to the "problem of imperceptible differences" (2011, 341).
- 14 For an application of symbolic value theory to questions about which foods to consume, see Cuneo 2016.
- 15 I attempted some of that in Chignell 2016. See Nefsky 2018 for a helpful overview of such efforts. Kingston (2020, draft) argues in an opportunistic spirit that a focus on individual consumption patterns is a distraction. He calls instead for a "relinquishment" approach whereby we take the benefits (in time, money, and resources) that we gain from consuming readily available industrial products and try to "remediate" in more effective (often political or activist) ways.
- 16 Note that this means that the inefficacy doesn't even need to be real; the supply chain and causal situation just has to be *perceived* as insensitive in the relevant ways.
- 17 Kant uses this term in a variety of contexts, and there is no good English translation of it. In other work I use the technical (and Teutonic) capitalization "Belief" and will adopt that practice here too, in order to avoid the overly religious connotation of "faith." Note that Kant has a non-doxastic conception of Belief (which he sometimes also calls "acceptance" [*Annehmung*]). It is not a response to the presence of perceived evidence but rather a voluntary positive "holding-for-true" (*Fürwahrhalten*) regarding a proposition based on "subjective" and sometimes context-sensitive grounds. See Chignell 2007a, 2007b for more on this notion. Also Pasternack 2011, 2014 and Chance and Pasternack 2018, as well as Wood 2020, who likewise adopts the convention of translating the term as "Belief". For a contemporary discussion of non-doxastic conceptions of faith, see Howard-Snyder 2016, 2019.
- 18 To establish this, we would need sociological studies of consumer and activist behavior in response to perceptions of inefficacy. Interestingly, the North American Meat Institute itself cites recent CNN and USDA survey data showing that of the approximately 5 percent of Americans who claim to be vegetarians, *around 65 percent will confess, when pressed, to having eaten meat products in the past 24 hours*. The Meat Institute's conclusion? "Bottom Line: Meat Is Amazing – And Irresistible" (North American Meat Institute, n.d., 2). www.meatinstitute.org/index.php?ht=d/sp/i/101931/pid/101931, accessed 15 Oct 2020.
- 19 Camus (1942) famously rejects this: he considers the same predicament and says that we must both accept the demands of the moral law and embrace the absurdity of a world in which justice never prevails: "One must imagine Sisyphus happy".
- 20 See Pace and McKaughan (2022) for an argument that faith has typically been regarded in Jewish and Christian traditions as involving trust and a kind of loyalty.
- 21 The "effective altruism" movement introduces itself as follows:

Most of us want to make a difference. We see suffering, injustice, and death, and are moved to do something about them. But working out what that "something" is, let alone actually doing it, is a difficult problem. It would be easy to be disheartened by the challenge.

(www.effectivealtruism.org/articles/introduction-to-effective-altruism, accessed 1 November 2020)

- 22 Adams (1979) sketched an empirical-psychological way of reconstructing Kant's proof a long time ago. My presentation here is influenced by his, though also different in many key respects. In the first *Critique*, Kant presents an early version of the moral proof according to which hope for happiness seems to be part of the *incentive* for acting rightly. In the second *Critique*, he provides the more canonical articulation of the proof that starts not with (1) previously but with the more controversial claim that we ought to will the highest good, and this involves, as a sort of rational presupposition, moral Belief in the existence whatever is required to make the highest good really possible (i.e. God and the afterlife). For articulations of that form of the proof, see Wood 1970; Willaschek 2016. Less attention has been devoted to the moral-psychological argument from despair in the third *Critique* and *Religion*, although see Van Impe 2014; Fugate 2014; Ebels-Duggan 2015 and Chance and Pasternack 2018.
- 23 Kant is a conceptual evidentialist about what *we* would call "belief" – the kind of holding-for-true (*Fürwahrhalten*) that can count, if true and justified, as knowledge (*Wissen*) (Kant calls it "conviction" (*Überzeugung*), as do many contemporary German epistemologists). See Chignell 2007a.
- 24 Rational Belief, for Kant, is a voluntary state of holding-for-true (*Fürwahrhalten*) that, for non-epistemic reasons, a subject uses to guide deliberation, action and assertion in certain contexts. See Chignell 2007b.
- 25 For the claim that there is a distinction between "superficial" hope and "substantial hope" and that the latter has extra conditions on it, see Pettit 2004; McGeer 2004. In my view, one of the main differences is found in the cognitive condition: substantial hope involves a more articulate sense of how the outcome is really, practically possible.
- 26 Compare Willaschek 2016 on "practical possibility".
- 27 For more on how "Shotz-n-Wingz Nite" can destroy a person's resolve, see Halteman and Halteman Zwart 2016, p. 131. Those authors claim that such events occur regularly at a place called "Baloneez", but Barnhill et al. (2016, 171) also reference a sister establishment called "Jimmy's You-Hack-it-Yourself BBQ".
- 28 See Chignell 2022.
- 29 Obviously Hope could try to *establish* such causal connections (and thereby shore up her resolve) by signaling her choices publicly when she can, *attempting* to influence others, becoming part of a broader cultural movement, seeking to become or enlist a social media influencer and so on. And obviously the question of whether her actions have made any causal difference must be considered over the longer haul. It's not clear how much this will help in private one-off moments under conditions of perceived inefficacy, however. Thanks to Elizabeth Harman, Renée Bolinger, Teresa Morgan and Judith Gundry for discussion here. Compare Lawford-Smith 2015.
- 30 There are many dissimilarities between the cases, but just by analogy, think of someone in, say, 18th-century South Carolina who decides, on moral grounds, to free the slaves that he inherited from his family. Those reasons were good, and in order to sustain his resolve, he might also have taken on Belief that his choice was strong evidence that others would (at some point) do the same. But it took a *long* while.
- 31 Thanks to Victoria McGeer for discussion here.
- 32 This problem, which has now generated thousands of articles, was originally stated by Nozick in 1969.
- 33 Jon Elster cites empirical evidence for the claim that people use "diagnostic thinking" to move from the assumption that they are "fairly typical members"

of a reference group to the conclusion that others “will tend to act like me”. Elster sees no flaw in this sort of reasoning *unless* explicitly causal claims are made – that is, unless people start thinking that “my action will bring it about” that others do the same. But the EDT conception of difference-making avoids that sort of “interpersonal magic”. See Elster (1985, 142–145). Thanks to Ewan Kingston (who pointed me to the Elster paper), Philip Pettit, Richard Bradley, Victoria McGeer and Kian Mintz-Woo for discussion here. Bradley and Mintz-Woo argue that this interpersonal “evidentialist” approach to consumption and voting cannot rationally be extended to Prisoner’s Dilemma cases, however, since defecting is *always* the best thing to do in the latter, even in EDT. This is presumably why followers of EDT have resisted David Lewis’s claim (in 1979) that “Newcomb’s Problem is a Prisoner’s Dilemma.” Compare Pittard 2018.

- 34 If this is correct, then the present broadly Kantian response to the psychological inefficacy problem has the added benefit of providing *prima facie* moral justification for Belief in a principle that supports one-boxing. There is something right about the prosperity gospel after all: Belief can indeed make you rich. . .
- 35 For discussions of these ideas, I am grateful to the editors of this volume and to Lara Buchak, Anne Barnhill, Renée Jorgensen, Mark Budolfson, Gabriel Citron, Silvia De Toffoli, Tyler Doggett, Gabriele Gava, Matthew Halteman, Daniel Howard-Snyder, Ewan Kingston, Tania Lombrozo, Victoria McGeer, Dan McKaughan, Kian Mintz-Woo, Philip Pettit and Leigh Vicens. I am particularly indebted to McGeer for lengthy written comments and long conversations. I also thank members of audiences and working groups at Cornell University, Humboldt University in Berlin, Goethe University in Frankfurt, the Hungarian Academy of Sciences in Budapest, Georgetown University, University of Witwatersrand, the Center for Human Values at Princeton, the Oxford University “Trust” workshop and the Eighth Kant Multilateral Colloquium in Catania, Italy. Some of this work was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

References

- Adams, R.M. 1979. “Moral Arguments for Theistic Belief.” In C.F. Delaney (ed.), *Rationality and Religious Belief*. Notre Dame: University of Notre Dame Press, 116–140.
- Albertzart, M. 2019. “A Kantian Solution to the Problem of Imperceptible Differences.” *European Journal of Philosophy* 27: 837–851.
- Allais, L., and Callanan, J. (eds.). 2020. *Kant and Animals*. Oxford: Oxford University Press.
- Almeida, M., and Bernstein, M. 2000. “Opportunistic Carnivorism.” *Journal of Applied Philosophy* 17(2): 205–211.
- Appiah, K.A. 1986. “Racism and Moral Pollution.” *Philosophical Forum* 18(2): 185–202.
- Barnhill, A., Budolfson, M.B., and Doggett, T. (eds.). 2016. *Food, Ethics, and Society*. New York: Oxford University Press.
- Barnhill, A., Budolfson, M.B., and Doggett, T. (eds.). 2018. *Oxford Handbook of Food Ethics*. New York: Oxford University Press.

- Boey, G. 2016. "Eating Animals and Personal Guilt: The Individualization of Responsibility for Factory Farming." In A. Barnhill, M.B. Budolfson, and T. Doggett (eds.), *Food, Ethics, and Society*. New York: Oxford University Press, 22–29.
- Bramble, B. 2015. "The Case against Meat." In B. Bramble and R. Fischer (eds.), *The Moral Complexities of Eating Meat*. New York: Oxford University Press, 135–152.
- Budolfson, M.B. 2016. "The Inefficacy Objection to Utilitarian Theories of the Ethics of the Marketplace." In A. Barnhill, M.B. Budolfson, and T. Doggett (eds.), *Food, Ethics, and Society*. New York: Oxford University Press, 204–210.
- Budolfson, M.B. 2018. "The Inefficacy Objection to Consequentialism, and the Problem with the Expected Consequences Response." *Philosophical Studies* 176: 1711–1724.
- Budolfson, M.B. "The Inefficacy Objection to Deontology: What It Is, Why It Is Important, and the New Type of Deontological Reason Required to Reply to It." Draft.
- Camus, A. 1942 [trans. 1955]. *The Myth of Sisyphus*. London: Hamish Hamilton.
- Chance, B., and Pasternack, L. 2018. "Rational Faith and the Pantheism Controversy: Kant's 'Orientation Essay' and the Evolution of His Moral Argument." In D. Dahlstrom (ed.), *Kant and His German Contemporaries*. New York: Cambridge University Press.
- Chignell, A. 2007a. "Kant's Concepts of Justification." *Nous* 41(1): 33–63.
- Chignell, A. 2007b. "Belief in Kant." *Philosophical Review* 116(3): 323–360.
- Chignell, A. 2014. "Rational Hope, Possibility, and Divine Action." In G.E. Michalson (ed.), *Religion within the Bounds of Mere Reason: A Critical Guide*. New York: Cambridge University Press, 98–117.
- Chignell, A. 2016. "Can We Really Vote with Our Forks? Opportunism and the Threshold Chicken." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 182–202.
- Chignell, A. 2022. "The Focus Theory of Hope." *The Philosophical Quarterly*, online first, print version forthcoming.
- Cuneo, T. 2016. "Conscientious Omnivorism." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 21–38.
- Ebels-Duggan, K. 2015. "The Right, the Good, and the Threat of Despair: (Kantian) Ethics and the Need for Hope in God." In J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion*, vol. 7. New York: Oxford, 81–109.
- Elster, J. 1985. "Rationality, Morality, and Collective Action." *Ethics* 96(1): 136–155.
- Frey, R.G. 1985. *Rights, Killing, and Suffering: Moral Vegetarianism and Applied Ethics*. Oxford: Blackwell.
- Fugate, C. 2014. "The Highest Good and Kant's Proof(s) of God's Existence." *History of Philosophy Quarterly* 31(2): 137–158.
- Halteman, M., and Halteman Zwart, M. 2016. "Philosophy as Therapy for Recovering (Unrestrained) Omnivores." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 129–148.
- Halteman, M., and McMullen, S. 2018. "Against Causal Inefficacy Objections: The Real Economic Impact of Individual Consumer Choices on Animal Agriculture." *Food Ethics* 2–3: 93–110.

- Harman, E. 2016. "Eating Meat as a Morally Permissible Moral Mistake." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 215–231.
- Hill, T.E. 1979. "Symbolic Protest and Calculated Silence." *Philosophy and Public Affairs* 9(1): 83–102.
- Howard-Snyder, D. 2016. "Does Faith Entail Belief?" *Faith and Philosophy* 33(2): 142–162.
- Howard-Snyder, D. 2019. "Three Arguments to Think that Faith Does Not Entail Belief," *Pacific Philosophical Quarterly* 100(1): 114–128.
- Kagan, S. 2011. "Do I Make a Difference?" *Philosophy and Public Affairs* 39(2): 105–141.
- Kingston, E. 2020. "Shopping with a Conscience? The Epistemic Case for Relinquishment over Conscientious Consumption." *Business Ethics Quarterly* 31(2): 242–274.
- Kingston, E. "Consumer Ethics: Relinquishing Benefits from Unruly Commerce." Draft.
- Lawford-Smith, H. 2015. "Unethical Consumption and Obligations to Signal." *Ethics & International Affairs* 29(3): 315–330.
- Lewis, D. 1979. "Newcomb's Problem is a Prisoner's Dilemma." *Philosophy & Public Affairs* 8: 235–240.
- Martin, A. 2016. "Factory Farming and Consumer Complicity." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 203–214.
- Martin, M. 1976. "A Critique of Moral Vegetarianism." *Reason Papers* 3: 13–43.
- McGeer, V. 2004. "The Art of Good Hope." *Annals of the American Academy of Political and Social Science* 592: 100–127.
- McPherson, T. 2016. "Why I Am a Vegan (and You Should Be One Too)." In A. Chignell, T. Cuneo, and M. Halteman (eds.), *Philosophy Comes to Dinner*. New York: Routledge, 73–91.
- Michalson, G.E. 1999. *Kant and the Problem of God*. New York: Blackwell.
- Michaelson, E. 2016a. "Act Consequentialism and Inefficacy." In A. Barnhill, M.B. Budolfson, and T. Doggett (eds.), *Food, Ethics, and Society*. New York: Oxford University Press, 210–214.
- Michaelson, E. 2016b. "A Kantian Response to Futility Worries?" In A. Barnhill, M.B. Budolfson, and T. Doggett (eds.), *Food, Ethics, and Society*. New York: Oxford University Press, 215–218.
- Nefsky, J. 2018. "Consumer Choice and Collective Impact." In A. Barnhill, M.B. Budolfson, and T. Doggett (eds.), *Oxford Handbook of Food Ethics*. New York: Oxford University Press, 267–286.
- Nobis, N. 2002. "Vegetarianism and Virtue: Does Consequentialism Demand Too Little?" *Social Theory and Practice* 28(1): 135–156.
- Norcross, A. 2004. "Puppies, Pigs, and People: Eating Meat and Marginal Cases." *Philosophical Perspectives* 18: 229–245.
- North American Meat Institute. n. d. *The Amazing Meat and Poultry Supply*. www.meatinsitute.org/index.php?ht=d/sp/i/101931/pid/101931
- Nozick, R. 1969. "Newcomb's Problem and Two Principles of Choice." In N. Rescher (ed.), *Essays in Honor of Carl Hempel*. Dordrecht: Reidel, 440–472.

- Pace, M., and McKaughan, D. 2022. "Judaean-Christian Faith as Trust and Loyalty." *Religious Studies* 58(1): 30–60.
- Parcell, J., and Pierce, V. 2000. "Factors Affecting Wholesale Chicken Prices." *Journal of Agricultural and Applied Economics* 32(3): 471–478.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, D. 2011. *On What Matters*. Oxford: Oxford University Press.
- Pasternack, L. 2011. "The Development and Scope of Kantian Belief: The Highest Good, the Practical Postulates and the Fact of Reason." *Kant-Studien* 102(3): 290–315.
- Pasternack, L. 2014. "Kant on Opinion: Assent, Hypothesis, and the Norms of General Applied Logic." *Kant-Studien* 105(1): 41–82.
- Pettit, P. 2004. "Hope and Its Place in Mind." *Annals of the American Academy of Political and Social Science* 592: 152–165.
- Pittard, J. 2018. "Evil and God's Toxin Puzzle." *Nous* 52(1): 88–108.
- Singer, P. 1980. "Utilitarianism and Vegetarianism." *Philosophy and Public Affairs* 9(4): 325–337.
- Raz, J. 1980. "Utilitarianism and Vegetarianism." *Philosophy and Public Affairs* 9(4): 325–337.
- Raz, J. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
- Van Impe, S. 2014. "Kant's Moral Theism and Moral Despair Argument against Atheism." *Heythrop Journal* 55: 757–768.
- Willaschek, M. 2016. "Must We Believe in the Realizability of Our Ends? On a Premise of Kant's Argument for the Postulates of Pure Practical Reason." In T. Höwing (ed.), *The Highest Good in Kant's Philosophy*. Berlin: De Gruyter, 223–244.
- Wood, A. 1970. *Kant's Moral Religion*. Ithaca, NY: Cornell University Press.
- Wood, A. 2020. *Kant and Religion*. New York: Cambridge University Press.

LYING, DECEPTION AND DISHONESTY

Kant and the contemporary debate on the definition of lying

Stefano Bacin

1. Introduction

Lying is currently an important focus of the philosophical debate, after having been rather neglected for quite some time.¹ One of the very few classical authors who consistently feature in the recent discussion, along with Augustine and, more occasionally, Aquinas, is Kant.² Remarkably, most of the history of moral philosophy is entirely neglected, as it is apparently believed to have contributed nothing to the matter. Even the few writers who are in fact considered, however, are not always examined in depth. Among the classical authors, Kant is arguably the one whose views play the most prominent role in contemporary discussions. Nevertheless, the frequent interaction with his examination of lying is mostly confined within rather specific boundaries. In fact, Kant's view enters the stage of the contemporary discussion basically in only three not mutually exclusive ways. First and foremost, it is customary to refer to Kant's view as the paramount example of the perplexing absolutist claim that lying is never allowed. According to many contemporary writers, Kant's remarks display most clearly the counterintuitive traits of such a notion. Second, and related to that, writers often point out passages from the lecture notes where Kant is reported to have allowed exceptions to the universal prohibition of lying, as this would show that Kant himself entertained a less rigoristic notion at some point. Third, the increasingly unpopular idea that lying is by definition morally worse than deception is mostly discussed, and often rejected, as a Kantian idea.

If these ways to consider Kant's contribution to the philosophical analysis of lying do show some promise, they are nevertheless significantly restricted.³ Kant's arguments for the apparently counterintuitive claim that lying is always wrong are hardly taken under scrutiny. The understandable attention, even

puzzlement, toward the passages from the lectures that allow us to reconstruct a much more nuanced view should rather be modulated by the consideration that those passages cannot lessen the official absolutist claim. In fact, the considered view in the later writings should be taken even more seriously in light of the earlier differentiated explorations. The most notable trait in the predominant way of treating Kant's view from the standpoint of the current debate, however, is that there seems to be no genuine interest for his perspective on one prominent topic in the contemporary discussion itself, namely the definition of lying. Most current debates primarily concern this complex issue, under the apparent assumption that the entire previous history of philosophical investigations of lying share one and the same traditional notion of what a lie is. Thus, surprisingly little attention is paid to the care that Kant devotes to how lying is best defined.⁴ His view on the foremost issue in the contemporary discussion is implicitly assimilated to that of the other classical writers without any qualification. The limits of this consideration hinder the possibility of a fruitful exchange between current perspectives and Kant's take on the matter. Both our understanding of Kant's view and the discussion on the defining features of lying would gain from a closer dialogue.

For the sake of a fruitful dialogue, the most relevant question to address is how lying in general is construed. The contemporary discussion is for the most part focused on this central issue, with the aim of finding a definition that is able to account for most, if not all, features of the highly complex activity that is called lying. The interest is thus primarily descriptive, in contrast with the primarily normative interest of Kant's examination, whose main aim is to establish whether lying is morally acceptable, and why. This contrast, however, should not be overemphasized, but it should rather be clarified. Also, a normatively minded analysis like Kant's entails an understanding of what is to be called lying at all. I will leave aside many interpretive questions and disregard the broader context of Kant's moral theory, which should be taken into account for a full examination. More specifically, the aim of my examination will be to clarify Kant's account of lying with regard to the terms of contemporary debates, not to corroborate or justify it, which would require further work.

I shall thus first consider where Kant stands in respect to the most central questions in the contemporary discussion on lying, that is, whether lying is to be understood as a form of deception. A closer look into the role that deception plays in lying will allow clarification of Kant's construal of lying in general. I shall suggest that, in Kant's view, deception is not, and cannot be, the defining feature of lying, in spite of several passages in which Kant appears to advocate a so-called "traditional definition" of lying, according to which it essentially requires an intent to deceive. To the contrary, Kant's view is able to accommodate non-deceptive lies, such as so-called "bald-faced lies". I shall suggest, thus, that Kant's view should be construed as a specific variant of intrinsic anti-deceptionism. Then, I shall examine where

Kant's account stands with respect to a recent proposal to consider dishonesty, instead of deception, the pivotal notion in discussing lying. Finally, I shall comment on some important methodological differences between Kant's account and the contemporary discussion with regard to the role of intuitions and definitions. Last but not least, a comparison with the terms of the contemporary discussion is helpful to highlight some notable methodological features of Kant's approach to this issue, which provides a good example of his way to develop a moral theory, in contrast with currently widespread methodological assumptions.

2. Lying and deception

Lying is a topic of interest from many different standpoints. Linguistics, psychology, philosophy of language, and moral, legal and political philosophy share a common interest in the widespread activity that is usually called lying. This multidisciplinary relevance has contributed to bringing the topic to the foreground in the recent philosophical discussion. Because of the variety of perspectives that come together in sharing that interest, the issue of a satisfactory definition of lying, which should provide the necessary common ground, has been attracting much attention.

According to the general outlook of the contemporary discussion, lying has been traditionally defined as a kind of deception or essentially based on an intent to deceive. Views that understand lying in these terms are thus usually labeled deceptionist. It is mostly assumed that this holds true for Kant's view, too, as representative for the traditional conception of lying. Furthermore, it is usually assumed that absolutist views that regard lying as unconditionally wrong have to entail a commitment to a deceptionist notion of lying. As Don Fallis observes,

“[d]ivorcing lying from deception” only looks like an unhappy result if we have a preexisting commitment to the wrongness (or at least, *prima facie* wrongness) of lying. Of course, many philosophers (most notably, Augustine and Kant) clearly do have such a commitment.

(Fallis 2015, 93)

Accordingly, Kant is widely considered to hold a deceptionist conception. In the following, I shall put this assumption under closer scrutiny.

Kant's remarks on lying do include formulations of some version of a traditional definition. In the *Doctrine of Virtue*, the first of the casuistical questions that follow the treatment of lying in § 9 asks: “Can an untruth from mere politeness . . . be considered a lie?” Kant answers: “No one is deceived by it” (6:431).⁵ The remark seems to follow from the assumption that one statement cannot count as a lie if it cannot possibly deceive anyone.⁶ Yet the

remark is not unambiguous, given that in other passages, Kant argues against the permissibility of lies said out of politeness (see e.g. 27:701). A deceptionist thought is expressed much more explicitly earlier in the same section, however. When Kant introduces the possibility of what he calls an “inner lie”, he observes that a lie “requires a second person whom one intends to deceive” (6:430). Other passages are analogously suggestive of an intrinsic connection between lying and intent to deceive. In the *Vigilantius* lectures notes from the early 1790s, Kant is reported to have explained to his students that “an untruth differs from a lie in this, that both, indeed, contain a *falsiloquium*, i.e., a declaration whereby the other is deceived, but the latter is uttered with an associated intention to injure the other by the untruth” (27:700). It soon becomes clear that this definition is specifically relevant to the juridical domain, since Kant is reported to have further observed: “In ethics . . . every *falsiloquium*, every knowing deception [*jedes wissentliche Hintergehen*] is impermissible, even though it be not immediately coupled with an injury” (27:700).⁷ The main focus of attention, here as in other texts, is the distinction between cases in which the possible harm done to another person through an intentional untruth is relevant or not. In both sorts of cases, however, lying is an intentional deception, according to these remarks.

A deceptionist account of lying, however, cannot merely amount to providing a definition in terms of the intent of deceive. Two further important elements belong in such an account: (1) the thought that the intent to deceive is what explains the wrongness of lying and (2) a more precise characterization of the nature of the deception involved in lying.

As to the first point, a deceptionist conception of lying traditionally holds that the wrongness of lying is determined by the intent to deceive (see Mahon 2018b, 51). A classical attempt is to construe lying as deception in terms of manipulation, to which the wrongness of lying would ultimately come down. In this view, “deceiving people (or at least some people, in some circumstances) is an example of using or manipulating them, and that that is what is wrong with it” (Williams 2002, 93).⁸ But Kant never suggests such a construal of lying. This would fall under what he understands as the possible harm done to others by lying to them. His main point, however, is that what makes lying *as such* wrong, both in the juridical and in the ethical sphere, is independent from any harm (see 6:430, 8:426). There are two main issues with this way of framing a deceptionist conception that could apply to Kant’s view. First, this view would reduce lying to a more general vice, that is, the infringement of others’ right to autonomous choices. This does not correspond to Kant’s perspective though, which always underscores the immediate wrongness of lying *per se*. Second, this view would make lying a merely other-regarding wrong, whose moral salience would only be in the relationship to others. But this is not what Kant maintains. Not only he construes lying as a violation of an ethical duty to oneself (see 6:420 and 429; I shall come back to this central thought in the following section). Also

in the juridical case examined in the *Supposed Right to Lie* essay, lying is condemned regardless of the relationship of the liar to another person, since Kant argues that it is wrong towards “humanity in general” (8:426; cf. Kae- hler 328, 27:447). In neither case is the wrongness of lying determined by the intent to deceive someone else, thereby harming a particular other person.

As to the second point, deceptionism should take a distinction of two kind or layers of deception into account. The intent to deceive can concern both the state of affairs at issue (the content of the communicative act) and one’s own thoughts (the deceiver’s belief). The main focus of attention is usually the first kind, or layer, of deception, namely how the liar aims at causing a false belief in someone else regarding a state of affairs. In contrast to this, Kant’s view concentrates on the second sort of deception, that is, on how lying infringes on the communication of one’s thoughts. A liar is not someone who deceptively generates false beliefs in other people but someone “who does not himself believe what he tells another” (6:429). Lying is thus not necessarily about causing false beliefs in others, which is how deception is mostly understood in the contemporary debate.⁹ The “second person whom one intends to deceive” (cf. 6:430) is, in fact, merely the addressee of an untruthful statement regarding one’s thoughts. Kant’s view excludes that the epistemic harm done to another person’s beliefs must be considered the distinctive feature of lying.

In spite of the passages that I have first mentioned, thus, Kant’s view of lying cannot be merely construed as a version of deceptionism, since it holds that (1) the moral worth of lying is not determined by its being deceptive, and (2) the deception that Kant refers to as belonging to lying primarily (or even solely) regards one’s thoughts. We have thus to consider further elements in Kant’s view.

3. Anti-deceptionism in Kant’s ethics

In spite of the prominent role given to deception in the passages that I have considered so far, any reference to the intent to deceive is remarkably absent from the definition of lying in the *Doctrine of Virtue*, which should be regarded as the main statement of Kant’s considered view on the matter, from the standpoint of ethics.¹⁰ There Kant characterizes lying as an “intentional untruth [*eine vorsätzliche Unwahrheit*] in the expression of one’s thoughts” (6:429). The *Supposed Right to Lie* essay presents an equivalent definition: “a lie, defined merely as an intentionally untrue declaration to another” (8:426). Here Kant does not hint at any intent of deceiving, nor does he leave room for it. The constitutive feature of lying is simply to declare an “untruth” in giving voice to one’s mind.

As in the clarification given in the *Vigilantius* lectures (27:700), the definition in the *Doctrine of Virtue* differentiates the ethical from the juridical aspect of lying. The need for such a distinction stems from the different relevance of the harm done to other persons. From the juridical standpoint,

lying must be condemned insofar as it infringes on another's individual rights, whereas this is not relevant to the ethical appraisal, which regards it as blameworthy independently from that consideration. With regard to this difference, the Vigilantius notes and the *Doctrine of Virtue* are in agreement. The most notable difference between them, however, concerns exactly whether Kant's view can be appropriately construed as deceptionism, since the *Doctrine of Virtue* drops any reference to deception, whereas the lecture notes still present lying as "knowing deception" (27:700) from the ethical perspective.

Contrasting the passages from the Vigilantius notes and the *Doctrine of Virtue* makes it apparent that Kant's treatments of lying include apparently diverse elements, which in terms of the contemporary discussion would be considered incompatible. A closer inspection shows that Kant's view, however, must not be regarded as inconsistent. Kant rather deploys a traditional definition to make an original point. He does present lying as deceitful at some point but develops an account that differs significantly from a conception centered on the intent to deceive.

Kant construes lying in a way that also covers cases in which the addressee is not justified in believing that the liar is truthful, nor is any intent to deceive in place. Beyond the nominal clarifications in terms of deception he sometimes deploys, Kant's account of lying, thus, is in fact able to accommodate cases of non-deceptive lying. Those are lies, even if they cannot intend, or hope, to deceive. Indeed, Kant mentions such cases, for instance, in the following example, which is presented in lecture notes from the 1770s:

The Inquiring Thief: "Somebody who knows that I have money asks me: Do you have money on you [*bey dir*]? If I keep silent, the other concludes that I do. If I say yes, he takes it away from me; if I say no, I tell a lie; so, what is to be done? So far as I am constrained by force against me to make a confession [*ein Geständniß von mir zu geben*], and some unlawful use is made of my statement, and I am unable to get out of this by remaining silent, the lie is a defensive measure. The declaration extorted, which is to be misused, permits me to defend myself; for whether my admission or my money is extracted is all the same."

(Kaehler 330, cf. 27:448)¹¹

Kant presents this case to discuss the permissibility of a "necessary lie", a *Notlüge*. Unlike in his considered view, at this point, he is still willing to acknowledge that an untruthful statement might be allowed or excused when the statement is given under duress. Kant's construal of the cases, however, shows that he does not understand the potential lie in terms of deception. The possibility of deceiving the thief is *per hypothesi* excluded, as he "knows that I have money". Kant points out, instead, that what would

make the statement a potential lie, is that it is “a confession [*ein Geständniß von mir*]”, that is, a declaration of the speaker’s mind. If here Kant holds that that statement does not count as a lie, it is only because the declaration has been coerced and thus cannot be taken as genuine. In absence of this condition, an untruthful declaration to someone who could not be deceived (and whom the speaker accordingly cannot intend to deceive) counts as a lie. Kant’s take on the case, thus, is significantly different from a traditional deceptionist account.

Similar cases belong to the non-deceptive lies, the so-called “bald-faced lies”, which are currently one of the most intensely debated issues.¹² Even if lies are often, or mostly, motivated by an intent to deceive, the possibility that this does not have to be the case brings the alternative between deceptionism and anti-deceptionism to the fore. The Inquiring Thief is a case in point because of the features that I have highlighted.¹³ Along similar lines, one of the most frequently examined cases in the current literature goes as follows:

The Cheating Student: “Suppose that a college Dean is cowed whenever he fears that someone might threaten a law suit and has a firm, but unofficial, policy of never upholding a professor’s charge that a student cheated on an exam unless the student confesses in writing to having cheated. . . . A student is caught in the act of cheating on an exam by copying from a crib sheet. . . . The student is privy to information about the Dean’s *de facto* policy and, when called before the Dean, he (the student) affirms that he did not cheat on the exam. . . . The student says this on the record in an official proceeding and thereby warrants the truth of statements he knows to be false. He intends to avoid punishment by doing this. He may have no intention of deceiving the Dean that he did not cheat.”

(Carson 2010, 21)

The Cheating Student does not aim to conceal his guilt, if only because he cannot expect to deceive the addressee of his statements, the dean, about the matter at hand. One could suggest that a deceptionist view can account for such cases too, after all, since the bald-faced liar does aim to “conceal information” (Lackey 2031, 241). Still, her acts are not about generating or communicating a false belief about the matter at issue. The dispute is thus about why bald-faced lies do attract blame, even if they do not follow from an intent to deceive.

Now, it is often assumed that “absolutist” conceptions of lying as such are not able to acknowledge bald-faced lies, primarily because a rigorous prohibition would have to be based on a strict definition, which is supposed to be unable to include such cases. This would accordingly apply to Kant’s view as well, in its traditional rendition.¹⁴ In contrast to that assumption, I suggest that Kant’s mature conception can in fact account for such cases,

even if differently than contemporary anti-deceptionism. An “error *in alio*”, as Kant puts it in the Vigilantius lectures, or even the mere intent to cause a false belief, is here not the issue. A bald-faced lie like that of the Cheating Student is blameworthy, even if harmless, because it is an intentionally untrue expression of his thoughts, in spite of not including any realistic intent to deceive. Importantly, such cases are not even about beliefs or information, since they reveal a mere unwillingness to manifest one’s awareness of one’s own acts. The Inquiring Thief case is not different, in this respect, as Kant stresses that the one condition for an untrue statement to be a lie is that it is taken to be a declaration of one’s mind.¹⁵

In the terms of the lectures from the 1770s, the main requirement for a statement to be a possible lie is particularly demanding, since it should be explicit that the statement has to be taken as a genuine declaration. (“Not every untruth is a lie; it is so only if there is an express declaration of my willingness to inform the other of my thought [*seinen Sinn zu verstehen geben*]”; Kaehler 329; cf. 27: 448.)¹⁶ The unrealistic insistence on this prior quasi-contractual agreement of sorts is not present in the same terms in later texts.¹⁷ Still, the general point remains very much in place. Someone lies when his statement presents his thoughts untruthfully. This is what Kant specifically calls declaration, which is the central notion in his official definitions of lying, quoted before.¹⁸ An intentionally untrue expression of one’s thoughts, however, does not have to aim at deceiving anyone.

Lying is thus not about deceiving but first and foremost about an intentional misrepresentation of the speaker’s mind. This is exactly what bald-faced liars do. Their lies are a good showcase for the self-regarding nature of lying that Kant highlights in his ethics as its key moral feature. Kant is reported to have observed already in lectures from the 1770s that lie “is more of a violation of duty to oneself than to others” (Kaehler 172; cf. 27:341; see also 27:604; 11:332). Then his considered view in the *Doctrine of Virtue* centers on the thought that lying is exactly such a violation, in that it belongs to the ways of conducts that “make it one’s basic principle to have no basic principle . . . , that is, . . . make oneself an object of contempt” (6:420; see already Kaehler 172; cf. 27:341). The intentional untruth of a declaration infringes in the moral status of the liar because of his unwillingness to express his own thoughts. Bald-faced lies are exactly cases in which, independently from other considerations, someone “makes himself an object of contempt . . . in his own eyes” (6:430; cf. 27:700), thereby incurring the reproach of “worthlessness” (8:426; cf. 6:403) Kant considers distinctive of lying. If the morally crucial feature in any lie is how it affects the liar and his moral capacity, the intent to deceive, or even to conceal information, cannot be regarded as the defining trait of lying. Rather than from an intent to deceive and thereby harm another person, lies stem primarily from the unwillingness to express one’s thoughts truthfully. Rather than disregard for the truth of another’s beliefs, they show disregard for

truthfulness in declarations. The impossibility of deceiving anyone, which is distinctive of bald-faced lies, does thus not prevent the possibility of lying, in Kant's view. On the contrary, it can account for such lies as noteworthy examples of the central feature of lying in general.

This strand of Kant's view is especially apparent in his ethical theory, that is, in his treatment of lying as an ethical wrong. What about the juridical domain, however? Should we infer that Kant's moves away from the traditional deceptionism only in ethics but not in his examination of juridically relevant lies? As I have mentioned, however, the discussion of the *Supposed Right to Lie* essay is based on the same definition as the *Doctrine of Virtue*, in which there is no mention of an intent to deceive as a necessary condition for treating a statement as a lie. That a lie is juridically wrong in that special sense, even if it does not harm anyone, because it violates "humanity in general", implies that here, too, its wrongness is not construed in deception-related terms. A declaration is expected to be a truthful communication of one's thoughts.

Notably, Kant's take on the issue entails a rejection of the claim that deceitful lies and bald-faced lies are morally wrong for different reasons.¹⁹ To the contrary, Kant's view accounts for the wrongness in both deceitful and non-deceitful lies, which, for him, share a feature that makes them blameworthy. In this respect, the difference between Kant's normative approach to the definition of lying and the contemporary descriptive approach begins to show. I shall come back to this important contrast in § 5.

Contemporary anti-deceptionists hold that the attempts to construe all lying as cases of deception are ultimately motivated by "the worry that if lies are not always intended to deceive, it is difficult to explain th[eir] *prima facie* wrongness" (Fallis 2015, 82). In contrast, Kant's view on lying is motivated by the aim to acknowledge both that (1) lies do not have to be deceitful and, in this respect, harmful to other people, even if only to their epistemic stance, and that (2) lies are nevertheless always morally wrong, independently from further considerations. Such a view is thus a version of intrinsic anti-deceptionism, that is, a view maintaining that not only lying does not necessarily involve the intention to deceive but which also holds that "lying itself is a morally relevant factor, or as we might put it, lying itself is a moral wrong-maker".²⁰

4. The peculiar gravity of lying as dishonesty

A recurring topic in the recent discussion is how lying and deception compare as to their moral worth. If their relation is a key to understanding them, how would this affect their normative status? In this respect, the supposed traditional view that lying is morally worse than misleading as a specific kind of deception is taken to be a Kantian idea. It is a currently widespread assumption that "[p]hilosophers who endorse the perceived moral distinction between lying and misleading almost uniformly derive their view from a particular strand in Kant's thought" (Berstler 2019, 11f.).

The alleged “Kantian idea” has recently been rejected by arguing that lying and misleading can differ either as to their outcome or the “method” they use (see Saul 2012, 69ff.). Since neither of them clearly separates lies and cases of misleading, a clear-cut difference between them should be discarded. As to Kant’s view, which is assumed to be the paradigmatic version of the traditional thought, the difference cannot be based on the outcomes, intended or actual, of the acts at issue. “The moral status of any particular deception depends on such things as its goal or its consequences” (Saul 2012, 99) only if we are willing to embrace a broadly consequentialist conception, which cannot be ascribed to Kant. In this picture, then, the feature of lying that makes it a distinctive wrong, worse than misleading acts, should be, for Kant, the use of language in assertions. Lying would be considered worse than misleading because it comes down to a misuse of the crucial function of language, which has to be sanctioned as such. In the same spirit, Bernard Williams accordingly sees in Kant a “fetishization of explicit assertion” (2002, 81), since it seems that the salient distinction is between saying vs not saying. Also, those who, against Saul and others, endorse the alleged Kantian idea do it with regard to the use of language, maintaining that “conventional language matters, morally speaking” (Berstler 2019, 331). It is usually assumed that in a view “commonly attributed to Kant”, “falsely asserting degrades the practice of assertion. If even one person lies, she harms everyone else’s ability to assert. If everyone lies, our capacity for assertion disappears altogether” (Berstler 2019, 17).

Kant, however, does *not* hold that lying is worse than deception because of its nature of assertion. In other terms, he does not understand lying as deception plus assertion, as it were. Kant does not seem to be prone to the “fetishization” that Williams points out, since what is important in a declaration, for Kant, is not that it is a speech act and that its misuse causes the common “capacity for assertion” to “disappear altogether”. A declaration matters, instead, insofar as it provides a representation of the self as a subject that presents himself through his genuine thoughts. A declaration is thus morally salient as the manifestation of one’s mind. In fact, nothing in Kant’s observations on lying appears to commit him to exclude the possibility of even non-linguistic declarations. The distinctive feature of a declaration is its presumed task of presenting one’s thoughts in a shape that is assumed to be truthful. This task is most commonly performed through language, for sure, but it does not have to be so. Even the one passage in Kant’s mature writings in which he seems to come closer to presenting language as the distinctive ground of lying does nothing more than emphasize that any moral subject is “bound . . . to the condition of the agreement of the declaration” of himself as a moral subject (6:430).²¹ Thus the linguistic character of lying does not play a decisive role in Kant’s considered view. Language might prove to be the best means to that end, if only because, as Kant suggests at some point, thinking is deeply informed by language (see e.g. 7:192), but

it is not the only means available.²² The peculiar gravity of lying does not derive from its being an essentially linguistic act.

In Kant's view, thus the difference between the two ways of acting, lying and misleading, does not concern what Saul calls the "method", that is, how the wrong is done. It is, rather, a different kind of wrong. Misleading and lying might share some descriptive features. From the standpoint of Kant's moral theory, however, their normative status makes an important difference apparent. Compared with misleading, lying is an altogether different activity, which is intentionally aimed at others and may well include an intent to deceive them. Nevertheless, its distinctive wrong consists in injuring the subject's own capacity to act morally. The difference from misleading resides, thus, not so much in features that can be discovered in a descriptive investigation but in a different normative status. (I shall come back to this important underlying difference from contemporary debates in § 5.)

Jennifer Saul argues, furthermore, that the traditional thought that misleading is generally better than lying has to do with the fact that "when we consider the morality of particular acts, as we do when presented with cases of lying and misleading, we actually think about more than just the morality of the acts", that is, about "the virtuousness of the *actor*" (2012, 86). Saul thereby comes closer to Kant's own perspective on lying. What matters in lying and makes it wrong, is, for Kant, not the consequences of the corresponding acts but that it determines a trait of the liar that affects his moral status. Kant's central claim that the wrong of lying consists in its self-regarding character is here crucial. Although one lies to another person,²³ lying is nevertheless construed by Kant as "the greatest violation of a human being's duty to himself regarded merely as a moral being (the humanity in his own person)" (6:429).²⁴ This is why, for Kant, lying is deeply wrong, in a way that is specifically different from that in which deceiving or misleading is wrong. Since lying affects the moral standing of the liar, it represents a violation of one of the duties to oneself, which, in Kant's view, enjoy a specific priority over other obligations.

The thought that the "method" of lying, that is, verbal statements, cannot be the ground of its wrongness leads some writers to argue that "lying is not a distinct moral category" (Barber 2020, 8), which can in turn lead one to suggest that the morally relevant features of the corresponding acts should be framed differently than within the current alternative between deceptionist and anti-deceptionist accounts. This happens in a recent proposal, which is helpful to consider in comparison with Kant's view, since this new angle to the contemporary discussion provides a further opportunity to characterize Kant's notion more precisely.

A limit of describing a view as anti-deceptionism is that the label only expresses that that view does not reduce lying to deception or deceitful intent and misses how it positively characterizes lying. With regard to the contemporary discussion, it might thus seem that the debate should move

past the focus on lying, deception and misleading. As has been recently proposed, one way to do this would be to shift the perspective and look at cases of lying through the lens provided by the broader notion of dishonesty. Whereas the discussion concentrates “on the boundary between saying and merely intimating, insinuating, etc.”, a more productive examination should focus instead “on the boundary between what is and is not expressed in a communicative act, irrespective of whether it is actually said” (Barber 2020, 2). Focusing on dishonesty entails that the alleged prejudice in favor of language would be overcome but also that lying should be regarded as a derivative, secondary notion. Alex Barber thus defines dishonesty as “expressing that p when one knows p to be untrue” (2020, 12). The main feature of this alternative construal of untruthful expressive acts is that they would then not be about deceiving, as dishonesty is “not equivalent to seeking to cause someone to believe that p when one knows p to be untrue” (Barber 2020, 13). In Barber’s account, lying as well as other dishonest acts would be characterized by the distinctive wrong of consisting in an “abuse of communication”. Thus, we would have to “include expressive meaning as well as the literal use of language” (Barber 2020, 17).

Barber follows Saul’s critical references to Kant and his supposed proneness to a fetishizing idea of language as the only means of morally relevant communication. In fact, however, Kant frames the issue analogously to the dishonesty-centered suggestion, with an important difference. His genuine view differs not because it holds to a merely assertive notion of lying, since it is in fact able to acknowledge as lies non-assertive, maybe even non-linguistic, declarations. Kant’s view does not explain the wrongness of lying through an abuse of communication but in purely moral terms. What is abused is the moral standing and, consequently, the participation in relationships between moral subjects. Untruthful declarations prevent the liar from legitimately presenting himself as a subject with a moral standing. Barber remarks that the dishonest “will in the past have falsely represented herself as trustworthy” (2020, 18). Analogously, the Kantian liar expresses untruth under the shared tacit assumption that what he says represents his thoughts, which corresponds to the oddly unrealistic condition of an antecedent pact in the lectures from the 1770s that I have mentioned before. That is in fact the condition that is built in the notion of declaration, in Kant’s specific sense.

The definition of these closely related words has to be stipulative at some point if they are to serve as univocal technical terms, as both Kant and the contemporary discussion intend. Barber presents dishonesty as a characteristic of acts, whereas the term more often denotes a quality of agents, as opposed to “insincerity”, which is then used to mean a property of assertions (see e.g. Stokke 2018). More explicitly, others present dishonesty as a character trait (see Carson 2010, 257–265). For Kant, talking of dishonesty is clearly about a quality of the actor instead of the act. In fact, Kant uses the term (or a German term that it is plausibly rendered as “dishonesty”) in

a crucial passage to denote exactly the general moral quality that a subject determines for himself in truthfully communicating his thoughts:

Truthfulness in one's declarations is also called *honesty* [*Ehrlichkeit*] and, if the declarations are promises, *sincerity* [*Redlichkeit*] (6:429).

(Carson 2010, 21)

Here sincerity is not presented as a quality of truthful statements but as the moral quality of a person who refrains from a specific kind of lies, that is, "lies whereby the other is cheated" (Kaehler 330; cf. 27:449), which corresponds to the lying promise discussed in the *Groundwork* (4:403, 419). Honesty, and its vicious counterpart dishonesty, are thus more basic moral qualities that are determined by the truthfulness, or untruthfulness, of one's declarations. Kant's account of lying can thus be regarded as a version of intrinsic anti-deceptionism that regards dishonesty as the distinctive feature of lies and liars.

5. Defining lying by normative principles: concluding remarks on method

The assumption that Kant's view should be assimilated to the deceptionism of the traditional conception, analogously to the other classical authors, proves not merely inadequate with regard to taxonomy. More importantly, it prevents understanding Kant's remarks enough to allow a dialogue with contemporary perspectives on the issue. Although some tension between deceptionist and anti-deceptionist elements is still detectable in Kant's remarks on lying, the most distinctive features of his account bring him in the anti-deceptionist camp. In light of the terms of current debates, thus, I have suggested that the view put forward by Kant can be understood as a variant of intrinsic anti-deceptionism, because it does not construe lying in terms of an intent to deceive, while maintaining that lying is wrong independently from any other considerations. The claim that lying is morally worse than misleading is also based, in his account, on the peculiar self-regarding character of lying, which consists, for Kant, in disregarding the moral necessity of a truthful expression of one's thoughts. This idea suggests, finally, that in Kant's view, lying can be positively characterized as dishonesty.

The comparison with the terms of the contemporary discussion, however, is also helpful to highlight some notable methodological features of Kant's approach to this issue, which provides a good example of his way to develop a moral theory, in contrast with currently widespread methodological assumptions. In the contemporary discussion on the definition of lying, the different positions unfold mainly from a consideration of particular examples. A concurrent account is rejected as inadequate because it "fail[s] to count as lies

assertions that clearly are” (Lackey 2013, 245). An unsatisfactory account is either too broad or too narrow, because it either equates to lying statements of a different sort or proves unable to cover some examples, which are supposed to be uncontroversial cases of lying. Most of the sophisticated recent debate is thus about how to account for the features that are apparent in those examples. Accordingly, the discussion relies heavily on intuitions (“That’s clearly lying!”)²⁵ or on the ordinary use of language,²⁶ with the primary aim to formulate definitions that are able to accommodate them. Contemporary philosophers appeal to empirical evidence, observing, for instance, that “there is empirical evidence that most people are disposed to count such statements as lies” (Fallis 2009, 42, fn. 47), in order to justify that the salient features of the corresponding cases are to be accommodated by a satisfactory definition.²⁷ Although this strategy is related to the interdisciplinary interest in a descriptively adequate account of lying, the methodological background and the problems related to an intuition-based approach are usually not discussed.²⁸

Kant does *not* develop his view in a similar way. His main interest in the topic is not descriptive but normative. The aim of his examination of lying is not to sketch a definition able to cover all particular cases but to explain the moral significance of lying. Kant regards lying as an inherently normative notion, which must thus be treated accordingly. Here an important difference emerges. In general, Kant’s moral theory does not start off from intuitions about particular cases. In Kant’s view, such a method would amount to trying to infer normative standards from examples, which is not a viable option, as he argues in the *Groundwork* (see 4:408f.).²⁹ Instead of referring to intuitions about what a lie is, thus, Kant examines lying on the basis of normative standards.

A normatively minded account of lying can be developed in at least two different ways: either by taking the normative reaction to which lying is subject (e.g. blame) as the starting point from which lying is construed or by focusing on an obligation in order to determine which kind of conduct violates it.³⁰ Kant takes the second path. In the lecture notes from an early course, Kant is reported to have remarked about the white lie: “it is an *untruth* that breaches no *obligation*, and thus is properly no *lie*” (27:62). The remark stems from the early 1760s and pre-dates many significant developments in Kant’s thought, also specifically concerning his account of lying. Still, it provides an apt phrasing for the distinctively normative take on the subject that Kant followed on in the next decades. From this standpoint, the defining feature of lying is that it is a violation of an obligation. Thus it must be defined not by providing an accurate description of the various instances of what is usually called lying but by clarifying what obligation is violated by lying, and in which respect. A normative approach to the definition yields the definition via a clarification of the wrong-making feature of lying. Note that also Kant’s emphasis on the distinction between ethical and juridical considerations of lying follows from this normative

approach. Maintaining that “Kant does not give a single answer to the question of what is a lie”, because “he tells us what a lie is in the ethical sense, what a lie is in the juristic sense, and what a lie is in the sense of right” (Mahon 2009, 209),³¹ is thus somehow misleading. Kant’s view does not aim at generating multiple notions of lying, as if the descriptive difficulties in the traditional approaches could be overcome simply by substituting three different notions for the concept at issue. Still, the care to distinguish between different perspectives is necessary, as it reflects Kant’s distinctive approach. The two perspectives have to be distinguished (even if not separated) because they follow from the different application of moral laws in the ethical or juridical appraisal.

Kant famously held that philosophical investigation cannot, and should not, start off from providing definitions: “in philosophy the definition, as distinctness made precise, must conclude rather than begin the work” (KrV A731/B759). This general thought is relevant here as well. In fact, it helps to shed light on a significant fundamental difference between Kant’s view and the contemporary debate. A descriptive approach, which focuses on particular instances, “can often infer much from some marks that we have drawn from an as yet uncompleted analysis before we have arrived at a complete exposition” (KrV A730/B758) of the concept of lying. It cannot include all relevant traits, though. Accordingly, a proper definition of lying, for Kant, cannot be the starting point for an investigation into the morality of lying, but it should rather be the outcome of a philosophical examination that has first clarified the standards of moral obligation. Only once it has been made clear which obligation applies to cases of what is usually called lying is it possible to provide a proper definition of what lying is.

Some writers suggest that Kant belongs to a long tradition in which lying is defined very strictly, so that many deceptions cannot count as lies and many false statements can be considered permissible. An absolutist take on lying would thereby be made consistent, in spite of the most common linguistic use, which, the other way around, displays a broader understanding of lying, combined with the willingness to make room for permissible sorts of lies.³² This way of framing Kant’s approach, however, obscures a crucial point, namely that, in his perspective, there is no other starting point than a normative principle. It is not the case that a strict definition is given in order to easily accommodate the conditions posed by an absolute prohibition. Rather, Kant cannot but provide a specific definition of lying, since the definition follows from the morally relevant traits to which principles give significance. The aim of this approach is not to protect an arbitrary absolutist assumption but to account for the proper significance of lying. If lying is first and foremost a normative notion, why should we consider cases in which the notion does not apply? There is nothing artificial in examining lying through the lens provided by moral principles if the moral meaning of the notion is ultimately the relevant one.

Unlike contemporary discussions on the definition of lying, thus, Kant's normative approach entails a (mildly?) revisionist perspective on the ordinary understanding of what a lie is. Here a distinctive trait of Kant's moral philosophy overall becomes apparent. Its general aim is to clarify the standards of moral appraisal that are available to ordinary moral thinking without necessarily confirming the consolidated outcomes of the ordinary application of those standards. By rectifying the notion of what makes lying morally wrong, Kant's moral theory eventually provides a definition of lying that may partially revise the ordinary understanding of what a lie is. Such a revision is possible because the examination draws not on the features of empirical instances but on the underlying standards through which a statement is judged to be a lie.³³

Notes

- 1 The state of the art is best represented by the recent *Oxford Handbook of Lying* (Meibauer 2018b). An helpful general survey is provided by Meibauer 2018c and, more specifically, Mahon 2018b. Mahon even suggests that “today, more philosophers than ever before are working on the subject of lying” (Mahon 2018b, 32).
- 2 See, for example, Meibauer 2018a, 334. What counts as the classical view on lying is summarized in Mahon 2018a.
- 3 As James Mahon has pointed out: “More than any other element of his moral philosophy, Kant's writings on lies have elicited an unprecedented amount of abuse” (2009, 201). The remark still holds true ten years later.
- 4 One partial exception is Carson 2010, 67ff.
- 5 All references to Kant's writings are given by volume and page number of the Academy Edition. For the *Ethik Kaebler*, I follow the edition provided in Kant, *Vorlesung zur Moralphilosophie*. Ed. Werner Stark. Berlin – New York: De Gruyter 2004. The English translation of the quotations is taken from the Cambridge Edition of the Work of Immanuel Kant, where available, and, for the *Groundwork for the Metaphysics of Morals*, from Jens Timmermann's revision (Cambridge University Press, 2011).
- 6 This is the reading of Mahon (2009, 207), who in general holds that for a statement to be a lie, in Kant's view, “it must be intended that the untruthful statement be believed to be true” (2009, 207).
- 7 Here I cannot comment on Kant's changing usage of the Grotian distinction between lie (*mendacium*) and *falsiloquium*, which is a crucial aspect of Kant's work on a more precise determination of the notion of lying. On this, see Timmermann, forthcoming.
- 8 See, for example, also Bok, 1978, 21–22.
- 9 See Mahon 2016, § 3. Positions differ on the matter only with respect to further conditions and qualifications concerning how a false belief is caused in others.
- 10 A general disregard for Kant's (admittedly difficult, even perplexing) account in the *Doctrine of Virtue* is widespread in the current discussion, in which the lecture notes and the notorious essay on the *Supposed Right to Lie* are often considered the only relevant sources of Kant's thought, whereas the *Doctrine of Virtue* is most often entirely ignored. See, for example, Fallis 2009. A partial exception in this regard is Williams 2002, 106f., in addition to Mahon 2009.

- 11 I have modified the Cambridge Edition translation, following Timmermann (forthcoming), § 3.
- 12 See Sorensen 2007; Fallis 2009; Carson 2010, 2018; Lackey 2013; Fallis 2015; Mahon 2016; Berstler 2019.
- 13 The Inquiring Thief case is briefly mentioned as an example of non-deceptive lie in Fallis 2009, 43 fn. 48. Fallis, however, does not discuss the complication that the statement is given under duress, which Kant there treats as morally salient.
- 14 See Sorensen 2007, 263: “The plausibility of a strong condemnation of lying is normally protected with a narrow definition of ‘lie’. Since no bald-faced lie involves the intent to deceive, I suspect Kant and Ross would regard the bald-faced lie as no more a lie than metaphor, hyperbole, or sarcasm.”
- 15 Berstler (2019, 29) presents another instance of bald-faced lie, which can be helpful to further clarify the matter: “suppose that a defendant, Jane, is testifying at a trial. The opposing counsel has just played a videotape of Jane robbing a bank. Everybody knows that Jane is in the video, and everybody knows everybody knows this (and so on). Nonetheless, Jane’s attorney has cautioned her to admit to nothing on the stand. When the opposing counsel asks Jane whether she can identify the woman in the video, Jane says that she can’t.” Here the juridical setting is relevant and might be confusing, since Jane’s lie must also be regarded as false testimony. Nevertheless, the general point is clear enough and applies outside of trial situations as well: Jane is unwilling to be truthful about herself, although it is impossible to deceive or conceal information about her acts.
- 16 Note that the same phrase, “zu verstehen geben”, occurs in the lectures on natural right from the 1780s: “If I imply something to him [*gebe ich ihm etwas zu verstehen*], but I mean something else by that [*verstehe was andres darunter*] then it is *falsiloquium*” (27:1340). The point is obscured, I believe, in the Cambridge Edition translation, which reads: “If I get him to understand me about something one way but I understand it another way”. Kant’s remark is not about succeeding in causing a false belief in another but intimating something different from one’s thoughts, that is, “declarare mentem suam”, per the Latin precept that Kant is commenting on with that remark.
- 17 For a more extended assessment of this point, and the view presented in the Kaehler lectures in general, see Timmermann (forthcoming), § 3.
- 18 The crucial role of the term declaration in Kant’s view has been emphasized by Wood 2008, 240ff.
- 19 See Sorensen 2007, 263: “You have good reasons to refrain from bald-faced lying but these are not the moral reasons that condemn disguised lies”.
- 20 I borrow the label from Stokke 2019, 331.
- 21 For further comments on this passage, see Bacin 2013, 249f.
- 22 Here I find myself in disagreement with Mahon 2009, 203f.
- 23 I leave aside here the issue of the supposed lies to oneself, which requires deeper examination.
- 24 On this important claim, on which here I cannot comment further, see Bacin 2013.
- 25 See, for example, Fallis 2009, 33, 43, 46, 51, 53; Lackey 2013, 238.
- 26 See, for example, Lackey 2013, 238 fn. 7; Williams 2002, 96 f.
- 27 On the state of the empirical studies on lying, see Wiegmann and Meibauer 2019.
- 28 To the best of my knowledge, the only exception to this general trend is Carson, who discusses the problematic status of appealing to intuitions and rejects that approach (see 2010, chap. 6). This might be related to the fact that Carson’s primary interest goes to the ethical dimension of lying, not to a descriptively

adequate taxonomy. Saul takes also a critical stance towards intuitions in her discussion of the lying vs misleading preference; see 2012, 70 f. For further critical remarks on the appeal to intuitions in the discussions on the definition of lying and misleading, see Cappelen and Dever 2019, 40f. (Thanks to Stefano Lo Re for referring me to Cappelen and Dever's book.)

- 29 A full clarification of the grounds and the extent of the difference between the method of Kant's moral philosophy and the contemporary appeal to intuitions requires a broader examination, which I shall present in a separate paper.
- 30 In the recent discussion, a normative approach to the definition of lying can be found in Cuneo 2014, whose approach follows the first option. Pallikkathayil (2019) is very close to taking the other path, as she distinguishes, in a broadly Kantian fashion, a duty not to deceive from a duty not to lie but then argues that "it seems both possible and appropriate to conduct our moral inquiry without settling the definition of lying".
- 31 Mahon rightly observes that the three senses of lying correspond to three duties not to lie. He, however, does not further investigate the grounds for introducing the distinction.
- 32 See, for example, Sorensen 2007, 263; Saul 2012, 69.
- 33 I should like to thank Jens Timmermann for many conversations on this topic through different times and places. Lately, the discussion in a workshop on his manuscript on *Kant's Supposed Right to Lie*, held in St Andrews in December 2019, has been extremely helpful to me. To Jens's manuscript I owe many important clarifications, even beyond what I have made explicit in the footnotes here. Thanks also to Sorin Baiasu and James Mahon for discussion at the workshop and to Stefano Lo Re for very helpful remarks on this chapter.

References

- Bacin, S. 2013. "The Perfect Duty to Oneself Merely as a Moral Being (TL 6:428–437)." In A. Trampota, O. Sensen, and J. Timmermann (eds.), *Kant's "Tugendlehre". A Comprehensive Commentary*. Berlin: De Gruyter, 245–268.
- Barber, A. 2020. "Lying, Misleading, and Dishonesty." *The Journal of Ethics* 24(2): 141–164.
- Berstler, S. 2019. "What's the Good of Language? On the Moral Distinction between Lying and Misleading." *Ethics* 130(1): 5–31.
- Bok, S. 1978. *Lying: Moral Choice in Public and Private Life*. Hassocks: Harvester Press.
- Cappelen, H., and Dever, J. 2019. *Bad Language*. Oxford, New York: Oxford University Press.
- Carson, T.L. 2010. *Lying and Deception: Theory and Practice*. Oxford: Oxford University Press.
- Carson, T.L. 2018. "Lying and Ethics." In J. Meibauer (ed.), *The Oxford Handbook of Lying*. Oxford: Oxford University Press, 469–482.
- Cuneo, T. 2014. *Speech and Morality: On the Metaethical Implications of Speaking*. Oxford: Oxford University Press.
- Fallis, D. 2009. "What Is Lying?" *Journal of Philosophy* 106(1): 29–56.
- Fallis, D. 2015. "Are Bald-Faced Lies Deceptive after All?" *Ratio* 28(1): 81–96.
- Lackey, J. 2013. "Lies and Deception: An Unhappy Divorce." *Analysis* 73(2): 236–248.

- Mahon, J.E. 2009. "The Truth about Kant on Lies." In M. Clancy (ed.), *The Philosophy of Deception*. Oxford: Oxford University Press, 201–224.
- Mahon, J.E. 2016. "The Definition of Lying and Deception." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), <https://plato.stanford.edu/archives/win2016/entries/lying-definition/> (accessed on February 22, 2021).
- Mahon, J.E. 2018a. "Classic Philosophical Approaches to Lying and Deception." In J. Meibauer (ed.), *The Oxford Handbook of Lying*. Oxford: Oxford University Press, 13–31.
- Mahon, J.E. 2018b. "Contemporary Approaches to the Philosophy of Lying." In J. Meibauer (ed.), *The Oxford Handbook of Lying*. Oxford: Oxford University Press, 32–55.
- Meibauer, J. 2018a. "Lügen." In F. Liedtke and A. Tuchen (eds.), *Handbuch Pragmatik*. Stuttgart: J.B. Metzler, 334–341.
- Meibauer, J. (ed.). 2018b. *The Oxford Handbook of Lying*. New York: Oxford University Press.
- Meibauer, J. 2018c. "Introduction." In J. Meibauer (ed.), *The Oxford Handbook of Lying*. Oxford: Oxford University Press, 1–10.
- Pallikkathayil, J. 2019. "The Truth about Deception." *Philosophy and Phenomenological Research* 98(1): 147–166.
- Saul, J.M. 2012. *Lying, Misleading, and What Is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press.
- Sorensen, R. 2007. "Bald-Faced Lies! Lying without the Intent to Deceive." *Pacific Philosophical Quarterly* 88(2): 251–264.
- Stokke, A. 2018. *Lying and Insincerity*. Oxford: Oxford University Press.
- Stokke, A. 2019. "Lies, Harm, and Practical Interests." *Philosophy and Phenomenological Research* 98(2): 329–345.
- Timmermann, J. forthcoming. *Kant's Right to Lie*.
- Wiegmann, A., and Meibauer, J. 2019. "The Folk Concept of Lying." *Philosophy Compass* 14(8): e12620. DOI: 10.1111/phc3.12620.
- Williams, B. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.
- Wood, A.W. 2008. *Kantian Ethics*. Cambridge; New York: Cambridge University Press.

THE DUTY AND THE MAXIMS

Elements for a morality and culture of sustainable development

Anselmo Aportone

1. From the UN Agenda to the “duties of virtue” of Kant

On September 25, 2015, the United Nations General Assembly adopted the resolution “Transforming our world: the 2030 Agenda for sustainable development”.¹ It is a program of action based on the general principles adopted by the United Nations² for people and the planet in the name of prosperity, peace and global collaboration. It identifies objectives to pursue to make the economic, social and environmental dynamics in our world sustainable. Whatever its real prescriptive power,³ it is a solemn declaration of a commitment to overcome an idea of progress marked by mere economic growth (measured at least unilaterally through gross domestic product) and therefore to consider the synergistic and systemic relationships of the three aforementioned dimensions: from the perspective of the whole of humanity, a development is sustainable in the long run, perhaps already in the medium term, only if it produces general environmental, social and economic well-being and the prospect of leaving future generations a quality of life at least not inferior to the current one in an intact ecosystem, which ultimately is also the least risky and the richest in opportunities.

The assembly that brings together the representatives of (almost) all countries of the world, which should therefore represent the common interest of humanity, has long recognized⁴ and, with the 2030 agenda, reaffirmed with greater strength and ambition that – above all in a fully globalized world and in the era of the so-called anthropocene with its climatic emergency⁵ – the aforementioned sustainable development is, in the idea and the realization, the condition and result of moral respect, of cultural growth and perhaps even of the physical survival of humanity in the near future. It therefore calls for rapid assumptions of responsibility and a concrete commitment to direct the action according to conscious and adequate maxims. The idea of global sustainable development as a general purpose of the United Nations in fact places objective constraints on the power of choice of all the particular

subjects that concretely shape the relationships between human beings and between them and nature. The ends to be pursued are to be considered parts of a system of means favorable or contrary to the achievement of the most comprehensive aim of sustainable development, through a conscious and competent evaluation of the consequences of the actions.

On the one hand, starting from the heritage of the 20th century, the concept of sustainable global development exposes a general end, which should really be pursued by all people and institutional subjects and yet sufficiently determined and concrete. It draws its matter from particular interests, potentially and in fact conflicting, but from these, it emancipates itself as a rational determination of the will in view of an objective interest of the whole of humanity. In this way, the UN, an institution largely inspired by the Kantian idea of the *Völkerstaat* or *Völkerbund*,⁶ still echoes, up to the sphere of global governance of the planet, the practical philosophy⁷ and the wishes of the philosopher of Königsberg. This occurs through the solicitation to determine the political will according to a pattern that corresponds to the profound structure, informed by the moral law, of the process of determining the good will analyzed by Kant (on which we will dwell shortly). But it happens already in the invitation to grasp rationally the “opportunity” offered by the crisis of the current development model (“crisis” understood, in the etymological sense of the κρίσις, as a decisive phase of a disease but also as discernment and choice that can be transformed into the necessary prerequisite for improvement). The serious risks impending for humanity due to a one-sided development (determined by a prevalent instrumental use of reason for the unruly exploitation of a geoid of limited size and resources) could perhaps empower the *pure* practical reason (the one which is not conditioned by selfish motives) so that, directing development to the common good, it can realize itself in the direction of man’s moral destination.

For, the very opposition of inclinations to one another, from which evil arises, furnishes reason a free play to subjugate them all and, in place of evil, which destroys itself, to establish the rule of good, which, once it exists, continues to maintain itself of its own accord.
(TP 08: 312).

On the other hand, precisely by virtue of its rational character (in the sense mentioned previously of a harmonization of particular ends according to an *objective* general interest of humanity), this end unifies (at least in concept) the will of the particular subjects from the perspective of an empirical synthesis of the variety of actions according to objective rules. The corresponding *subjective* purpose is an increase in happiness in accordance with the obligation imposed by the normative principle of a good will, the will of humanity immanent in the will of people as they recognize themselves in principle as free and autonomous. If this is true, the idea of sustainable development contributes to constituting something similar

to a real historical consciousness of humanity as a subject and object that acts and suffers in exchange with nature. In fact, humanity in the person (*Menschheit*, understood as a set of fundamental dispositions and as an idea that gives them a horizon of possibilities and development) gives the latter her inalienable dignity, making it an end in itself, an individual who can never be deprived of her fundamental rights and reduced to a simple means, and thus determines, in a distributive sense, the qualitative unity of human beings on the moral and juridical level. Instead, humanity understood from the point of view of praxis (*Humanität*) as the real collective unity of a plural subject is based not only on the idea of *Menschheit* but also on the set of relationships between people, established by the actions performed in compliance with the principles which are consistent with the same idea. In a late handwritten note, Kant defines it as “reciprocal benevolence, combined with mutual respect” (1797, Refl. 1531, AA 15: 959).

The way of thinking characteristic of the union of good living with virtue in sodal intercourse is humanity. What matters here is not the degree of good living, since one person requires much, another little, depending on what seems to him to be necessary. Rather, what matters is only the kind of relationship whereby the inclination to good living is limited by the law of virtue.

(*Anthr.* § 88, 07: 277)

In other words, the human quality of the natural and social life of people does not depend univocally on the quantity and quality of the goods available to them but on the ethical quality of the maxims that determine the will and guide the action according to well-being and prosperity, to the satisfaction of needs, but from a general perspective resisting selfish inclinations and interests. This means choosing ends compatible with everyone’s freedom and well-being, in short, those that contribute to the common good and sustainable well-being. The ends, whose realization significantly conditions the development of society and the environment by promoting human relationships, are not arbitrary or merely subjective but possess the generality and necessity of duties, which oblige the moral conscience of people and can therefore be defined as “duties of virtue”, that is, “an end that is also a duty” (MSTL 06: 383), if the virtue is defined as “the strength of a human being’s maxims in fulfilling his duty” (06: 394; vgl. 06: 405).

There are many different duties, corresponding to the different ends prescribed by the law, which are called duties of virtue (*officia honestatis*) just because they are subject only to free self-constraint, not constraint by other human beings, and because they determine an end that is also a duty.

(*ibid.*)

The Kantian conviction is therefore that humanity in relations between human beings requires the strength of virtue in pursuing the well-being of each one,⁸ in acting in a manner worthy of *Menschheit*, which is inherent in each person and obliges mutual respect, that is, without reducing anyone to a simple means for attaining one's goals, rather than settling down to this way to act in one's exclusive immediate interest (today it is evident that this happens in fact even when we do not care about the negative effects of our actions on the environment in which people live).

Kant evokes the strength of will because the form of "private" action (that which assumes the hypothetical-instrumental function of reason as determining in the context of human relations) actually shows a substantial weakness of the will, which originates in the misunderstanding of the unconditional value of the autonomy of one's own and others' will based on the original fact of the obligation towards the moral law, and therefore towards humanity, which every practical reason (every will worthy of the name)⁹ imposes on itself as the will of a sensitive as well as rational subject. Reason should govern human nature, to which it also belongs as its spontaneous component together with the receptive one, not only in view of the high moral destination but also because human action is not predetermined by nature. The latter places vague and uncertain internal and external limits on the first, and thus the human being needs an autonomous faculty of ends, one which is also able to set objective limits on the position of the ends and the relative choice of means.¹⁰

2. The choice of ends: from the moral law to the determination of the will

The moral law according to Kant notoriously does not dictate the ends but commands one to act freely and autonomously. It is the objective law of positive freedom; it is a law, as a norm that applies to all rational beings regardless of the empirical conditions of its application, and therefore in a general and necessary way, and pertains to the possibility that characterizes the will to determine itself according to an a priori form. This is the original nature of the will itself and allows it to exercise a free causality, that is, to be the first and autonomous condition of human actions. It establishes the possibility of emancipating one's behavior from determination through empirical conditions but does not finish with this negative freedom:

So act that the maxim of your will could always hold at the same time as a principle in a giving of universal law.

(KpV, 0 5: 30)

This law of self-determination cannot be derived from the objects of the will, whose concept does not rather presuppose universal legislation; it assumes

rather than that of a will that determines its intention through the representation of ends and implies the practical necessity of the means to achieve them. The moral law is the completely formal principle of pure practical reason, insofar as it constitutes the will, that is, the faculty of rational beings to determine themselves autonomously, which is actualized as free power of choice (*arbitrium*).

The power of choice is indeed an *arbitrium sensitivum*, yet not *brutum* but *liberum*, because sensibility does not render its action necessary, but in the human being there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses.

(KrV, A 534/B 562)

Consequently, the awareness of the moral law, which Kant recognizes as “fact of reason” (KpV, 05: 31), makes it possible for people to act (in the proper sense) according to maxims which are chosen (of course) in reference to what practical experience presents to them as concretely determinable, but freely and not by instinct or heteronomous determination. We therefore speak of the actions which (unlike the operation of “blind” causes that produce “mechanical” effects) can be classified as good or bad and involve the agent’s responsibility. At the same time, the moral law, as a principle that gives form to the will in general, prescribes the form of *human* action to the particular wills of the rational beings which are free and yet belong to the world of the laws of phenomenal nature (that always determine the external conditions and the empirical motives of action), commanding them to act according to maxims compatible with the universal and necessary form of the law, that is, with freedom understood positively and in an objective sense. The moral law as an unconditional formal principle with respect to particular motives and purposes, that is, to the necessary matter of actions,¹¹ prescribes therefore the autonomy or self-determination of the will that is sensibly situated, conditioning and qualifying the choice of ends and means. The autonomy of a finite subject cannot evidently be thought of as an unconditional position and realization of purposes but rather as a choice of ends in relation to “what ought to happen” (KrV, A 802/B 830) and therefore based on the principles of pure practical reason that orientate the volition towards the unity and systematicity of ends, that is, ultimately towards the self-realization of humanity. There are simple technical-practical rules under a problematic condition of the will, that is, particular purposes intended as hypothetical conditions of the action and therefore considered by putting in brackets the question of the effective rational deliberation or self-determination of the will in the specific case. The technical-practical rules concern only the relationship between possible ends and the means to achieve them.

Here, however, the rule says: one ought absolutely to proceed in a certain way. The practical rule is therefore unconditional and so is

represented a priori as a categorical practical proposition by which the will is objectively determined absolutely and immediately (by the practical rule itself, which accordingly is here a law). For, pure reason, *practical of itself*, is here immediately lawgiving. The will is thought as independent of empirical conditions and hence, as a pure will, as determined *by the mere form of law*, and this determining ground is regarded as the supreme condition of all maxims. (KpV, 05: 31)

If the will were simply a faculty to represent and pursue goals, it would be only an aspect of the theoretical use of reason, a technical capacity in a broad sense, useful for thinking about the means required by the desired end. It would therefore be determined heteronomously by the interests or by “the satisfaction that we combine with the representation of the existence of an object” (KU, 05: 204), and therefore from the objects that sources external to reason present to consciousness, such as passions. Its highest task could then be to make the objects of desire clear, also with respect to their consequences, and therefore to “rationalize” the passions by promoting consistency and agreement among them. This, although in perhaps too simple and crude terms, is Hume’s conception, according to which reason *cannot* be practical, since it has no causal power with respect to the choice and pursuit of ends, as is illustrated brilliantly in the following quote from the *Treatise*:

’Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. ’Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. ’Tis as little contrary to reason to prefer even my own acknowledg’d lesser good to my greater, and have a more ardent affection for the former than the latter. A trivial good may, from certain circumstances, produce a desire superior to what arises from the greatest and most valuable enjoyment; nor is there any thing more extraordinary in this, than in mechanics to see one pound weight raise up a hundred by the advantage of its situation. In short, a passion must be accompany’d with some false judgment, in order to its being unreasonable; and even then ’tis not the passion, properly speaking, which is unreasonable, but the judgment. (Hume 1896, 416)

A volition insofar as determined solely by a desire could never be considered irrational, if not – in a derived and in any case very weak sense – if the desirability of the object depended on a false judgment and would disappear following the rectification of the judgment. Conceptions of this type present, however, a serious problem: if one does not recognize a pure practical reason with an empirically unconditional principle (the moral law), not only

is the idea of freedom emptied of its substantial meaning, since the laws of nature alone would then govern the events, but also the notion of end (of practical intentionality) loses much of its meaning, since it is inseparably connected to rationality criteria.¹² In fact, desires and passions, if they are not freely accepted by the will through a maxim, do not properly constitute ends and remain on the level of the *arbitrium brutum* as data that certainly affect the faculty of desiring but do not belong to a *subject* and are not located in a horizon of meaning. All of this contradicts practical experience without offering another better understanding of the phenomena it presents (i.e. responses actually corresponding to the kind of problems that arise in the concrete determination of wills and in considering responsibilities). Every choice seems thus deprived of any objective justification.¹³

Leaving aside here the inquiry on theories informed by a highly speculative logic and remaining on the line of equilibrium of Kantian philosophy, it can be said that the notion of end does not refer properly to an object independent of reflection and will but is related to an effect “the representation of which is at the same time the determining ground of its production in an intelligently acting” (KU, 05: 426). It means a conformity of the contingent to rules or laws insofar it depends on a free intentional causality or is at least referred to it in reflection: we can speak properly of ends – even if only in an analogical and regulatory sense, such as when, for example, we refer to ends of nature – only if we assume (the idea of) a potentially free and conscious action, guided by representations of reason and maxims of the will, in other words, a practical reason that is determined in a rational, not in a random or hetero-directed, way by a pure principle and is therefore autonomous in the choice of its ends. “We must be free in order to be able to make use of our powers *purposively* in freedom” (*Religion within the Boundaries of Mere Reason*, 06: 188 Fn., italics added). On the basis of this assumption (or better, of the rational principle that justifies it, of which we have an immediate awareness, analogous to that of apperception, “the self-consciousness of a pure practical reason”, KpV, 05: 29) we can (1) conceive of ends, because the action of the free subject is mediated by a maxim that gives a rule to the action and determines it as the effect of a free causality; (2) think (through the principles of theoretical reason and of the faculty of judgment) of the same ends as elements of an aggregate or of a orderly and unitary system of ends and means; and finally (3) distinguish the different ends according to their conformity or non-compliance with an apical norm, valid for itself as an immediate expression of the pure practical reason (the moral law that commands us to assume itself as the principle and determining ground of the determinations of the will). This allows us to justify or prohibit the assumption of certain maxims, to evaluate the possible ends of the actions as good or bad with respect to the highest purpose of the human will and to consider a duty pursuing good ends according to the criterion that we find in our moral consciousness.

If that which is to be promoted as an end *through the human being's connection to nature* is to be found within the human being himself, then it must be either the kind of end that can be satisfied by the beneficence of nature itself, or it is the aptitude and skill for all sorts of ends for which he can use nature (external and internal). The first end of nature would be the happiness, the second the culture of the human being.

(KU, § 83, 05: 429–430, italics added)

Along these lines, through an analytical reflection, Kant defines happiness as the natural end of the human being, of a practical reason necessarily combined with nature, and culture all as sorts of means to pursue it, which in turn is to be considered a purpose of the free way of acting. It goes without saying that the development of happiness and culture are also the general aims that are historically articulated in the objectives of the 2030 agenda. The following quote instead synthetically formulates the general maxim through which these ends are taken as objects of the will or, more precisely, by means of which the will acting in the world must self-determine in accordance with the moral law.

The proposition, “Make the highest possible good in this world your own final end!”, is a synthetic proposition a priori which is introduced by the moral law itself, and yet through it practical reason reaches beyond the law. And this is possible because the moral law is taken with reference to the *characteristic, natural to the human being, of having to consider in every action, besides the law, also an end* (this characteristic of the human being makes him an object of experience). The proposition itself is possible (just like the theoretical yet synthetic propositions a priori) only because it contains the a priori principle of the cognition of the determining grounds of a power of free choice in experience in general, so far as experience, by exhibiting the effects of morality in its ends, gives an objective, although only practical, reality to the concept of morality as having causality in the world.

(*Religion*, 06: 6 Fn., italics added)

The imperative to pursue the good in the sensible world as much as possible (see KU, 05: 453) evidently implies the command to adopt a coherent system of maxims that is not in contradiction with the principle of legislation valid unconditionally for all human beings and therefore prescribes the compatibility of the maxims of all free rational subjects, so that their actions do not enter into conflict but are favored or at least not hindered from each other and promote as much as possible a solidarity context for the achievement of common goals and individual ends (through a rational

use of the available resources). Respecting the moral law in the choice of the maxims is therefore the way to pursue the highest possible good on earth. If this is correct, the initial proposition of the passage seems to be a simple corollary of the same categorical imperative (“so act that the maxim of your will could always hold at the same time as a principle in a giving of universal law”). Instead Kant presents it as a synthetic proposition a priori, which the practical reason derives from the formulation of the law referring it to a general quality of the sensitive nature of the human being. Since the concept of the latter is external to the immediate awareness of the moral law (proper to all rational beings, whatever their nature is otherwise), the proposition is synthetic. It is, however, also a priori since it derives its content from the possibility of human experience in general and not from any particular empirical experience; therefore, it claims to be valid for all human beings. Yet the moral law, as “fundamental law (*Grundgesetz*) of pure practical reason” (KpV, 05: 30), or a priori principle that constitutes a practical rational faculty and makes its use possible, *is* pure practical reason. The fact that it is present to our practical consciousness of *natural* rational beings in the imperative way (since the motives of our will are also pathologically determined, and therefore it is possible for us to determine the empirical will through maxims that are not compatible with the universality of the law, although the latter is and remains the general condition of the maxims) shows that, for us, the law has always already referred to the phenomenal nature of the human being. In what sense, then, does practical reason broaden itself beyond the moral law through an a priori synthesis of the latter with a property of human-sensitive nature?

The completely general characteristic – but specific and given together, although not unitarily, with the original consciousness of the law – to which reason refers in the passage from the first to the second a priori proposition (“a priori” understood in different ways as pure and impure, which allows us to qualify one proposition as law and the other as maxim) is the sensitive nature of the power of choice of the human being. It means that she must “*consider in every action, besides the law, also an end*”. To clarify this passage, it should be remembered that in later works, Kant makes a more distinct and consequent use of the terms “will” (*Wille, voluntas*) and “power of choice” (*Willkür, arbitrium*). The meaning of the first word is determined in close and immediate reference to the moral law:

The will is therefore the faculty of desire considered not so much in relation to action (as power of choice is) but rather in relation to the ground determining choice to action (*Bestimmungsgrund der Willkür*). The will itself, strictly speaking, has no determining ground; insofar as it can determine [the power of] choice, it is instead practical reason itself.

(MSTL, 06: 213, italics added).

The will can be called “higher” faculty of desire. The adjective means that it is the source of the pure principle or “*inneren Bestimmungsgrund*” of the will in general, since it does not refer directly to individual actions (and to the contingent ends through which we determine the “lower” faculty of desire, the empirical will) but prescribes the fundamental law for the choice of the maxims of the actions.¹⁴ In this sense, “the will, which is based on nothing other than the law, can neither be called free nor unfree, because it does not focus on actions, but directly on the legislation for the maxim of action (i.e. practical reason itself)” (MSTL, 06: 226). In other words, two elements can be distinguished in the faculty of desire. The first, the will, is defined as that for which the sensitive part, or the lower faculty of desire, is determined in the human being as free power of choice. This, as we have already seen, is in turn defined as the faculty of desire “in relation to the action” and – since “the determining ground of the same [faculty] for the act is encountered in itself, not in the object” – even more precisely as “a faculty to do or to refrain at one’s leisure . . . provided it is combined with the consciousness of the ability of its action to produce the object” (MSTL, 06: 213). The power of choice can therefore be said to be free if the faculty of desire is determined through concepts, or ends, which are contingent in themselves and yet conform to the form of action commanded by the moral law, so that the subject’s actions are not to be considered reactions to stimuli but executions of maxims which are freely formulated and adopted in the process of self-determination of the will.

This property of being able to act or refrain from action in relation to an object of the will is not analytically implicated in the proposition that expresses the law and can emerge only from the specification of the structure of human action as determined by a free will, that is, in general and under different respects by the laws of nature and by the moral law, but immediately by the end that he had to think and choose for himself and therefore by his own maxim. The concept of this common property must therefore be combined with the pure one of the moral law to unify the different elements of the faculty of desire that the philosopher’s analysis has distinguished to clarify their different status and function. On the basis of this synthesis, the practical reason formulates the maxim of the “higher” end that the power of choice of our experience must first of all make its own to be able then to orient itself in the choice of the maxims of the actions referred to the sensible world and to remain coherent with its own determining ground. The concept of free will defined thus exposes the idea of “morality” (*Sittlichkeit*), that is, of “the conformity (*Übereinstimmung*) of the maxim of an action with the law” (MSRL, 06: 225) not only in reference to the subjective moment of the assumption of the law by an empirical consciousness but also to the specific freedom of a rational *natural* being. Through it, the practical consciousness or reason actually “extends” itself with (1) the reference to the possible empirical matter of the will and (2) the definition of the way in which “the effects of morality” (*Religion*, 06: 6) are

manifested in the phenomenal practical experience: in the choice of ends or maxims. The same concept and the explanation of its possibility therefore give objective consistency to the causality of freedom in the sensitive world, although only from a practical point of view, it being understood that its theoretical status is and remains that of a transcendental idea.

3. The two maxims of happiness and the autonomy of practical reason

To underline its difference with the empirical general maxim pragmatically founded on the principle of happiness understood as the orientation of acting for the satisfaction of inclinations (see KrV, A 806/B 834), the previous general synthetic proposition of free power of choice could be called a morally qualified maxim of happiness. In fact, it is founded on the moral law, and its function is to direct the free choice towards a happiness that can be achieved first of all by making oneself worthy of it through a conduct informed by specific maxims that can be subsumed under the moral law. One might wonder, however, whether at the end of the day, the practical function of the two maxims is not equivalent and whether therefore the former is not to be preferred, because it presupposes a naturalistic structure of action, certainly simpler and devoid of the (both in theory and in practice) demanding assumptions of the latter. The price to pay would be the renunciation of our “natural” consideration of ourselves as free and rational beings according to a strong and challenging meaning, which would be a part of a folk psychology, one that is not adequate to the objective reality of things and therefore to be traced back to a more scientific theory of human nature. This is not the place to discuss this perspective and its problems, except to note the doubt that, if it is correct, and if it is true that – as emerges in a clear way from Kantian analysis – the notions of freedom and rationality implicate each other,¹⁵ then the eliminativistic scientific requirement itself could be unmotivated and devoid of practical motivations. In any case, Kant has a direct response to this objection.

The principle of happiness can indeed furnish maxims, but never such as would be fit for laws of the will, *even if universal happiness were made the object*. . . . Just because an object of choice is here put at the basis of its rule and must therefore precede it, the rule can be referred to and can be based upon nothing other than what one approves, and so it refers to and is based upon experience, and then the variety of judgment must be endless. This principle, therefore, does not prescribe the very same practical rules to all rational beings, even though the rules come under a common heading, namely that of happiness. The moral law, however, is thought as objectively necessary only because it is to hold for everyone having reason and will. (KpV, 05: 36, italics added)

A will founded on the pragmatic principle of happiness would be heteronomous and therefore incapable of thinking and a fortiori of realizing that the general unity of aims can be the object of progress towards a moral happiness (whose concept is determinable in practical perspective) which does not demand satisfaction of any particular inclination and interest (theoretically impossible and practically inconsistent) but the achievement of the greatest possible number of ends according to a general rational order. This progress actually seems to be the only “higher” purpose that can be rationally assumed by the will in general but particularly in the maxims of humanity’s development in the era of globalization. The real representation of the folk psychology to be overcome would therefore be not the immediate awareness of the moral law, which is placed on a qualitatively different level with respect to that of knowledge (whether it belongs to common sense or science), but rather the naturalistic concept of happiness, which correctly starts from the human condition, constantly affected by needs, but then operates a generalization that rests on an indefinite concept, on “a fluctuating idea” (GMS, 04: 399). By happiness is meant, in fact, the state of affairs of “satisfaction of all of our inclinations (*extensive*, with regard to their manifoldness, as well as *intensive*, with regard to degree, and also *protensive*, with regard to duration)” (KrV A 806/B 834), which is impossible to determine conceptually. Furthermore, even if it were, the realization of this purpose by a natural rational being would depend “on the harmony (*Übereinstimmung*) of nature with his whole end as well as with the essential determining ground of his will”, which is very problematic since he is certainly not the only determining cause of his natural world (KpV, 05: 124). In the aforementioned § 83 of the KU, Kant clarifies that actually

the concept of happiness is not one that the human being has, say, abstracted from his instincts and thus derived from the animality in himself; rather, it is a mere idea of a state to which he would make his instincts adequate under merely empirical conditions (which is impossible).

(KU, 05: 430)

In short, happiness is not the sum of the individual satisfactions of desires and needs, and therefore, in an empiristic sense, it cannot be anything real. It is not a concept that can be abstracted from phenomenal experience. It is a general and speculative notion that cannot be immediately reflected in the sensible sphere or an idea of practical reason, the indeterminate concept of lasting and necessarily general well-being (see KrV A 809/B 837). However, this idea is founded on the *rational* maxim of happiness, which, as we have seen, combines a priori two elements that are, on the contrary, well determined, the moral law and the sensible nature of human power of choice. Such a maxim can therefore determine the faculty of desire towards

a search for happiness according to the general law of freedom through the pursuing of an end which is both natural and objectively necessary but understood as the regulative ideal of the rational subject acting in the phenomenal world, an end that our free power of choice recognizes as a duty by determining itself in its aforementioned general higher maxim, according to the principle of morality, and not by following the inclinations *assumed as general principles of actions*, that is, according to the *pragmatic* maxim of happiness.

To sum up, a rational agent is a free agent and vice versa. He has a will, that is, the faculty of self-determination through concepts. That is, he formulates freely the maxims of his actions and is responsible for their consequences.

Thus the moral law expresses nothing other than the autonomy of pure practical reason, that is, freedom, and this is itself the formal condition of all maxims, under which alone they can accord with the supreme practical law.

(KpV, 05: 33)

Here the dual character of autonomy¹⁶ finds expression: On the one hand, as a formal condition, autonomy coincides with the theoretically “incomprehensible” freedom of will. It can be understood through a negative particular judgment referring to the thesis of an exclusive dependence of human action on the laws of nature (which we understand and know), since the moral law commands a form of action that implies negative freedom, that is, an escape from the necessity of the series of phenomenal events governed by natural causality in which our actions are nevertheless located.¹⁷ Being free allows the will not to be inevitably determined by nature, and therefore it can be ruled through spontaneous representations and act according to maxims. On the other hand, as the first principle of practical reason, the moral law indicates the essential quality that an action must have in order to take the form of free causality, and in this way, autonomy also determines the criterion of the choice of particular ends. The maxims consistent with the moral law share their form (the unconditional and universal character) at their specific level of generality and can therefore be considered laws of will (objective rules of freedom, necessary with respect to what should be). It seems therefore clear that they are not universalized empirical rules but – in the sense mentioned previously – a priori synthetic propositions in which an act of reason (i.e. spontaneous or autonomous) connects the universal form of the law (the principle of autonomy) with matter and empirical conditions of the will in particular fields and circumstances. These propositions are therefore born as self-formulated specific laws of our power of choice.

An imperative is a practical rule by which an action in itself contingent is *made* necessary. An imperative differs from a practical law in that a law indeed represents an action as necessary but takes no account of whether this action already inheres by an *inner* necessity in the acting subject (as in a holy being); or whether it is contingent (as in the human being); for where the former is the case there is no imperative. Hence an imperative is a rule the representation of which *makes* necessary an action that is subjectively contingent and thus represents the subject as one that must be *constrained* (necessitated) to conform with the rule.

(MSRL, 06: 222)

The imperative form of the rule pertains, we could say, to the passage from the moral law, which would govern without exception the pure practical reason or will, to the maxim of the will of a natural rational being, which in the aforementioned law has the principle of its own autonomy, whose legislative form should always be preserved in the practical principles which she considers valid for herself.¹⁸

Practical principles are propositions that contain a general determination of the will, having under it several practical rules. They are *subjective, or maxims*, when the condition is *regarded by the subject as holding only for his will*; but they are objective, or practical laws, when the condition is *cognized as objective, that is, as holding for the will of every rational being*. (KpV, § 1, 05: 18, italics added)

Practical laws, insofar as they are at same time subjective grounds of actions, i.e., subjective principles, are called maxims.

(KrV A 812/B 840)

The *arbitrium sensitivum liberum* is determined through the not necessary but personal assumption of the maxims and thus chooses the particular ends by conforming to the prescriptions of the practical laws (the value of which can be unconditional or private, depending on their ground of determination), showing, however, in the choice the causal power of the pure will that makes the formulation of the same maxims possible and meaningful. The use of expressions that refer to the laws of will, in the plural, can therefore be explained, despite the principle of pure practical reason being the only moral law. “Laws proceed from the will, *maxims* from choice (*Willkür*)”. From a subjective point of view, the maxims are an expression and manifestation of the will that independently determines itself with respect to a phenomenal matter and adopts its own particular laws. From the objective one, they identify the ends that give meaning to free actions as contingent phenomenal events, which nevertheless have their own practical necessity.

This notion therefore illustrates, more than those of desire or ends, the profound – formal, non-empirical, normative, originally relational, both subjective (as autonomous) and objective (as general) – structure of will and action.¹⁹ So let us dwell a little longer on the concept of the maxim to indicate its role – according to the Kantian concept of self-education and realization of the human being – in pursuing by means of our faculties the “ultimate end” of nature (culture of mind, skill, knowledge, science) and the “final end” of humanity (autonomy, freedom, morality). These ends are distinct but combined in the ideal of wisdom, which will certainly be needed to govern sustainable development.

4. On the notion of maxims

The Kantian concept of maxim has many meanings,²⁰ but its use in works dedicated to practical philosophy has been more widely articulated and discussed. Collecting the threads of this debate, Jens Timmermann distinguishes at least three main meanings of the term and shows how each of them allows us to focus better on texts and problems of Kantian ethics.²¹

The term “maxims” denotes first of all the principles of individual human actions, understood in the proper sense, that is, as free actions. As we have seen, Kant thinks that human action – all his actions in the world of phenomena – is never only the product of natural causes but the result of an act (of the subject of acting as a noumenon) or the free choice of the rule to follow in determining one’s own will in one way or another with respect to a certain circumstance (at least if, as man says, the person “is in himself” and “has not lost his reason”; otherwise, it is no longer freedom but nature that operates in her).

In other words, the maxims provide a criterion for distinguishing actions and events. . . . Kant distinguishes, in accordance with school terminology, acting according to principles (according to maxims, voluntary) from the one that takes place on the basis of simple stimuli; in this second case there is no rational mediation.

(Bacin 1999, 338)

Defining a choice as free means, we repeat, thinking it in the negative as constrained but not subjugated to natural causes and in the positive as determined instead by an original causality, specifically freedom, which manifests itself in the duty to pursue the general interest of reason or to obey the moral law.

It could be said, by vague analogy, that the unity of the maxim, by virtue of the principle of autonomy, accompanies and guides the actions of the human being, just as the transcendental unity of apperception accompanies and unifies all my representations, meaning that in both cases, a relationship

takes place between a non-empirical (pure or non-pure) condition, necessary with respect to an empirical condition, and this conditioned itself as made possible by that condition, even if the aforementioned condition is not always the object of a conscious reflection or a manifest content of the consciousness that the subject has of that conditioned. The “I think” must *be able* to accompany all my representations, but it does not necessarily become part of the content of the subject’s representations; similarly, the guiding maxim of an action must *be able* to be formulated and eventually become the object of reflection, but we do not always act with the awareness of the maxim that gives unity to our action, that is, of the rule of a certain behavior.²² The fundamental relationship between the faculty of desire and the end of the will is nonetheless established in the maxim “through a representative mediation”, whereby the desires and inclinations that make up the subject matter of the maxims are “raised to concept, so that they do not merely give rise to a solicitation according to stimuli” (Bacin 1999, 341).²³

The actions, however, usually do not depend on an isolated maxim, but – above all – the choice of specific maxims takes place in the light of more general principles that guide us and allow us to establish a (tendentally) systematic order among the maxims available to the power of choice. They are higher-level maxims – in a second sense – which do not refer directly to individual volitions but to the acts or resolutions that direct one’s will to act in a certain way in given or possible situations. We do not dwell here on the complex theme of character and its intertwining with that of the maxims,²⁴ but it is clear that, for Kant, the latter are “determination of the will that does not remain unrelated to each other, devoid of subject: on the contrary, they . . . build . . . a moral identity” and are not to be considered generic rules of impersonal life. “A maxim can be specified and developed in another, but . . . it is never a question of the relationship that is established between a general principle and its application at a lower level of generality; instead, it is about the acts of the will and the dynamics of their succession” (Bacin 2006, 203f.).²⁵ We can define this kind of maxims as principles of determination of will and acting in a defined context, which can also *depend* on more general maxims, which in turn can *refer* to even more comprehensive principles and so on, up to the more general practical principle, defined from the fundamental choice of wanting to act as the moral law commands us or, on the contrary, as our particular private interest suggests (italics emphasize the choice of verbs that should not suggest subsequent moments of application, much less a deductive deliberation).

On the one hand, we always act according to maxims: human actions in the strict sense . . . as such are based on self-imposed principles – in a weak sense . . . ; we are responsible for our action and its subjective principles. On the other hand, precisely from the fact that our faculty of desiring is accessible to rational influences

derives the duty to examine, order and revise these rules: we must consciously and autonomously set the right moral maxims.

(Timmermann 2003, 156)

As a result of this process of reflection and determination of the will, we obtain maxims of which we are aware in a clear and distinct way, and this is the third specific meaning of the term: more or less general and solid but aware and constant subjective principles of our way of acting.

In any case, the Kantian maxims are subjective principles, with this adjective indicating not a limitation of the scope of the rule (which, as subjective – according to a common meaning – would not be general, that is, objective, valid always and for all but merely particular) but his active assumption by a subject to immediately determine his own will with respect to the ends that he intends to pursue with his own actions (to be considered therefore in reflection examples or particular cases of the application of a practical law generated and taken as a principle). With Timmermann, we can say that “the maxims essentially specify an end which is preordained [to the actions], which coordinates morally neutral rules. Freedom in the choice of the maxims is a freedom in the choice of the fundamental purposes of one’s own action” (2003, 186) and consequently of the means necessary to achieve them. In essence, choosing a maxim means deciding that – obviously not in the abstract and arbitrarily but in a specific situation or in circumstances of a certain type or in general in my life – I intend to act following the technical-practical rules (the hypothetical imperatives of GMS) necessary to achieve a certain end, rather than another, in accordance or not with the moral law (with freedom in the Kantian sense), that is, obeying or not obeying the categorical imperative.²⁶

The term “maxim” indicates a rule “that the power of choice gives to itself for the use of its freedom”,²⁷ that is, that the person autonomously formulates and makes his own to give form and content (and therefore moral or amoral quality) to his own act (for which he is therefore responsible). The maxims, even before the actions, are therefore the result of our response to the command of the moral law, which is renewed in every concrete circumstance of our practical experience. Their choice marks the articulation of the general interest of reason with respect to the particular areas of human experience and therefore the concrete becoming and realization of the same reason in an adult individual, hopefully also autonomous in concrete deliberations (not only in principle).

Maxims must originate from the human being himself. . . . In the beginning these are school maxims and later maxims of humanity. In the beginning the child obeys laws. Maxims too are laws, but subjective ones; they originate from the human being’s own understanding.

(Päd 09: 481)

5. “Mankind must progressively pull out of itself the whole natural disposition of humanity”

In conclusion, however paradoxical it may seem at first sight, the formal and a priori nature of human practical reason characterizes the spontaneity of the individual subject of experience and the autonomy of the person as a faculty to act, which is free and at the same time needful of concrete contextual determination. Hence, the human being is constitutively determined and guided by laws and turned towards the world in the theoretical interest and in the choice of purposes. He is a rational being thanks to the faculties of his mind, but this does not mean that he always thinks, knows, acts or has feelings in accordance with the principles that make him precisely rational, that is, free; on the other hand, he is only, according to the well-known Kantian expression, a “rationable” being²⁸ and therefore constantly on the path towards his own freedom. This is reflected in the fact that

the human being is the only creature that must be educated. . . . Animals use their powers as soon as they have any in a regular manner. . . . An animal is already all that it can be because of its instinct; a *foreign intelligence* (*Vernunft*) has already taken care of everything for it. But the human being needs *his own intelligence* (*Vernunft*). He has no instinct and must work out the plan of his conduct for himself. However, since the human being is not immediately in a position to do this, because he is in a raw state when he comes into the world, others must do it for him.

(Päd 09: 441)

We highlight only some elements of this passage, without going into the merits of the “Kantian ethology”, which is however useful here for a clarification by contrast: animals are what they are by nature, they behave in a certain way because their mind is made this way, and instinct guides them to the use of their forces “in accordance with the rules”. It cannot therefore be said, strictly speaking, that they follow a rule, because an action of this type at least implies the possibility of compliance with rules subject to reflection, at least in order to choose between alternative dispositions. Animals would therefore have no “maxims”, subjective principles; it is as if an external reason had established for them rules of behavior to which they are simply subject, governed by laws of nature, without having to worry about formulating them, choosing them or putting them in order, but also without the possibility of doing these things. Since he belongs to the phenomenal world, the human being is not necessarily spared this destiny, but he is human precisely by virtue of his freedom, and this means that he must use his own faculties, dispositions and forces to make himself. “The human being can only become human through education”. “The human species is supposed to

bring out, little by little, humanity's entire natural predisposition by means of its own effort" (Päd 09: 443, 441).

This means that when a human being comes into the world, his abilities are "raw" and must be "cultivated" so that they develop and are applied in a correct and not (self) destructive way in accordance with the rules that must be set by oneself; therefore, he needs reason, or rather his own reason, which, however, beyond the first formal principles and the disposition to use them, is given him above all as a task, as something that he must make his own, according to the delphic-pindaric maxim: know yourself, recognize what you are and try to become one. The human being has to make himself free in his cognitive and practical action through the self-knowledge and culture of his mind, forming his own nature and defining the principles of his freedom,²⁹ the principles of his action and therefore the ends to be pursued. Herein lies the cultivation (*cultura*) of human "natural powers (powers of spirit, mind, and body) as means to all sorts of possible ends", which is "a duty to himself" (MSTL 06: 444).³⁰

In the Kantian conception, culture is considered, as is known, the ultimate end of a teleologically understood nature, that is, of man as a natural rational being; through culture, nature predisposes him to his final destination of moral being.³¹ In addition, the human being by nature has "a powerful propensity towards freedom", although this disposition is also affected by "a certain roughness". This is not "a noble propensity towards freedom, as Rousseau and others believe. . . , in that the animal in this case has so to speak not yet developed the humanity inside itself" (Päd 09: 442). In general, "many germs lie within humanity, and now it is our business to develop the natural dispositions proportionally and to unfold humanity from its germs and to make it happen that the human being reaches his vocation (*Bestimmung*)" (Päd 09: 445). Adopting a teleological language (within the limits set by the Third Critique), it can be said that nature has endowed the human being with dispositions to develop the faculties whose use, appropriately and critically "cultivated" (which is possible only in society),³² takes him beyond the limits of the same nature (*natura naturata*). One aspect of this acquisition concerns what is useful or necessary to live in society ("*Zivilisierung*")³³ and finally, on the horizon of his final destination,

the human being should not merely be skilled for all sorts of ends, but should also acquire the disposition to choose nothing but good ends. Good ends are those which are necessarily approved by everyone and which can be the simultaneous ends of everyone.

("Moralisierung", Päd 09: 450)

The capacity to set oneself an end – any end whatsoever – is what characterizes humanity (as distinguished from animality). Hence there is also bound up with the end of humanity in our own person the

rational will, and so the duty, to make ourselves worthy of humanity by culture in general, by procuring or promoting the capacity to realize all sorts of possible ends, so far as this is to be found in the human being himself. In other words, the human being has a duty to cultivate the crude predispositions of his nature, by which the animal is first raised into the human being. It is therefore a duty in itself.

But this duty is a merely ethical one, that is, a duty of wide obligation. No rational principle prescribes specifically how far one should go in cultivating one's capacities (in enlarging or correcting one's capacity for understanding, i.e., in acquiring knowledge or skill). Then too, the different situations in which human beings may find themselves make a human being's choice of the occupation for which he should cultivate his talents very much a matter for him to decide as he chooses. – With regard to natural perfection, accordingly, there is no law of reason for actions but only a law for maxims of actions, which runs as follows: "Cultivate your powers of mind and body so that they are fit to realize any ends you might encounter," however uncertain you are which of them could sometime become yours.

(MSTL, 06: 392)

If the "the cultivation of any capacities whatever for furthering ends set forth by reason" (MSTL, 06: 391) is the ultimate end of nature in view of the ultimate goal of morality (of acting on the basis of what is right, of duty), and if education aims at its development, it is also possible to invert the point of view, as this passage of the *Doctrine of virtue* bears witness to, and to recognize that the development of culture is itself a duty, however "imperfect" or "of wide obligation", since it corresponds to a general maxim that refers to possible actions and not to specific maxims imminent to concrete actions aimed at pursuing particular purposes. As already mentioned, rules must be subordinated to the maxim of an action (as determination of one's own will) in order to put it into practice and achieve the end, and if the end is a duty, it is also a duty to develop the skills and means to achieve it. Culture is therefore inherent in the dimension of the passage from the determination of the will in general through the choice of concrete ends to the realization of the latter as a process, which is technically necessary with respect to the achievement of the ends and so dutiful in itself.

The faculty to set oneself ends becomes concrete and develops when its principles (moral law and maxims) are applied to the matter of practical experience, which is certainly not a brute datum: obviously intellect and judgment are needed to "read" the circumstances and the alternative possibilities, to formulate appropriate maxims and examine their general consistency and validity. Culture, especially that of cognitive faculties, intervenes necessarily in the process of will determination and choice of ends. To be able to act well, it is obviously also necessary to think well, but since – as we have seen – for

Kant, human forces and faculties do not work automatically by instinct but are always to be acquired, developed and exercised in the first person, which requires *knowing* how to think well or a normative reference to rules to be taken as principles of one's own thought action, of which we can become aware. Therefore, the previous general maxim of culture, in the passage just cited in the *Doctrine of Virtue*, implies or corresponds to the general maxims of the empirical intellect referring to the correct way of thinking.³⁴

The reasonable being must learn to make use of his own reason and cannot do it with a solipsistic exercise: in the theoretical and technical-practical use, he is educated to follow from the beginning rules established by the reason of others, until he is able to form or assume his own maxims; however, with respect to the practical use of one's freedom, no one can ever dictate the rules. A child rather must be educated to act in general according to maxims and therefore must be treated from the outset as an autonomous being.³⁵ Moral education is aimed directly at autonomy and self-thinking; education to culture indirectly, since it is guided by and falls under the responsibility of those who protect the minor until he reaches maturity. The first form of education should lead, upon reaching the age of majority, to the definitive assumption of moral and legal responsibilities; the second can be open and permanent, because in knowledge and in society, one is never self-sufficient. However,

it is because of laziness and cowardice that so great a part of human-kind, after nature has long since emancipated them from other people's direction (*naturaliter maiorennes*), nevertheless gladly remains minors for life, and that it becomes so easy for others to set themselves up as their guardians.

(WA 08: 35)

Hence the famous appeal "Sapere aude! Have the courage to use your own intellect!" This is the motto of the Enlightenment, as it demands the man's exit from a state of minority for which he himself is responsible. Claudio La Rocca appropriately draws attention to two late reformulations of this exhortation:³⁶

(*Sapere aude*) Try to use your own reason for your true absolute ends.

For all knowledge (*scientia*) that the reasoning man can use for his own well-being, self-knowledge (*nosce te ipsum*) is a commandment of reason which contains everything; *sapere aude* sey wise: a possession that, if it is not already there, you cannot even acquire it.³⁷

In these late fragments, the tone is more forgiving than in the just-cited essay on enlightenment of 1784, in which Kant however denounces how

the laziness and cowardice of those who remain in a state of minority are ultimately consequences of cultural and civil (mis)education given to human beings by their guardians, who have treated them as and continue to consider them “pets”. The lexical passage from “intellect” to “reason” in the texts of the two periods is not particularly significant, since in the writing on enlightenment, the first term was used in its broadest meaning and did not refer to a specific higher faculty of knowledge (intellect, faculty of judgment, reason); in fact, Kant, repeating his appeal in the continuation of the essay, also uses “reason”, always in the broad meaning of general faculty that makes human beings rational. On the contrary, the specification of the essential moral and cosmopolitan nature of this faculty, to which appeal is made here, is relevant.

Reason has been very often misunderstood in the sense of an instrumental rationality characterized by a formalism devoid of content, a mere formal ability; instead (*repetita iuvant*), the formalism of the Kantian reason has to do with the universal and normative character of the autonomy of the will (and of the principles of thought and knowledge), which constitute the conditions of possibility and the ultimate horizon of the sense of human action. As mentioned previously, reason is above all the autonomous faculty of the will to determine itself, setting ends, and of the power of choice to formulate the maxims and direct the action to their realization. Precisely because and by virtue of its freedom and spontaneity, reason is not something that we can take as given by nature, but it must reach itself, cultivating itself. In this sense can be read the second of the last quotes: to exit from the state of minority means to appropriate a faculty which we always possess on an intelligible level (as a kind of simply “legal” possession or *possessio noumenon* to be assumed a priori, because otherwise, we as empirical subjects could not even claim it or acquire it) but which we are unable to exercise in a mature way from the beginning. To use one’s reason in the first person means to gain possession of it in the sphere of concrete life and in the phenomenal world (*possessio phaenomenon*). If I were prevented from using something I own, this would infringe a right. In the case of reason, claiming the right to autonomy is even a fundamental duty commanded by the categorical imperative, because it is linked to the final end of humanity. However, in order to be able to claim one’s rights, it is necessary to know them; therefore, the command of reason to know oneself (i.e. above all our reason) is basic. Making use of one’s reason autonomously, as it should be according to our “final nature”, our right and moral law, means not letting oneself have the ends of one’s own action dictated from the (un)reason of others and not determining one’s own will according to particular interests that, following false needs and fictitious aims, do not correspond to the full development of humanity in us and in the species to our moral destination. Really making use of one’s reason means, therefore, pursuing “your true absolute ends” and, since “wisdom is the relationship to the essential ends of humanity”, Kant finally translates the Horatian *sapere aude* as “be wise”.

6. Kant's practical philosophy and the maxims of humanity for the 21st century

Kant's practical reason is formal in that it is essentially autonomous and critical; it neither dictates a priori the ends of actions nor limits itself to guiding action pragmatically to the achievement of all possible ends but constantly refers the use of the principles a priori of reason to its "true absolute ends".

Autonomy is not simply "the decision not to be governed" . . . , but also the need, felt by reason itself as part of its logic, to contribute to the production of the historical conditions of a generally shared possibility of freely giving shape to meaningful experiences.

(La Rocca 2004, 130)

The two interests of reason converge in the "aspiration to wisdom, which is always unfinished" (OP, AA 20: 6). The position of the final end requires the ability to pursue it in practice through the necessary means, which in turn become essential ends, giving rise to a system of ends and maxims, which, in order to be conceived and implemented, requires skills and conditions which are summarized in the concept of culture, and among these, not least, are knowledge and "science (critically sought and methodically introduced)", which Kant considers "the narrow door that leads to the doctrine of wisdom" (KpV, 05: 163). The reference to wisdom, "but through the path of science, the only one that once opened never closes again and does not allow to get lost" (KrV, A 850/B 878), also constitutes the cosmopolitan concept of philosophy,³⁸ which therefore is naturally suited to contribute significantly to the formation of a new culture of sustainable development.

The latter – spreading especially in the younger generations – obviously requires ethical choices and assumptions of responsibility that appear "natural" if considered in the context of the reflections on rationality, will, power of choice, action and praxis carried out by Kant. For example, the people of Fridays for Future demand the adoption of measures that will quickly lead to sustainable development, simply because this is the only rational and morally right thing to do for positive development and for the future of humanity. They then appeal to science and ask politicians to be guided by the best theories available today and do so not for ideological scientism (unaware of the complexity and necessity of social and political mediations, which are necessary also to achieve the most complete and objective description and understanding of the "matter" of choices), but because they see and feel the need to combine morality and rational reflection guided by knowledge in the formulation of the maxims appropriate to the idea of a just government of the interests of all the subjects involved (in this case, literally of all humanity, including future generations). If they adopt the cosmopolitan perspective and

share the characterization of Kant's wisdom, they attest *eo ipso* the *practical* relevance of Kantian practical philosophy. The theoretical one is implicitly thematized in the previous exposition, in which too many themes – some very contentious and controversial among the interpreters – were certainly touched upon with insufficient arguments, but only to try to compose a schematic but not one-sided design of the Kantian point of view in order to highlight the usability of his philosophy in the contemporary world.

However, what has just been mentioned does not apply to all the positions in which ethics is articulated? In the moment in which the moral conscience is listened to, it becomes evident that the ends are not homogeneous and that it is right to pursue those who respect and favor humanity and to reject the selfish ones who ignore or exploit it. Of course, it may be difficult to achieve clarity and distinction in the representation of the good ends for humanity, ones adequate to the objective principle of freedom and preparatory to a morally qualified happiness, independent of narrow pragmatic interests. It seems to me, however, that the conditions action must confront today, which require the exercise of a mature and self-aware rationality, highlight the merits of the Kantian approach. Let me be extremely schematic: Utilitarian and generally teleological theories hold that the fundamental criterion of what is morally right is the benchmarking of the good produced. The good is defined independently of the right, and then the right is defined as that which maximizes the good. Deontological theories affirm that in addition to the goodness and wickedness of the consequences of actions, other criteria must be considered: natural rights, values, social contract and so on. But both positions presuppose a general agreement on some fundamental values with respect to nature and social life, which today is continually questioned or ignored, making clear the impossibility, already diagnosed by Kant, of building a general will starting from the grounds of heteronomous determination of will. In particular, the idea of acting towards sustainability is evident today as a “true” and necessary end of humanity and is contradictory with respect to the unconditional affirmation of particular wills. It can perhaps be said that today's historical development forces us to seek the principles for our deliberations on the only apparently algid and abstract level of the pure practical reason recognized by Kant. This is useful to give our practice the character of universality and objectivity that on the one hand is necessary to successfully pursue the common and essential purposes and on the other corresponds and appeals to the formal structure which is immanent to the actions of people but easily blurred in individual consciences by the variety and fragmentation of the matter of will (multiplication and atomization of needs). Kant's moral philosophy can support the trust in the ability of every human being to overcome selfish motives, and its purpose is precisely to lead reason to itself so that human beings can be free and pursue happiness in harmony with each other.

Kant recognizes a pure moral obligation, which must be fulfilled through an open and perfectible system of subjective maxims that follow the general

principle of practical reason, the moral law. To respect and promote humanity, the agent can, and has the duty (to try) to, orient and to ethically constitute his will and action in a concrete and non-contradictory way according to the principles of shared freedom and happiness. We could say that Kant's ideal is not mutual conformity but – if not harmony – at least the compatibility in freedom of the persons' actions, or their moral sustainability, that combines the multiplicity of maxims in reference to moral law. Recall that the moral law is the practical pure principle, which makes all maxims possible and fixes their criteria of objectivity and evaluation; it requires the power of choice to formulate for itself maxims that have its practical pure form (as subjective universal rules) and can inform free actions that pursue “absolute ends”, mine and particular, yet shared by an unlimited community of free subjects. The categorical imperative does not prescribe specific duties but the formal criterion for deciding whether a maxim is moral (sustainable for humanity and “its” world). It commands one to act on the basis of a universalizable maxim, but the same maxim can be followed differently according to specific circumstances. We can therefore hypothesize that the answer to the question “if sustainable development presupposes ethical actions, which ethics should we develop in view of sustainability?” could be “a Kantian maxims ethics that combines autonomy, rationality and moral sensitivity to specific contexts”.³⁹

A certainly superficial, but in my opinion not negligible, verification of this hypothesis is offered by the fact that the goals formulated in Agenda 2030 present imperatives, which should be assumed to be “higher” maxims of the power of choice and of rational action by everyone on this planet, first of all by the nations and subjects that govern development. Those goals can be considered maxims that claim to be valid as practical laws, since they formulate – in accordance with a priori normative principles – objective rules for the factual exercise of the freedom of all rational beings and thus indicate (very general) true ends of a good will and therefore ends that are also duties. Obviously this interpretation does not exempt a reflection on the content of those specific practical laws (and on their eventual systematic order) and does not ensure that each of them is congruent with what Kant wrote. As an example, the following concluding remarks address briefly these issues about the first of the goals listed in the Agenda: “end poverty in all its forms everywhere”.⁴⁰

At first glance, it, like many of the other goals, aims to make people happy and therefore seems to be more utilitarian and paternalistic than conforming to the spirit of Kantian morality. Instead, we have already seen that the general maxim of morally qualified happiness really is so, as MSTL also confirms.

What are the ends that are also duties? They are one's own perfection and the happiness of others. Perfection and happiness cannot be interchanged here, so that one's own happiness and the perfection

of others would be made ends that would be in themselves duties of the same person.

(MSTL 6:385)⁴¹

It could still be argued that not every moral obligation results in a right. So it would be a duty to help the poor, but in the liberal sphere of charity, not in that of “duties of right (*officia iuris*), that is, duties for which external lawgiving is possible” (MSRL 6: 239), which only guarantees mutual respect for the freedom of others and therefore does not impose duties depending on the “desires” of others (MSRL 6: 230). But Kant views poverty from a dual perspective. Certainly from that of charity, understood, however, as a real and anything but supererogatory duty that is exposed in the MSTL with a progressive shift of accent from generosity to the maxim, which turns out also to be a practical law, and finally to a kind of duty of compensation:

We shall acknowledge that we are under obligation to help someone poor; but since the favour we do implies that his well-being depends on our generosity, and this humbles him, it is our duty to behave as if our help is either merely what is due him or but a slight service of love, and to spare him humiliation and maintain his respect for himself (6:449–450).

Everyone who finds himself in need wishes to be helped by others. But if he lets his maxim of being unwilling to assist others in turn when they are in need become public, that is, makes this a universal permissive law, then everyone would likewise deny him assistance when he himself is in need, or at least would be authorized to deny it. Hence the maxim of self-interest would conflict with itself if it were made a universal law, that is, it is contrary to duty. Consequently the maxim of common interest, of beneficence toward those in need, is a universal duty of human beings, just because they are to be considered fellow human beings, that is, rational beings with needs, united by nature in one dwelling place so that they can help one another (6:453).

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? . . . 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 favoured 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?

(6:454)

On the other hand, consistent with what has just been read, it does not seem that Kant considers it improper to deal with poverty in the sphere of law, as attested by the following passage, well known and much discussed.

The general will of the people has united itself into a society which is to maintain itself perpetually; and for this end it has submitted itself to the internal authority of the state in order to maintain those members of the society who are unable to maintain themselves. For reasons of state the government is therefore authorized to constrain the wealthy to provide the means of sustenance to those who are unable to provide for even their most necessary natural needs. The wealthy have acquired an obligation to the commonwealth, since they owe their existence to an act of submitting to its protection and care, which they need in order to live; on this obligation the state now bases its right to contribute what is theirs to maintaining their fellow citizens.

(MSRL 6:326)⁴²

Now we cannot dwell further on these passages. However, it seems implausible that, according to the Kantian conception of justice, the fight against existing poverty should be only a matter of charity and not of justice or that there are duties of distributive justice to beneficially eliminate poverty only within a particular community, and in this regard tasks and rights of the state, but that this does not apply to the rest of humanity. Merten Reglitz, for example, reconstructs a Kantian argument against world poverty, which aims to integrate and correct arguments of current liberal theorists who refer to Kant and is so summarized by himself:

If we accept the claims (A) that according to the original contract test, Kant's theory includes a duty to eliminate poverty as a necessary condition of the legitimacy of any legal order, and (B) that humanity has the duty to create a coercive global condition to enable conclusive external possessions, it follows that no one must live in poverty anywhere within this global civil condition lest the entire scheme of justice and property rights be illegitimate.⁴³

Notes

- 1 <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- 2 S. sect. 10–13 of the Agenda 2030.

- 3 www.un.org/sustainabledevelopment/development-agenda/: “The Sustainable Development Goals are not legally binding. Nevertheless, countries are expected to take ownership and establish a national framework for achieving the 17 Goals”.
- 4 The concept of sustainability was introduced during the first UN conference on the environment in 1972. In 1987, with the publication of the so-called Brundtland report, the goal of sustainable development was clearly defined and, after the UN conference on environment and development in 1992, has become the new paradigm of development itself.
- 5 www.nationalgeographic.org/encyclopedia/anthropocene/: “Officially, the current epoch is called the Holocene, which began 11,700 years ago after the last major ice age. However, the Anthropocene Epoch is an unofficial unit of geologic time, used to describe the most recent period in Earth’s history when human activity started to have a significant impact on the planet’s climate and ecosystems. The word Anthropocene is . . . coined and made popular by biologist Eugene Stormer and chemist Paul Crutzen in 2000. . . . In 2016, the Anthropocene Working Group agreed that the Anthropocene is different from the Holocene, and began in the year 1950 when the Great Acceleration, a dramatic increase in human activity affecting the planet, took off.”
- 6 See Corradetti 2020.
- 7 Practical is “all that is possible through freedom” (KrV, A 800/B 828), and practical philosophy is concerned with the practical use of reason, that is “the determining grounds of the will” (KpV, 05: 15 and 20).
- 8 “Strength of any kind can be recognized only by the obstacles it can overcome, and in the case of virtue these obstacles are natural inclinations, which can come into conflict with the human being’s moral resolution; and since it is the human being himself who puts these obstacles in the way of his maxims, virtue is not merely a self-constraint (for then one natural inclination could strive to overcome another), but also a self-constraint in accordance with a principle of inner freedom, and so through the mere representation of one’s duty in accordance with its formal law” (ibid.).
- 9 “This principle of morality, just on account of the universality of the lawgiving that makes it the formal supreme determining ground of the will regardless of all subjective differences, is declared by reason to be at the same time a law for all rational beings insofar as they have a will, that is, the ability to determine their causality by the representation of rules, hence insofar as they are capable of actions in accordance with principles and consequently also in accordance with a priori practical principles (for these alone have that necessity which reason requires for a principle)” (KpV, 05: 32).
- 10 See *infra* § 5, and here Päd 09: 441.
- 11 “The subjective conditions of freedom are the ability to act, and further, that we know what pertains thereto, that we are aware of the motivating ground and the object of the action” (V-Mo/Collins 27: 291), otherwise it would be “an *empty* will, without object” (FM 20: 342). The maxims “are synthetic a priori propositions which, unlike the supreme principle, cannot be pure, since they always have to do with a matter . . . that is, they are ‘material’ laws that express certain duties with respect to certain fields of action of a subject in general (which therefore becomes effectively autonomous)” (Bacin 1999, 357).
- 12 Kant says explicitly in the essay *On the Use of Teleological Principles in Philosophy* that “ends have a direct relationship to reason, be it foreign reason or our own. Yet, even in order to place them in foreign reason, we must presuppose our own reason at least as an analogue to the latter, since those ends cannot be represented at all without such an analogy” (ÜGTP 08: 182).

- 13 On the same line of thought is Kant's reflection on happiness, which appears to be the natural purpose of every human being, see (KU, § 83, 05: 430), quoted *infra*.
- 14 "All material practical rules put the determining ground of the will in the lower faculty of desire, and were there no merely formal laws of the will sufficient to determine it, then neither could any higher faculty of desire be admitted". "Then only, insofar as reason of itself (not in the service of the inclinations) determines the will, is reason a true higher faculty of desire, to which the pathologically determinable is subordinate, and then only is reason really, and indeed specifically, distinct from the latter, so that even the least admixture of the latter's impulses infringes upon its strength and superiority, just as anything at all empirical as a condition in a mathematical demonstration degrades and destroys its dignity and force. In a practical law reason determines the will immediately, not by means of an intervening feeling of pleasure or displeasure, not even in this law; and that it can as pure reason be practical is what alone makes it possible for it to be lawgiving" (KpV, 05: 22, 24 f.). "In fact the law, here again, prescribes only *the maxim of the action*, that of seeking the basis of obligation solely in the law and not in sensible impulse (advantage or disadvantage), and hence *not the action itself*" (MSTL, 06: 392).
- 15 In general, it can be remembered that, for Kant, practical freedom is not understood except as the faculty of the will to self-determine itself through reasons or rational determining grounds. Here I cannot articulate the aforementioned implication and neither can I explain how it emerges in Kantian Criticism. I will limit myself to report two quotes. "To admit that the moral law within us is itself deceptive would call forth in us the wish, which arouses our abhorrence, rather to be rid of all reason and to regard ourselves as thrown by one's principles into the same mechanism of nature as all the other species of animals" (MSRL, 06: 355). The human being recognizes herself both as "a sensible being" and as "an intelligible being", that "can be cognized only in morally practical relations, where the incomprehensible property of freedom is revealed by the influence of reason on the inner lawgiving will" (MSRL, 06: 418). On the same page, Kant distinguishes the aspect that makes the human being "a natural being that has reason (*homo phaenomenon*)", to which one can attribute an instrumental rationality, from what makes him "a being endowed with inner freedom (*homo noumenon*)", or rather a "personality" bearer of the practical interest of reason, to which as we know Kant attributes a primacy over the exclusively theoretical one, because its object is the (moral) destination of humanity and because "all interest is ultimately practical and even that of speculative reason is only conditional and is complete in practical use alone" (KpV, 05: 121). See Fn. 20 and Fn. 40.
- 16 On this complex and divisive concept, see Sensen 2013.
- 17 Of course this is not about an exception to the universal validity of the laws of nature but a sort of overcoming of the natural causes of the empirical will in the practical perspective, conceived by virtue of the fact that the concept of phenomenal causality, necessarily conditioned by the laws of nature, and the concept of the unconditional causality of freedom are not contradictory and can, indeed must, be referred to the same sphere, that of human actions, to satisfy two essential needs, different but intertwined, of reason. This is the main theme of the third antinomy, on which I allow myself to refer to Aportone 2018.
- 18 This way of considering their relationship reduces a tension within the normativity of the maxim. In fact, in the strict sense, the latter does not have a prescriptive character, since it indicates "not how the subject should act, but how he actually acts (in a practical-moral sense) or according to which directives he will act (in

the sensitive world)” and “no maxim has as an alternative a violation . . . of what is expressed in it . . . but only another maxim” (Bacin 1999, 346–347). However, they “possess a certain prescriptive charge in that they express the choice of the subject to act in a certain direction and for certain reasons, and the commitment it entails, together with the practical consequences implied by it” (Bacin 2006, 202). A clear passage on the aforementioned relationship between law, imperative and maxims is in *GMS*, 04: 420–21: “When I think of a hypothetical imperative in general I do not know beforehand what it will contain; I do not know this until I am given the condition. But when I think of a categorical imperative I know at once what it contains. For, since the imperative contains, beyond the law, only the necessity that the maxim be in conformity with this law, while the law contains no condition to which it would be limited, nothing is left with which the maxim of action is to conform but the universality of a law as such; and this conformity alone is what the imperative properly represents as necessary”.

- 19 I agree in this with Bacin 1999, 362: The maxims “are the essential structure in act of the action itself for the practical-moral aspect, the complete determination of it according to the object and the determining reason [. . . They are] the ultimate and concrete result of practical activity in the proper sense, as self-determination of the subject, and therefore . . . the primary data for the consciousness of the acting subject”.
- 20 The concept of maxim as a subjective rule of action is primarily a concept of practical philosophy, and, in fact, already in the *KrV*, it is introduced and defined in this sense. However, this does not mean that the Kantian use of this concept in other contexts has little to do with the plexus of meanings that we are about to summarize, because it is reason itself that is essentially practical, and experiencing and knowing are actions that refer to “categorical” laws and to subjective principles of their application, too. Furthermore, if on the one hand, the true nature of practical reason shows itself in the formalism of the moral law, which must be realized in a coherent system of maxims, on the other hand, the character of theoretical reason is revealed by its equally formal architectonic interest, which guides the intellect and the faculty of judgment: “The greatest systematic unity, consequently also purposive unity, is the school and even the ground of the possibility of the greatest use of human reason” (*KrV*, A 694/B 722). This also requires the autonomy of the reflective faculty of judgment set out in the *KU* as a transcendental condition. A comparison between the maxims of practical reason and those of reflective judgment would quickly show the same dynamic of the transition from pure laws to the particular laws of experience through the formulation and application of subjective principles. Georg Kohler has dealt with it from another perspective: “The constant striving of the reflective power of judgement to go beyond the already achieved and fixed [empirical] generality, to ascend to the completed system of all experience, is thus the concrete ‘architectonic interest’ of reason itself”. “Reason – be it as practical or theoretical – actually only means the unconditional, apodictic mission: the ‘faculty’ of the encompassing ends. Reason itself, therefore, is always that which comes before the concretization of these goals, not the concretization itself; – when it comes to the latter, the reflective power of judgment must appear, which turns the end of total knowledge into the ‘heuristic hypothesis’ of the purposiveness of nature and which is ever concerned with individual things”, Kohler 1980, 18 and 20, Fn. In short, precisely the formalism of the principles of reason makes it possible and demands a striving towards the achievement of its highest goals, that is, a spontaneous or free action of rational subjects, which, in accordance with the

formal unity of experience in general in its different fundamental dimensions, can achieve ends by determining themselves in the assumption of the maxims. And since the subjects determine themselves in reflection, one could also say that the whole “system of criticism” exposes, in the words of Birgit Recki, “the practicality of all judgment in its action character” and consequently the need to “incorporate” the formal laws of reason – whether these are directed towards nature or customs – in the concrete acts of the subjects, conditioned by the context of the action, but also self-determining through the assumption of subjective principles. According to Recki, the practical character of judging in general “is underlined by the application of the concept of maxim to the question of the way of thinking” in § 40 of the KU (Recki 2001, 120). Subjectivity (in the positive and non-limiting sense) and the normative a priori relation to action are traits that also characterize the “maxims of the common human intellect” and probably all the contexts in which Kant introduces the maxims of the faculties of the mind, which are never mere impersonal rules or laws of nature. For example, consider those of speculative reason, described in the KrV: “I call subjective principles that are taken not from the constitution of the object, but from the interest of reason in regard to a certain possible perfection of the cognition of this object, maxims of reason . . . methods satisfying this interest” (A 666/B 694); or those of the faculty of judgment, remembered in the Introduction of the KU as non-empirical principles that this same faculty has laid down as a basis for investigations into nature (E §V, 05: 182).

- 21 Timmermann 2003, 145–188. For an overall discussion of the notion of maxim in Kant, including reference to its sources and recent interpretations, see Bacin 1999; Bacin 2006.
- 22 This idea, that the maxim must not be a conscious principle, does not originate with Kant. In the Wolffian school, for example, it was present as a premise of an ethics understood as a moral characteristic, a doctrine that teaches one to read the signs to know what is hidden in the human soul; for some textual references, see Bacin 2006, 186, Fn. 30. In Kant, however, it fuses with the following thesis, which is apparently so far from his ethical “rationalism”: “The real morality of actions (their merit and guilt), even that of our own conduct, therefore remains entirely hidden from us. Our imputations can be referred only to the empirical character. How much of it is to be ascribed to mere nature and innocent defects of temperament or to its happy constitution (*merito fortunae*) this no one can discover, and hence no one can judge it with complete justice” (KrV A 551/B 579, Fn), as it would be, if the maxims were always evident, first and foremost to the agent’s conscience. Consequently, “the First Command of All Duties to Oneself” is “know (scrutinize, fathom) yourself, . . . in terms of your moral perfection in relation to your duty. That is, know your heart – whether it is good or evil, whether the source of your actions is pure or impure, and what can be imputed to you as belonging originally to the substance of a human being or as derived (acquired or developed) and belonging to your moral condition. Moral cognition of oneself, which seeks to penetrate into the depths (the abyss) of one’s heart, which are quite difficult to fathom, is the beginning of all human wisdom. For in the case of a human being, the ultimate wisdom, which consists in the harmony of a being’s will with its final end, requires him first to remove the obstacle within (an evil will actually present in him) and then to develop the original predisposition to a good will within him, which can never be lost. (Only the descent into the hell of self-cognition can pave the way to godliness.)” (MSTL, 06: 441).
- 23 See also the Fn. 50 and 51. See KpV, 05: 65–66.
- 24 See Munzel 1999; Gigliotti 2001.

- 25 The difference between a conception of the maxims as rules of life conceived in a hierarchy of levels of generality, of neo-Aristotelian connotation and that properly Kantian can also be indicated by emphasizing the position in which they place the maxim in practical syllogism. According to the traditional view, up to the Wolffian school, the maxim is seen as a major premise, a general assumption to which the action can be traced back to consolidate the will and the habit of following the rule indicated in it. In Kant's conception, it should instead perform the function of the minor premise, to reflect the process of determination of the will in the "subsumption" of the action to the moral law; see Bacin 2006: 185–200.
- 26 "Within a pathologically affected will of a rational being there can be found a conflict of maxims with the practical laws cognized by himself. For example, someone can make it his maxim to let no insult pass unavenged and yet at the same time see that this is no practical law but only his maxim – that, on the contrary, as being in one and the same maxim a rule for the will of every rational being it could not harmonize with itself" (KpV, 05: 19).
- 27 It is worth noting that in the original text, Kant uses the verb "to make", not "to give": "Regel, die die Willkür *sich selbst* für den Gebrauch ihrer Freiheit *macht*", *Religion*, 06:21.
- 28 See *Anthropology from a Pragmatic Point of View*, 07: 321–22.
- 29 See *Lectures on Pedagogy*, AA 09: 469–70.
- 30 This framework also includes the definition of metaphysics as "the culmination (*Vollendung*) of all culture of human reason, which is indispensable even if one sets aside its influence as a science for certain determinate ends" (KrV, A 850–51/B 878–79).
- 31 On the Kantian conception of culture, condensed in § 83 of KU, see Petronzio 2018, 161–178; Orth 2008; Flach 2007. Allow me to refer also to Aportone 2009.
- 32 See, for example, KU, 05: 432–33, and *Idea for a Universal History with a Cosmopolitan Aim*, 08: 17–31.
- 33 See Anth 07: 323–24.
- 34 This brings us to the important third paragraph of § 40 of the KU: "The following maxims of the common human understanding [. . . can] serve to elucidate its fundamental principles. . . : 1. To think for oneself; 2. To think in the position of everyone else; 3. Always to think in accord with oneself. The first is the maxim of the unprejudiced way of thinking, the second of the broad-minded way, the third that of the consistent way. The first is the maxim of . . . enlightenment. . . . As far as the second maxim of the way of thinking is concerned, . . . he sets himself apart from the subjective private conditions of the judgment, within which so many others are as if bracketed, and reflects on his own judgment from a universal standpoint (which he can only determine by putting himself into the standpoint of others). The third maxim . . . is the most difficult to achieve, and can only be achieved through the combination of the first two and after frequent observance of them has made them automatic. One can say that the first of these maxims is that maxim of the understanding, the second that of the power of judgment, the third that of reason" (05: 294–95).
- 35 "Moral culture must be based on maxims, not on discipline" (Päd 09: 480).
- 36 See La Rocca 2004, 123ff.
- 37 "(*Sapere aude*) Versuche dich Deiner eigenen Vernunft zu Deinen wahren absoluten Zwecken zu bedienen". "Zu allem Wissen (*Scientia*) dessen sich der vernünftelnde Mensch zu seinem Wohlseyn bedienen kann ist das Selbsterkenntnis (*nosce te ipsum*) ein Gebot der Vernunft welches Alles enthält: *sapere aude* sey weise: Ein Besitz der wenn man an sich nicht schon in seinem Besitz ist zu ihm auch nicht gelangt" (OP 21: 117 and 134, my translation).

- 38 See Bacin et al. 2013, Bd. I, and Henrich 1966. See also the definition of philosophy in *sensu cosmico* in the *Jäsche Logic*: “we can also call it a science of the highest maxim for the use of our reason, insofar as we understand by a maxim the inner principle of choice among various ends. For philosophy in the latter sense is in fact the science of the relation of all cognition and of all use of reason to the ultimate end of human reason, to which, as the highest, all other ends are subordinated, and in which they must all unite to form a unity” (09: 24).
- 39 Mats G. Hansson (1991) applies, to my knowledge in a pioneering way, this perspective in the field of bioethics.
- 40 See *supra* Fn. 1. The other goals of the list are as follows: 2. end hunger, achieve food security and improved nutrition and promote sustainable agriculture; 3. ensure healthy lives and promote wellbeing for all at all ages; 4. ensure inclusive and equitable quality education and promote lifelong learning opportunities for all; 5. achieve gender equality and empower all women and girls; 6. ensure availability and sustainable management of water and sanitation for all; 7. ensure access to affordable, reliable, sustainable and modern energy for all; 8. promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all; 9. build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation; 10. reduce inequality within and among countries; 11. make cities and human settlements inclusive, safe, resilient and sustainable; 12. ensure sustainable consumption and production patterns; 13. take urgent action to combat climate change and its impacts; 14. conserve and sustainably use the oceans, seas and marine resources for sustainable development; 15. protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss; 16. promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels; 17. strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development.
- 41 As evidence of the fact that Kant assumes a maxim of happiness aimed at satisfying only the persons’ “true absolute ends” (*supra* Fn. 40), see (MST 6:388): “When it comes to my promoting happiness as an end that is also a duty, this must therefore be the happiness of other human beings, whose (permitted) end I thus make my own end as well. It is for them to decide what they count as belonging to their happiness; but it is open to me to refuse them many things that they think will make them happy but that I do not, as long as they have no right to demand them from me as what is theirs”.
- 42 However in these pages, as in others, Kant is instead very prudent in indicating the practical-pragmatic rules which should be coordinated to the maxims of charity and welfare state: “a problem which has not yet been solved in such a way that the solution offends against neither rights nor morality” (6: 327, see also 6: 367 and 9: 452). A further problem is that poverty is a concept related to the failure to satisfy needs, which can be considered “natural” in the specific context of exercising one’s freedom. “Now no one can think a negation determinately without grounding it on the opposed affirmation. The person blind from birth cannot form the least representation of darkness, because he has no representation of light; the savage has no acquaintance with poverty, because he has none with prosperity. The ignorant person has no concept of his ignorance, because he has none of science, etc.” (A 575/B 603).
- 43 M. Reglitz, “A Kantian Argument against World Poverty”, *European Journal of Political Theory* 2016, 18(4), 489–507, p. 501. See the notes of this essay for bibliographical indications on the broad contemporary debate on justice and poverty.

References

- Aportone, A. 2009. "Il controsenso delle disposizioni naturali. Sulle relazioni tra natura, ragione e moralità secondo Kant." In F. Biasutti (a cura di), *Ethos e Natura. Ricerche sul significato dell'etica per la modernità*. Napoli: Bibliopolis, 379–410.
- Aportone, A. 2018. "Libertà e natura nelle serie fenomeniche." *Rivista di storia della filosofia* 73: 79–118.
- Bacin, S. 1999. "Massime e principi pratici in Kant." *Istituto Italiano per gli Studi Storici* 16: 323–362.
- Bacin, S. 2006. *Il senso dell'etica. Kant e la costruzione di una teoria morale*. Bologna: Il Mulino.
- Bacin, S., Ferrarin, A., La Rocca, C., and Ruffing, M. (Hg.). 2013. *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses*. Berlin; New York: de Gruyter.
- Corradetti, C. 2020. *Kant, Global Politics and Law. The World (State) Republic as a Regulative Idea of Reason*. London; New York: Routledge.
- Flach, W. 2007. "Kants Begriff der Kultur und das Selbstverständnis des Neukantianismus als Kulturphilosophie." In M. Heinz and C. Krijnen (Hg.), *Kant im Neukantianismus. Fortschritt oder Rückschritt?* Würzburg: Königshausen & Neumann, 9–24.
- Gigliotti, G. 2001. "Naturale e artificiale: il problema del carattere." *Rivista di filosofia* 92: 411–433.
- Hansson, M.G. 1991. *Human Dignity and Animal Well-Being. A Kantian Contribution to Biomedical Ethics*. Uppsala: Almqvist & Wiksell.
- Henrich, D. 1966. "Zu Kants Begriff der Philosophie. Eine Edition und eine Fragestellung." In F. Kaulbach and J. Ritter (Hg.), *Kritik und Metaphysik. H. Heimsoeth zum 80. Geburtstag*. Berlin: de Gruyter, 40–59.
- Hume, D. 1896. *Treatise of Human Nature*. Edited by L.A. Selby-Bigge. Oxford: Clarendon Press.
- Kohler, G. 1980. *Geschmacksurteil und ästhetische Erfahrung. Beiträge zur Auslegung von Kants Kritik der ästhetischen Urteilskraft*. Berlin; New York: de Gruyter.
- La Rocca, C. 2004. "Was Aufklärung sein wird. Zur Diskussion um die Aktualität eines Kantischen Konzepts." In R. Langthaler and H. Nagl-Docekal (Hg.), *Recht – Geschichte – Religion. Die Bedeutung Kants für die Philosophie der Gegenwart*. Berlin: Akademie Verlag.
- Munzel, G.F. 1999. *Kant's Conception of Moral Character. The Critical Link of Morality, Anthropology, and Reflective Judgment*. Chicago: The University of Chicago Press.
- Orth, E.W. 2008. "Kant als Kulturphilosoph. Eine Interpretation der Kantischen Philosophie durch die Neukantianer?" In M. Kugelstadt (ed.), *Kant-Lektionen. Zur Philosophie Kants und zu Aspekten ihrer Wirkungsgeschichte*. Würzburg: Königshausen & Neumann, 231–240.
- Petronzio, S. 2018. *L'essere razionale nel mondo. Moralità e teleologia nell'Appendice alla terza Critica kantiana*, Doctoral thesis. Roma: La Sapienza.
- Recki, B. 2001. *Ästhetik der Sitten. Die Affinität von ästhetischem Gefühl und praktischer Vernunft bei Kant*. Frankfurt am Main: Klostermann.
- Sensen, O. (ed.). 2013. *Kant on Moral Autonomy*. Cambridge: Cambridge University Press.
- Timmermann, J. 2003. *Sittengesetz und Freiheit. Untersuchungen zu Immanuel Kants Theorie des freien Willens*. Berlin; New York: de Gruyter.

KANT AS AN ANTE LITTERAM THEORIST AND CRITIC OF THE MORAL ENHANCEMENT

Alberto Pirni

1. Premise

In this chapter, I shall present the result of a tentative conjunction of two research paths that I have been developing throughout many years: the first one concerning Kant and practical philosophy (with a specific focus on a communitarian interpretation of the categorical imperative doctrine) and the second one pertaining to ethics and emerging technologies (with a specific reference to human enhancement and roboethics).¹ Accordingly, this is the first attempt to orient my research agenda on Kant towards issues related to human enhancement and to the ethics of emerging technologies. The argumentative path here proposed is the following: after a preliminary framing of the issue, which makes reference to the most relevant and correlative key points within Kantian moral theory, I shall introduce a comprehensive two-fold argument in order to first consider what we could call “the internal enhancement” towards the asymptotic full adequacy to the moral law and then to discuss the “external-internal” help or integration to our moral adequacy and its alleged eligibility. Finally, I will consider two intrinsically problematic areas related to the same rationale in order to conclude by raising some correlative issues *within* and *beyond* the Kantian perspective.

2. A preliminary framing (a two-fold Kantian argument)

Prima facie, trying to couple the Kantian moral perspective with some aspects related to the present-day debate about emerging technologies, we may first of all refer to two main research areas. Let me try to offer a very schematic picture of them. The first area is related to the concept of *dignity* – and the correlative, comprehensive question might be how such a concept (understood in the specific Kantian definition) could be useful to confront new and emerging challenges related to biotechnologies. The second area

pertains the theory of *autonomy*, which might refer to robotic systems understood in the most wide-ranging sense of the term – as well as processes and procedures related to machine learning.²

In order to try to present and discuss a preliminary map concerning the potential research paths for a comprehensive “Rehabilitierung” of the Kantian thought within the Moral Enhancement Debate,³ we may start from an analysis of a very basic yet fundamental question. May we claim that Kant would be in favor of human enhancement, or may we not? And which definition – if any – of human enhancement would he be able to accept within his theoretical account of the moral sphere?

As a first approximation, we do affirm that Kant could be considered both a *theorist* and a *critic* of a certain idea of *human enhancement*. More specifically, we believe that Kant can be considered a paradigmatic scholar by those who sympathetically adopt a certain definition of *moral enhancement* but also by those who are also skeptical and critical towards any form of enhancement. In accordance with these premises, in the subsequent sections of this chapter, two paths will be explored.

3. The first path: internal enhancement towards asymptotic full adequacy to the moral law

The first theoretical path (and probably the most well known and discussed among Kantian scholars) can be recalled by referring to a specific meaning of the concept at hand: not just “human enhancement” but the very idea of “moral enhancement” could be seen as the asymptotic goal which is legitimized and, to some extent, coherently required by the same Kantian theory referred to as the “pure practical reason”.

Accordingly, a preliminary list of possible key concepts taken from the Kantian moral vocabulary can be gathered here. A *synthetic* list may include the paradigmatic distinctions between *Pflichtmäßig* or *aus Pflicht* (GMS 04: 397–399), *Maxime* and *Gesetz* (GMS 04: 400–404), *Pflichten gegen sich selbst* and *Pflichten gegen anderen*;⁴ the same ideas and the correlative rationale related to *Moralisches Streben*, *Absicht*, *Tugend*, *Höchstes Gut*.⁵ All these concepts relate to processes or to the profile of something that should be done, that is to be achieved first: to something which is possible for each single being capable of reason to insert within a universe of “having to be”, to be constantly put in front of such a being and potentially reached.

Yet all these concepts are emerging from my own *internal* enhancement or – to put the issue in terms more in line with the Kantian universe – from my own internal capability of isolating and evaluating the options available to act, to analyze the possible procedures and different opportunities at my disposal by balancing and comparing them in order to make a decision and, in the end, to act in accordance with the genuine requirement of the pure practical reason.

From a Kantian perspective, to trigger a moral attitude means to generate a new and incremental threshold towards a progressive but always unachievable total adequacy of our motivation to the requirements of the pure moral law – and this by keeping in mind a structural limitation: each step is systematically linked (*systematisch verbunden*) to an individual effort, and none of the steps is taken for granted, nor is its overcoming made easier or more feasible by the habit we can develop to commit ourselves to the respect of the moral law.

In this sense, we have a sort of duty to enhance our moral attitude; as agents capable of reason, we have to positively answer to an implicit reformulation of the categorical imperative: “do increase your own moral perfection as much as it is possible for you”.⁶ Up to this point, we can assume that Kant would have provided a positive answer to the issue raised at the very beginning of this work; as a result, we may affirm that Kant might be included in the class of scholars and theorists who are in favor of human/moral enhancement. This could be the first thesis we can deliver by working within the more traditional conceptual vocabulary pertaining to the Kantian moral perspective: in this sense, Kant is an *ante litteram* theorist of moral enhancement.

4. The second path: the “external” help or integration – and its alleged eligibility

In spite of the conclusion just reached, our assessment may lead to a different result if we consider the same rationale from another perspective. For instance, we may wonder: from a Kantian point of view, would the usage of “something” to help our capability to pursue a moral action or conduct be permissible? In accordance with a first approximation, with “something”, we refer to external items in charge of processing complex decision settings, to drugs and to possible technological devices. As a second approximation, we might maintain that this form of “help” would reduce – up to annihilating – the “moral fight” which has to be “celebrated” within the faculty of will among contrasting motives that are in competition to become the guiding ones for the individual conduct. As we know, from a Kantian point of view, the moral value or virtuousness of my action is not a function of the outcomes or the results that it may obtain. Conversely, the moral value of an action is exactly in proportion to the difficulties that our rational will has met in overcoming our pathological reasons (for example, dictated by the desire for a concrete result, honor, social recognition and so on). Let’s explore this point more carefully in what follows.

4.1. *Simply a hypothetical imperative?*

Starting from these premises, could what we might call an “external help” or “integration” be considered eligible from a Kantian moral stance? *Prima*

facie, if we consider the well-known theoretical framework of Kant's moral theory, we should acknowledge that, in case we admit the possibility of availing ourselves of external help to more adequately pursue a moral action, we do enter into the perspective of the *hypothetical imperative*. We could describe such an argument by using a linear implication: *if* you want to become a rational moral agent, *then* you must profit from (that is, you must accept the help or the integrative support which is coming from) this or that specific device, item or drug.

In accordance with this reasoning, we might come very quickly to a completely different conclusion from the one reached as a result of the path we followed in the previous paragraph: any form (device) that fits within the definition of the so-called "moral" enhancement has to be considered *outside* of the sphere of morality, at least as the latter is conceived from a Kantian perspective.

Nonetheless, should we consider the same rationale within the perspective of the categorical imperative, we might leave the question open to an alternative interpretation. Such an interpretation may be triggered by considering the same research question from a further different point of view: let's imagine that the "help" or "integration" we refer to is something which does not intend to substitute for my own reason or my rational capability to act but just to enhance its functioning or to boost my ability to distinguish between rational and not-rational drivers or motivations (*Triebfedern*). In this case, may such help or integration be considered eligible from a Kantian stance? The case may arise when we need a little support or help in order to make the better decision and to implement the best moral action that is possible for us to adopt. In this case, do we have the possibility to consider this support or help as negative or morally not admissible? In coherence with this approach, then, we should reopen and reshape our question by going more in depth and analyzing the same issue within the entire perspective of the categorical imperative.

4.2. Which enhancement?

In order to clarify our reasoning, a precise idea of what might be considered moral enhancement should be provided. To start framing the issue, we can mention at least three types of moral enhancement.⁷

The first one can be identified as *pharmacological enhancement*.⁸ The expected result of this kind of enhancement, coming from drugs, is to help the subject with what concerns their capability to analyze specific situations and to refine the relative decision-making process. The same result could also be achieved by increasing their capability of isolating and excluding not purely rational moves (e.g. coming from sensitivity or *arbitrium*) from the amount of possible motivations that are *in principio* at their disposal or that should be considered in order to initiate an action or a pattern of actions in coherence with the moral law.

A second category includes a potentially very relevant number of possible technological structures or items. We can label it *smart decision platforms* or *digital moral evaluators*.⁹ This means that we are dealing with devices, or applications (applications or apps), programmed to offer comparative moral evaluations or to realize a multiple reshaping of the same settings in a very short time and from a specific deontological, utilitarian or any other traditional moral perspective. We are referring to very big calculators or to very advanced artificial intelligence systems. Again, they would function as a sort of aid or tool that the subject can activate or deactivate whenever they want and which, ideally, should neither be addictive nor involve any of the side effects that drugs can provoke.

A third category of items that could be isolated offers, in a sense, a further in-between perspective. We are now referring to the cluster of the so-called *neuroprosthetic implants*. With such an expression, we refer to devices to be surgically implanted in our brain (or at least in the brain area) and specifically devoted to providing a person with sensorineural loss (with specific reference to cognitive capability or physical movement control) a partial or full recovery of that specific capability – or its enhancement.¹⁰

This kind of enhancement directly addresses the mind-body relationship. Neuroprosthetic implants can be basically classified into two different but interrelated categories. On the one hand, we can consider implants able to process massive numbers of already-solved moral dilemmas – and the correlative procedures followed to solve them. These would function just as additional memory to the human subject, as an additional recovery system to host a number of “standard” modalities to answer or to react *vis-à-vis* an enormous number of contexts, situations and decisions.

On the other hand, we can refer to implants able to do nothing more or nothing else than recording one’s action or reaction towards specific problematic issues. Such devices may become the best recorder of one’s moral identity by keeping track of the individual moral history, a history of their moral reactions towards specific issues or specific key points or moments in their life. In this case, it would no more be a matter of integrative or additional memory but rather a solution to preserve a vivid memory of what a subject has done and of single parts of an individual unrepeatable life-path.

This way, the individual moral identity would become technologically validated while all their constitutive elements are being selected and stored in accordance with sharp and externally programmed clusters that would guarantee both selective and comprehensive access to such elements.

5. Two problematic areas

Now, keeping this framework in mind, we can articulate a preliminary evaluation. *Prima facie*, we must admit that all the previously outlined categories

(namely drugs, platforms, implants) shall be considered outcomes and products resulting from the work and commitment of other human beings, for example, chemists or engineers, as well as from calculators or artificial intelligence devices that have been programmed by other human beings. In a sense, this means that we can consider ourselves subjects externally driven, externally guided or externally pushed to make certain decisions;¹¹ to be noted is that here we refer to decisions that are within a certain range of variables externally foreseen and settled. This – we can maintain – has nothing to do with individual choices, with decisions about specific actions from the perspective of the *self*.

In case we decide to go more in depth in our reasoning, from a Kantian perspective, we should consider this first evaluation the triggering point of – at least – two further and correlated controversial areas. First of all, by endorsing a positive consideration of the different forms of moral enhancement, we need to ask ourselves whether we are sufficiently aware of the risk of a potential infringement of the logic of autonomy. Second, we should wonder whether we are implicitly facing a potential infringement of the universalization principle. Let us start by considering the first challenging area.

5.1. *Infringing autonomy?*

The main theoretical question of our investigation can be summarized as follows: are any of the typologies of moral enhancement hereby mentioned able to infringe the very logic of *autonomy*, as articulated in Kantian terms? As for the second (*smart decision platforms* or *digital moral evaluators*) and the third type (*neuroprosthetic implants*), we have no doubts: they surely infringe the rationale of individual autonomy.

On the contrary, *pharmacological enhancement* surely deserves a wider analysis. In fact, such form of enhancement appears to leave to the individual agent appropriate room to maneuver to decide by herself, as drugs should be in charge only of offering a more accurate analysis of the present variables and the possible motivations which are at stake in a specific deliberative framework. As a result, in principle, such an enhancement seems not to be dismissible. Nonetheless, we still have to deal with new formulations of classic issues.

The first one regards the so-called “informed consensus”: how accurate and how strongly linked to a high standard of information is the consensus we are (explicitly and implicitly) giving by using such forms of pharmacological enhancement? Second, and consequently: Do we have a good idea, delivered by public protocols or experimental settings, of the collateral (short- and long-term) effects on individual behavior or personalities that result from prolonged usage of them? And third, does such usage question the very idea of dignity which is systematically linked to the idea of individual subjects as capable of reason?

Let us then introduce a closer concentric circle, focusing on single aspects of the individual path that drive to action. Would we be able to isolate a specific point within the process delivered by the faculty of desire – related to the selection of possible motivations (*Triebfedern*) or devoted to the capability of articulating maxims in accordance to the categorical imperative – in which the insertion of pharmacological enhancement might not infringe human dignity as well as the principle of autonomy of the human subject?

As for the present context, in the following section of this chapter, we will focus on the first fundamental aspect related to the faculty of desire, which we will do by recalling the systemic distinction delivered by Kant between *Autonomie des Willens* and *Heteronomie der Willkür*. In order to present such a distinction and the correlative conceptual issues it raises, two important excerpts will be analyzed. The first one is taken from the *Critique of Practical Reason* and the second from *The Metaphysics of Morals*.

Autonomy of the will [*Autonomie des Willens*] is the sole principle of all moral laws and of duties in keeping with them; *heteronomy of choice* [*Heteronomie der Willkür*], on the other hand, not only does not ground any obligations at all but is instead opposed to the principle of obligation and to the morality of the will. That is to say, the sole principle of morality consists in *independence from all matter of the law* (namely, from a desired object) and at the same time in the determination of choice [*Bestimmung der Willkür*] through the mere form of giving universal law that a maxim must be capable of. That *independence* [*Jene Unabhängigkeit*], however, is *freedom in the negative sense*, whereas this *lawgiving of its own* [*eigene Gesetzgebung*] on the part of pure and, as such, practical reason is freedom in the *positive sense*. Thus, the moral law expresses nothing other than the *autonomy* of pure practical reason, that is, freedom, and this is itself the formal condition of all maxims, under which alone they can accord with the supreme practical law. (KpV 05: 33)

The faculty of desire in accordance with concepts [*Das Begehrungsvermögen nach Begriffen*] . . . insofar as it is joined with one's consciousness [*Bewußtsein*] of the ability to bring about its object by one's action it is called *choice* [*Willkür*]. . . . The faculty of desire whose inner determining ground, hence even what pleases it, lies within the subject's reason is called the *will* [*Wille*]. The will is therefore the faculty of desire considered not so much in relation to action (as choice is) but rather in relation to the ground determining choice to action [*zur Handlung*]. (MS 06: 213; my translation)

In this context, the following distinctions will be made with reference to the main topic. When we speak about the “faculty of will” using a Kantian vocabulary, we must preliminary distinguish between *Wille* and *Willkür* – namely between *will* and *choice* (or *arbitrium*). Subsequently, we need to apply to both distinctions the two-fold reference to *heteronomy* and *autonomy*, which characterizes our actions as *within* or *outside* the sphere of the pure practical reason in Kantian terms. While in the *Metaphysics of Morals*, both *autonomy* and *heteronomy* are related by Kant to the faculty of *will* (*die Wille*), in the *Critique of Practical Reason*, the concept of *heteronomy* is related to that of *choice* (*der Willkür*), whilst the entire perspective of *autonomy* is exclusively related to the *Will*. Consequently, we should wonder which is the specific meaning of freedom (positive or negative, first) to which we should link the enhancement rationale.¹²

In the end, if we jump back to the previous question (that is, “which kind of pharmacological enhancement can be legitimately admitted within the Kantian moral universe?”), we have to clarify whether we are looking for a kind of enhancement which is focused upon the enforcement of individual *autonomy* or alternatively on the reduction of the attractive force of *choice*. Then we need to revise the borders and the distinction between *autonomy* and *heteronomy* in order to be clearer when referring our action to the specific definition of freedom that we want to realize through our agency.

5.2. *Infringing the universalization principle?*

By referring again to the entire controversial framework previously mentioned, let us now consider the second challenging area, which is related to the potential infringement of the universalization principle.

Prima facie, this area constitutes a sort of even narrower concentric circle related to the very perspective of *autonomy*. As we know from the third formulation of the categorical imperative,¹³ we have to consider every being capable of reason, that is, an agent ideally able to create a universal legislation while acting. This is the description of the concept of autonomy, with a more accurate formulation, namely: “the principle of every human will as a will legislating universally through all its maxims” (GMS 04: 432).

Let’s imagine integrating this capability to act and to legislate (in such a way that only a universal will would be able to do) through a pharmacological enhancement. Furthermore, let’s leave aside the issue we already raised with regard to the maintenance of autonomy in case we start using drugs to better evaluate, decide, and, in the end, act. Rather let’s start by considering pharmacological enhancement as a sort of vaccination which is universally guaranteed. In principle, every being capable of reason would have a stable entitlement, a universal (moral) right to have access to such form of enhancement.

Now, let's play the game of applying such a perspective in our real world. When we enter "the realm of the phenomenon", we could be forced to admit that at least a part of humankind might be excluded from this form of enhancement. Also, we can easily imagine that this inclusion/exclusion process would have the same trend as in the case of the right to health. In other terms, the spreading of pharmacological enhancement would likely proceed in coherence with the same dynamics that led to the exclusion of relevant percentages of the population in Western countries and of the majority of people in areas like the African or Asian continents from the basic levels of healthcare coverage (or at least to severe limitations of it).

6. A skeptical balance

Incidentally and from a comprehensive perspective, by admitting the realistic possibility of an external enhancement, we might wonder whether we are losing once and for all the possibility of building up a *pure moral philosophy* (*eine reine Moralphilosophie*) (GMS 04: 389), namely a moral philosophy which excludes any empirical element from the very outset. As we know, only this kind of moral philosophy was envisaged by Kant as the final purpose of his entire theoretical building devoted to the practical philosophy.

Furthermore, from a systemic perspective, we might still wonder if perhaps with our reasoning we are implicitly legitimizing a new divide. In other words, we need to ponder whether we are aware of the risks which result from our reshaping of the very idea of a "capable-of-reason being", thus possibly paving the way to a sort of "two-speed population" in the real world. Let's imagine that just a portion of humankind has daily access to such an enhancement, while a second (and not minoritarian) part of it will be simply or *per principio* excluded from it. In a sense, we could be forced to admit that universal access to such a form of enhancement is realistically not possible. In other words, we should admit that we are on the verge of creating a new and probably dangerous "moral divide", namely a divide that increases the number of the divides already operating, like the digital, technological, social and educational ones, together with any other that contributes to rendering the expression "humankind" just a rhetorical word rather than one with a pragmatic significance.

In this sense, after the brief exploration of a possible evaluation of the entire perspective opened up by moral enhancement hereby outlined, we can provisionally conclude with a statement of skepticism.

On the one hand, Kant is clearly in favor of the promotion of human and "moral enhancement" by reshaping the very meaning of the *Verbesserung* of humankind and asking every single agent to do everything possible to promote humanity in herself. Nonetheless, on the other, "everything possible" does not seem to foresee the possibility of any empirical or artificial help,

that is, support coming from someone or something else other than every being-capable-of-reason who is in struggle with herself, in search of a positive answer to the categorical and universal issue raised by the moral law.

Notes

- 1 As for the Kantian agenda, let me recall here at least Pirni 2000, 2006, 2015, 2016; Capasso and Pirni 2021. As for the agenda related to emerging technologies and human enhancement, see Koops and Pirni 2013; Pirni and Carnevale 2013, 2014; Pirni 2017, 2019.
- 2 The problematic conceptual linkage between autonomy and robotic systems is one of the most intriguing discussed in the present debate. Just to offer a preliminary framing of the issue: Stradella et al. 2012; Abbass et al. 2018; Lawless et al. 2017; Mecacci and Santoni de Sio 2020.
- 3 In this sense, we are alluding here to a sort of possible prolongation of the debate known as the “*Rehabilitierung der praktischen Philosophie*” (the rehabilitation of practical philosophy). As we know, such a debate started from the opportunity of rehabilitating two classical normative perspectives in practical philosophy by re-interpreting the theoretical path opened up by Aristotle and Kant. This theoretical account was articulated basically to counteract the (at that time) dominating analytical perspective in moral and political domain. By proposing a further line of interest along which Kantian practical thought might be investigated as a comprehensive toolbox able to play an orientative and evaluative role towards emerging technologies and the human enhancement challenge, we would like here to enlarge the picture of the possible theoretical usages of the rationale of such a debate by coupling the same attempt I have been trying to perform by rehabilitating the role of the Aristotelian theory of justice and equality with reference to the human enhancement debate. On this point, let me recall: Pirni 2013b, 2014, 2019.
- 4 Kant focuses on this fundamental distinction in a large part (*Doctrine of the Elements of Ethics*) of the “Metaphysical First Principles of the Doctrine of Virtue” in the *Metaphysics of Morals* (MS 06: 421–447).
- 5 As we know, those now recalled are concepts that Kant develops along the entire group of writings gathered under the label of “practical philosophy”. The context in which they receive the most systematic explanation within the moral understanding of the terms is the *Analytic* (specifically for the concepts of *moralisches Streben* and *Absicht*) and the *Dialectics* (as for the concepts of *Tugend*, *höchstes Gut*) of the *Critique of Practical Reason*.
- 6 Some important textual basis for this kind of imperative can be found in Kant’s *Metaphysics of Morals* (MS 06: 391–392, 444–446). On this specific point, see Pinzani 2018, spec. pp. 122–123.
- 7 A comprehensive critical discussion about the topic is offered by Fimiani et al. 2004; Domingues 2012; Henry 2013; Braidotti 2013; Battaglia and Carnevale 2014; Clarke et al. 2016; Pirni 2017; Totaro 2018.
- 8 For a comprehensive framing on the main issues within an increasingly complex debate, see Cacic 2009; Maslen et al. 2014; Ricci 2020.
- 9 A preliminary framing is offered in Earp et al. 2017; Cokely 2017.
- 10 For a preliminary orientation within another complex debate, see Glannon 2017; Cinel et al. (2019); Lavazza 2019.
- 11 In more Kantian terms, we might say “heteronomously driven” – as in the case of pathological motives.
- 12 I extensively explored this issue in Pirni 2013a.

- 13 “[T]he idea of the will of every rational being as a universally legislative will” (GMS 04: 432).

References

- Abbass, H.A., Scholz, J., and Reid, D.J. (eds.). 2018. *Foundations of Trusted Autonomy*. Cham: Springer.
- Battaglia, F., and Carnevale, A. (eds.). 2014. “Reframing the Debate on Human Enhancement, *Humana.mente*.” *Journal of Philosophical Research* 7(26) (special issue).
- Braidotti, B. 2013. *The Posthuman*. London: Polity Press.
- Cakic, V. 2009. “Smart Drugs for Cognitive Enhancement: Ethical and Pragmatic Considerations in the Era of Cosmetic Neurology.” *Journal of Medical Ethics* 35(10): 611–615. doi: 10.1136/jme.2009.030882.
- M. Capasso, and A. Pirni. 2021. “On the Kantian distinction between Ethics and Right: Linearity and Mediation.” *Con-Textos Kantianos*, 13: 344–364. doi: 10.5281/zenodo.4899211.
- Cinel, C., Valeriani, D., and Poli, R. 2019. Neurotechnologies for Human Cognitive Augmentation: Current State of the Art and Future Prospects.” *Frontiers in Human Neuroscience* 31. <https://doi.org/10.3389/fnhum.2019.00013>.
- Clarke, S., Savulescu, J., Coady, T., Giubilini, A., and Sanyal, S. (eds.). 2016. *The Ethics of Human Enhancement. Understanding the Debate*. Oxford: Oxford University Press.
- Cokely, E.T. 2017. Informing Ethical Decision Making. In L. Syd, M. Johnson, and K.S. Rommelfanger (eds.), *The Routledge Handbook of Neuroethics*. Abingdon: Routledge.
- Domingues, I. (ed.). 2012. *Biotechnologies and the Human Condition*. Belo Horizonte: Editora UFMG.
- Earp, B.D., Douglas, T., and Savulescu, J. 2017. Moral Neuroenhancement. In L. Syd, M. Johnson, and K.S. Rommelfanger (eds.), *The Routledge Handbook of Neuroethics*. Abingdon: Routledge, 166–184.
- Fimiani, M., Gessa Kurotschka, V., and Pulcini, E. (eds.). 2004. *Umano, post-umano. Potere, sapere, etica nell’età globale*. Roma: Editori Riuniti.
- Glannon, W. 2017. *Brain Implants. Implications for Free Will*. In L. Syd, M. Johnson, and K.S. Rommelfanger (eds.), *The Routledge Handbook of Neuroethics*. Abingdon: Routledge.
- Henry, B. 2013. *Dal Golem ai Cyborgs. Trasmigrazioni nell’immaginario*. Livorno: Belforte.
- Koops, B.-J., and Pirni, A. (eds.). 2013. “Ethical and Legal Aspects of Enhancing Human Capabilities Through Robotics.” *Law, Innovation and Technology* 5(2) (special issue).
- Lavazza, A. 2019. “Transcranial Electrical Stimulation for Human Enhancement and the Risk of Inequality: Prohibition or Compensation?” *Bioethics* 33: 122–131.
- Lawless, W.F., Mittu, R., and Russell, D.S.S. (eds.). 2017. *Autonomy and Artificial Intelligence: A Threat or Savior?* Cham: Springer.
- Maslen, H., Faulmüller, N., and Savulescu, J. 2014. “Pharmacological Cognitive Enhancement. How Neuroscientific Research could Advance Ethical Debate.” *Frontiers of Systemic Neuroscience* 8: 107. doi: 10.3389/fnsys.2014.00107.

- Mecacci, G., and Santoni de Sio, F. 2020. "Meaningful Human Control as Reason-Responsiveness: The Case of Dual-Mode Vehicles." *Ethics and Information Technology* 22: 103–115 <https://doi.org/10.1007/s10676-019-09519-w>.
- Pinzani, A. 2018. "Há um dever de ser um deus? Sobre o imperativo moral de ter uma mente brilhante num corpo perfeito." In R. Crisp, D. Dall'Agnol, and M. Consenso Tonetto (eds.), *Ética aplicada e políticas públicas*. Florianópolis: Editora UFSC, 119–136.
- Pirni, A. 2000. *Il regno dei fini in Kant. Morale, religione e politica in collegamento sistematico*. Genova: il melangolo.
- Pirni, A. 2006. *Kant filosofo della comunità*. Pisa: ETS.
- Pirni, A. 2013a. "Freedom of the Will in Communitarian Perspective." In S. Bacin, A. Ferrarin, C. La Rocca, and M. Ruffing (Hrsgs.), *Kant und die Philosophie in weltbürgerlicher Absicht. XI. Internationaler Kant-Kongress/Kant and Philosophy in a Cosmopolitan Sense. XI International Kant Congress*. Berlin: De Gruyter, Vol. 3, 509–520.
- Pirni, A. 2013b. "Il potenziamento umano come sfida di equità per la società tecnologica. Un percorso aristotelico." *Archivio di Filosofia* 81(3): 115–122.
- Pirni, A. 2014. Towards a 'Human Enhancement Society'? Opportunities for an Aristotelian Approach to Frame the Question, *Humana.mente.*" *Journal of Philosophical Research* 7(26): 239–252.
- Pirni, A. (ed.). 2015. "Kant y las declinaciones de la armonía." *Con-Textos Kantianos – International Journal of Philosophy* 2 (special issue).
- Pirni, A. 2016. "The Place of Sociality. Models of Intersubjectivity According to Kant." In A. Faggion, A. Pinzani, and N. Sanchez Madrid (eds.), *Kant and Social Philosophy*. London: Palgrave Macmillan, 65–92.
- Pirni, A. (ed.). 2017. "Il post-umano realizzato. Orizzonti di possibilità e sfide per il nostro tempo." *Nuova Corrente* 59(159).
- Pirni, A. 2019. "Il potenziamento umano nel solco della riabilitazione della filosofia pratica." In L. Cimmino, L. Alici (a cura di), *Mutamenti. Nuovi ambienti, nuove responsabilità*. Perugia: Aguaplano, 15–36.
- Pirni, A., and Carnevale, A. (eds.). 2013. "Robotics and Public Issues: A Philosophical Agenda in Progress." *Cosmopolis* 9(2) (special issue).
- Pirni, A., and Carnevale, A. (eds.). 2014. "Investigating the Relationship between Future Technologies, Self and Society." *Politica & Società* 3(2) (special issue).
- Ricci, G. 2020. "Pharmacological Human Enhancement: An Overview of the Looming Bioethical and Regulatory Challenges." *Frontiers in Psychiatry* 11: 53 doi: 10.3389/fpsy.2020.00053. www.ncbi.nlm.nih.gov/pmc/articles/PMC7037021/ (accessed on October 2020).
- Stradella, E., Salvini, P., Pirni, A., Di Carlo, A., Oddo, C.M., Dario, P., and Palmerini, E. 2012. "Subjectivity of Autonomous Agents. Some Philosophical and Legal Remarks." In O. Boissier, G. Bonnet, and C. Tessier (eds.), *Rights and Duties of Autonomous Agents. Proceedings of the 1st Workshop on Rights and Duties of Autonomous Agents, in conjunction with the 20th European Conference on Artificial Intelligence – ECAI 2012*. Montpellier: CEUR Workshop Proceedings, 24–31.
- Totaro, F. (ed.). 2018. *Il corpo e le sue trasformazioni*. Brescia: Morcelliana.

SPIELRAUM

Narrow and wide duties and their consequences

Claudio La Rocca

*Alle freye Handlungen sind nicht durch die Natur und durch
kein Gesetz bestimmt, also ist die Freyheit was schreckliches,
weil die Handlungen nicht determiniret sind.*

(V-Mo/Kaehler(Stark), 31)

Philosophy deals with concepts. Philosophical reflection on concepts is not merely aimed at attaining clarity, however. Indeed, its aim is sometimes achieved by assessing the adequacy of our concepts: by weighing them up and changing them when necessary, resulting in the creation of new ones. This is not merely an intellectual matter; concepts not only enrich our knowledge but can also influence our actions. This applies especially to those ethical terms that guide, orient and even determine our lives. In this respect, the “narrow duty”/“wide duty” concept pair,¹ to which Kant attaches great importance in his moral philosophical reflection, emerges as significant. Kant’s confrontation with this pair of concepts and with their associated conceptual constellation is particularly interesting insofar as it clearly illustrates the special nature of philosophical work on basic concepts. In particular, it provides a clear example of how existing concepts are reinterpreted and reformulated in philosophy: rather than taking place in empty space, philosophical thinking flows within conceptual traditions of which even the most creative thinking is not free.

In this (sometimes difficult) dialogue with tradition, efforts to shed light must contend not only with the shadows of the past but also with the new spots of darkness that form when the light is shifted in novel directions. Thus, in his attempt to draw on a body of thought from the philosophical tradition while simultaneously changing it, Kant encountered difficulties that he could not completely overcome, making it difficult to accomplish his goal of providing a new and important systematization of ethical concepts. The quite complicated and subtle problems that we will encounter in this chapter

tell us something not only about the moral philosophical content with which Kant had to contend but also about moral philosophical reflection as such.

Kant deals early on with the distinction between narrow and wide duties and with the associated further distinction between perfect and imperfect duties, which is considered “the most controversial of Kant’s divisions of moral principles” (Hill 1971, 55). The latter, as we will see, is actually the source of the former and an element of a broad philosophical tradition. The narrow/wide distinction is introduced by Kant and becomes more important – at least as far as the published works are concerned – in a later phase of his thinking, along with the development of the *Doctrine of Virtue*, which was published in 1797 as the second part of *The Metaphysics of Morals*. It is precisely in this work that we find passages that pose serious interpretative problems, however, indicating that these concepts may not be sufficiently developed in Kant’s work.

The relationship between the two pairs of concepts mentioned previously – namely between narrow and wide duties on the one hand and between perfect and imperfect duties on the other – is particularly problematic. The treatment of both conceptual pairs results in a puzzle or riddle of sorts. I will not pretend to solve this riddle here; instead, my analysis will point the way toward a new way of considering the problem. I will concentrate here on the main features of Kant’s argumentation without pretending to touch on the full range of interesting issues that it raises.

1. Perfect/Imperfect, Narrow/Wide

As mentioned previously, the text in which the relevant problems of interpretation arise is *The Metaphysics of Morals*. With this said, the first place in Kant’s works where the narrow/wide distinction occurs, together with the perfect/imperfect distinction, is the *Groundwork of the Metaphysics of Morals* (1785). Here, Kant directly connects these distinctions to his famous four examples, which he introduces as follows:

We shall now enumerate some duties, according to their usual division, into duties to ourselves and to other human beings, into perfect and imperfect duties.

(GMS 4: 421)²

Only perfect and imperfect duties are mentioned here. Nevertheless, as we will see, the connection between this pair of terms and the narrow/wide concept pair that will be discussed shortly thereafter is already established in this text. At the same time, the “duties to oneself/to others” pair also plays a role.

A first definition is provided in a footnote: a perfect duty is “one that allows no exception to the advantage of inclination”. At the same time,

Kant notes that in his view, there are also perfect inner duties – “which runs counter to the use of the word adopted in the schools” (GMS 4: 421).

The four examples are well known, but it is useful to briefly recall them. The first concerns a person who is tempted to take his own life. The second concerns the duty to repay borrowed money. The third has to do with the duty to “expand and improve” (GMS 4: 423) one’s fortunate natural predisposition, that is, to develop one’s own talents. The fourth raises the question of the duty not only to refrain from harming others but also to actively help them.

The first and third examples are, of course, examples of duties to oneself; the second and fourth are examples of duties to others. According to Kant’s classification, the first two are cases of perfect duty, whereas the third and fourth are cases of imperfect duty. This results in the following schema:

		<i>Perfect duties</i>			<i>Imperfect duties</i>
<i>to oneself</i>	1	not to take one’s own life	3		to develop one’s talents
<i>to others</i>	2	to keep promises	4		to help others

When Kant comments on these examples, especially with regard to the possibility of deriving duties from the categorical imperative, he uses the narrow/wide pair. The way in which the duties are established is decisive. In the case of suicide and the case of borrowed money, the moral judgment leads to the result that the maxim of this act “cannot even be *thought* without contradiction as a general law of nature” (GMS 4: 424); in the case of the development of one’s own talents and helping others, on the other hand, there is no “inner impossibility”: the impossibility attaches to *willing* “that their maxim be elevated to the universality of a law of nature” (GMS 4: 424). Not keeping promises destroys the very concept of promising: a maxim that entails this is unthinkable as a general maxim. On the other hand, I can think of a maxim according which there is no obligation to help others: I cannot, however, *want* this maxim. “I cannot want”, we read in the *Mrongovius II* lecture notes from the same period (1784–5), “that lovelessness would become a general law, for in that case I also suffer myself” (V-Mo/Mron II, 29: 609). Only after this clarification do we encounter the narrow/wide distinction:

It is easy to see that the first conflicts with strict or narrower (unrelenting) [*strengen oder engeren (unnachlässlichen)*] duties, the second only with wider (meritorious) [*weiteren (verdienstlichen)*] duties.

(GMS 4: 424)

In this context, Kant is not particularly interested in the division of duties, which, he says, he reserves for a future metaphysics of morals. Nevertheless, he notes that this division is not introduced merely instrumentally, “to order [Kant’s] examples” (GMS 4: 421, note). On the contrary, Kant clearly announces an argumentative goal prior to his classification and examples:

If from this one imperative all imperatives of duty can be derived as from their principle, then, even though we leave it unsettled whether what is called duty is not as such an empty concept, we should at least be able to indicate what we think by it and what the concept means.

(GMS 4: 421)

The success of the division as a consequence deriving from the imperative is therefore proof that the latter can “explain” the different kinds of duty, and thus that it can be thought of as their principle. After listing his examples in conformity with the division of the duties, Kant claims that the following result has been achieved:

We have thus established at least this much, that if duty is a concept that is to contain significance and actual legislation for our actions it can be expressed only in categorical imperatives, but by no means in hypothetical ones; likewise we have – and this is already a lot – presented distinctly and determined for every use the content of the categorical imperative, which would have to contain the principle of all duty (if there were such a thing at all).

(GMS 4: 425)

Concerning the narrow/wide distinction and its relation to the perfect/imperfect distinction, it should be noted that the definition of perfect duties provided here could be viewed as “misleading” (Timmermann 2007, 80) – or even, following Wolfgang Kersting, as “a logical absurdity” (Kersting 1993, 189). According to the previously mentioned definition, Kant understands a perfect duty as “one that allows no exception to the advantage of inclination”. The violation of a duty is nevertheless defined precisely as the act of making an exception in favor of one’s inclination (GMS, 4: 424).³ According to Jens Timmermann, Kant here means that “imperfect duties initially permit trade-offs between morally worthy ends (or ‘grounds of obligation’)” (2007, 80).⁴ We will return to this later. In any case, what is clear in the *Groundwork* is that there seems to be a strict correspondence between the narrow/wide distinction and the perfect/imperfect distinction. Narrow duties are perfect; wide duties are imperfect. In this respect, a more precise definition could be obtained by allowing the characteristics of both terms to flow together. In

the *Groundwork*, a narrow duty is a duty the maxim of which “cannot even be *thought* without contradiction as a general law of nature” (4: 424) – one which allows for “no exception to the advantage of inclination” (4: 421) and which is *streng*, “unrelenting” (4: 424), not “meritorious”. In any case, as Kant says, the distinction concerns “the kind of obligation (not the object of their action)” (4: 424).

As is evident, the *Groundwork* is not particularly illuminating when it comes to this distinction. Apart from its misleading characterization of perfect duties, it merely conveys that in the case of narrow duty, the very *conceivability* of the conflicting maxim is impossible, whereas with wide duties, it is impossible to *want* the maxim. The consequences of this are not made clear here. A further characterization can be drawn from the examples: perfect duties seem to *forbid* certain actions⁵ – albeit indirectly, by rejecting certain maxims. Taking one’s own life and not repaying borrowed money prove morally impermissible.⁶

2. The Puzzle

The importance of this distinction becomes clearer in light of these few hints. What is at stake is basically whether (and how) our moral life moves in two dimensions. The first dimension is guided by rules of action that determine with noteworthy precision what one morally ought to do; the second dimension is guided by rules which, on the contrary, give moral agents a certain amount of “leeway” (in German, “*Spielraum*”, a term that is crucial for Kant) in which they can and should consider different options. In our everyday, non-philosophical moral thinking – which nonetheless has a long ethical and religious tradition behind it – we distinguish between prohibitions and commandments and thus between bad deeds, which are defined precisely enough, and good behavior, which serves a good purpose but can still be manifested in different actions. Philosophical thinking has a long history of conceptual distinctions, one strand of which traces back to Cicero, flowing into the pre-Kantian philosophy of the 18th century, with its widespread and variously interpreted distinction between perfect and imperfect duties. This terminology was so ubiquitous that, as late as 1780, Ludwig Julius Friedrich Höpfner, a law professor in Gießen, claimed that “even those who do not know much else” are aware of the distinction between perfect and imperfect duties (201).

Kant adopts this terminology (he speaks, as we have seen, of the “usual division”, GMS 4: 421), but he also complicates it by introducing the narrow/wide distinction. This new terminology had been used with some hesitation in the *Groundwork*, where Kant spoke of “strict or narrower (unrelenting) duties” and “wider (meritorious) duties”. The conceptual pair makes a striking appearance in *The Metaphysics of Morals* but not necessarily with clearer contours. Kant introduces the narrow/wide distinction with much solemnity in the title of paragraph VII of the introduction to *The Doctrine of Virtue*:

“Ethical Duties Are of Wide Obligation, Whereas Duties of Right Are of Narrow Obligation” (MS 6: 390). This statement does not offer a definition, but it does provide a clear assignment of the narrow and wide duties to law and ethics, respectively. Nevertheless, this clarity in approach does not correspond to a comparable clarity in execution. *The Doctrine of Virtue* indeed also contains a distinction (also highlighted in the paragraph’s title) between perfect and imperfect duties – one which takes place entirely within ethics. The first part of the “Doctrine of the Elements of Ethics” contains a first book entitled “Perfect Duties to Oneself” and a second book entitled “On Man’s Imperfect Duties to Himself”. Perfect duties to oneself include the prohibition against “self-disembodiment” (*Selbstentleibung*),⁷ “defiling oneself by lust” (*wohllüstige Selbstschändung*) and “stupefying oneself” (*Selbstbetäubung*) (duties to oneself as an animal being), as well as lying, avarice, false humility and the commandment of self-knowledge (duties to oneself as a moral being). Imperfect duties include developing and increasing one’s natural perfection and increasing one’s moral perfection. Strangely enough, no analogous division is proposed for duties to others, although the perfect/imperfect distinction was also applied to this kind of duty in the *Groundwork*.

Thus we have two divisions which are made according to the pairs of concepts narrow/wide and perfect/imperfect and which raise a problem. If these pairs of concepts are related such that, even if they do not equate to each other, they at least imply each other (such that what is narrow must be perfect, and what is wide must be imperfect), this results in an inconsistency: certain ethical duties which are exclusively wide according to § VII must also be perfect, that is, narrow.

This inconsistency has been noted and much discussed in the literature.⁸ From a purely formal point of view, the problem can be resolved in various ways: by redefining or omitting one or the other pair of concepts, by rejecting their correspondence, by introducing further distinctions and so on. Of course, all of these solutions depend on an interpretation of the terms in question. So far, we have only briefly considered their function in the classification of duties, but we have yet to say anything about their definition.

3. Law, Ethics, and Duties

In both *The Metaphysics of Morals* and the *Groundwork*, the connection between the narrow/wide and perfect/imperfect distinctions must be inferred rather than gleaned directly from Kant’s words. The text in which this connection is addressed most directly is the preliminary work (*Vorarbeiten*) for the *Doctrine of Virtue*, which was published in Volume XXIII of the Academy Edition.

It is an open question whether these *Vorarbeiten* contain Kant’s final conception or simply a provisional view, which was then partly revised in *The Metaphysics of Morals*.⁹ In any case, the principle grounding both distinctions

is explicitly mentioned in this text, following a line of argumentation that corresponds in its main features to that found in *The Metaphysics of Morals*. Paragraph VI of the introduction to *The Doctrine of Virtue*, which is the premise of the distinction between narrow and wide duties (as stated in paragraph VII), claims that “Ethics . . . Does not Give Laws for *Actions* (*Ius* does that), but Only for the *Maxims* of Actions” (MS 6: 388). The same is said in the *Vorarbeiten*, which adds the idea of indeterminacy. The principle of the duties of virtue, Kant writes, “does not command actions whose maxim can be universally legislative, but leaves them indeterminate; it commands, on the contrary, the maxim of a certain kind of action” (VAMS 23: 391). This indeterminacy derives not from a defect but from an additional condition, namely the requirement that the action be not only in accordance with the law (the principle of freedom under general laws, *the principle of right*) but also based on a (morally tested) maxim. The moral – not merely legal – value of the action refers to the level of maxims and in this respect to another layer of the relationship between the particular (the action) and the general, which goes beyond the standard situation of the application of a (general) rule to a (special) case. Actions are directly regulated not by a law but by a maxim, which is the real object of the prescription.

This is how the indeterminacy, that is, the “wide” nature of duties, comes into play: the principle, Kant writes, can

only determine actions in general, not precisely, and the necessitation, which must be found in every duty, concerns only the way of thinking (maxim), while the capacity of choice is given leeway [*Spielraum*], only if it does not specifically change the way of thinking and its principle.

(VAMS 23: 391)

This *Spielraum* is limited by one’s way of thinking, or *Gesinnung*, which nevertheless includes a plurality of possible actions. According to the quoted passage, the maxim commands “a certain kind of action”; the text states in the following lines that morality “only sets maxims for the genre of action”, such that “there remains a certain leeway [*Spielraum*] for freedom to determine them” (VAMS 23: 393).¹⁰

This further generality of the ethical prescription can also be characterized by the directly related concept of purpose. The ethical maxim defines purposes, not actions. Kant continues: “Such a maxim, which determines nothing in terms of actions (in terms of type and degree), is a maxim of ends” (VAMS 23: 393).

Thus the thesis according to which all ethical duties are wide duties seems to be clearly expressed and justified here. Since they concern a second level of generality, that is, maxims (which set out a type or genre of action), and thus embrace and allow for a plurality of possible actions, they leave room

for maneuver. This margin is not the flexibility that the power of judgment always uses when applying a rule; it is rather the *Spielraum* that is opened up by the further layer of generality introduced by the maxim. It is this additional scope that narrow, legal duties lack.¹¹

In the *Vorarbeiten*, the link between the narrow/wide and perfect/imperfect distinctions is perhaps expressed more sharply than in other texts. Consider, for example, the following passage:

In all duties, however, one should also consider the obligation, i.e. the moral necessitation that concerns the form of the obligation, whether it is perfect or imperfect, determining *stricte* or *late*.

(VAMS 23: 393)

Kant is even clearer a few lines further, where both terms are connected with the justification or the difference we have emphasized, namely with the relation to the maxim and not to the action. The text reads:

But if the law does not directly command the action, but only the *maxim* of the action, if it leaves to the free judgment of the subject the way in which and the measure of the degree to which the command should be carried out, [if it commands] only that it is necessary to do what it commands as much as is possible for us under the given conditions, then the obligation is imperfect and the law is not narrowly but only widely binding, *late obligans*.

(VAMS 23: 394)¹²

Together with their necessary connection, this passage highlights the slightly different connotations of these terms. “Perfection” is a property that characterizes a kind of obligation. As Wolfgang Kersting emphasizes (1993, 187), this corresponds to the conventional use of the word in pre-Kantian philosophy. For Christian Wolff, that *obligatio* was considered *imperfecta*, “cui ut satisfiat nemo cogi potest” (1763, I, III, § 80, 39), which no one can be forced to fulfill. Accordingly, for Johann George Sulzer, a perfect obligation implied the possibility that one could be compelled by force to do the required action and thus that the duty “could be made into a law” (1773, 396). According to Kant’s previous lecture on moral philosophy in 1777, an obligation is perfect when external constraint is added to the inner obligation.¹³ Kant later departs from this characterization (enforceability becomes more a result of the direct relationship of duty to action), but “perfect” remains conceived as a property that belongs to obligation. Enforceability was also connected to the character of *Schuldigkeit*, the property of owing something to someone else, which is also part of Kant’s concept of perfect duty.

From the quoted text from the *Vorarbeiten*, we can see that if the obligation is imperfect, the law is also *late obligans*, that is, wide with regard to obligation.

The Metaphysics of Morals does not appear to contradict this view. The two terms seem to describe two aspects of the same phenomenon but from two different perspectives. The perfect/imperfect distinction, which at first referred to enforceability – which is only possible where *actions* are determined – shifts its focus in Kant from the conditioned to the condition, as it were: namely from enforceability (which is a consequence) to the property of being directed to the action (if the action is targeted, the obligation is enforceable). The narrow/wide pair, in turn, describes the relationship between the rule and the action and its consequence: if the relationship is immediate, the action is determined, and the obligation is in this sense narrow; if the relationship is indirect because the regulation concerns the nature of the maxim (or the purpose) – and only through this the action – then the obligation is wide, because a plurality of actions can correspond to the same maxim or serve the same purpose.

Thus far, I have deliberately bracketed other elements that play a role in the conceptual definition of the terms so that the structure of the terms becomes clearer. These include the concepts of negative and positive duty, commandment and prohibition, commitment and omission, guilt and merit, moral unworthiness and vice and the distinction between maxims of action and maxims of purpose. Indeed, in different ways, these all seem to be a consequence of the distinction between perfect and imperfect duties, which has its origin in the Kantian determination of the distinction between law and ethics.

4. Possible Solutions

The core structure of the distinction is therefore at odds with the conventional (albeit diverse) principles of differentiation in pre-Kantian philosophy because – as I have repeated often enough previously – it is based on the relation to actions and maxims that Kant introduces as the basic idea of his moral theory. If the ethical value of an action lies not in the nature of the action but in the inner disposition (*Gesinnung*) which is expressed in the maxim, a double distance between rule and case is introduced. The texts I have quoted, as well as many passages from *The Metaphysics of Morals*, show that Kant's train of thought moves in a direction that focuses on actions or maxims, which justifies the perfect/imperfect characterization and thus determines the narrow/wide characterization and which, as a consequence, essentially leads to the concept of "leeway". This, of course, exacerbates the riddle or "puzzle" of the existence of perfect duties to oneself, however, which should be part of ethics even though they are perfect duties.

If the core structure is that highlighted previously, then we can reject Onora O'Neill's view, according to which the narrow/wide distinction applies from two perspectives. In her account, it applies with regard to the obligation ("in obligation") depending on whether acts/omissions or purposes are commanded; this in turn depends on the derivation of the duty. By contrast, the distinction "in requirement" depends on the circumstances, in particular on

whether acts/omissions or the pursuit of certain policies are required to fulfill the duty. Thus, perfect ethical duties can be wide in obligation because they provide purposes but narrow in requirement since these purposes can be fulfilled by actions – and only by certain actions or omissions. For example, the duties not to take one’s own life and to avoid self-stupor do have a purpose (but which one?), but they require specific omissions insofar as it is only through such omissions that they can be achieved.

To put this into a coherent framework, the term “perfect” should be identified with the second meaning of “narrow”, that is, narrow in requirement. O’Neill is not entirely convinced of the possibility of this *identification* (although she allows that the two distinctions may *coincide*) and tends to think that the relevance of the perfect/imperfect distinction should be put in question and that the pair of terms is unnecessary, or at least not worthy of further investigation (2013, 121 ff).

A possible solution to the contradiction that arises with the concept of ethical *and* perfect duties is to decouple the two pairs of terms – which O’Neill attempts to do (with some hesitation, as we have seen). Nevertheless, my impression is that the connection between the pairs of terms cannot be sacrificed without doing violence to the Kantian texts. We have seen that there is a strict connection between the pairs of terms in the *Vorarbeiten*. This is confirmed in many parts of the Kantian corpus. In the Vigilantius lecture notes (1793/94), which were taken roughly around the same time as the preparation of *The Metaphysics of Morals*, there are formulations that suggest complete identity. Thus laws can be “[d]uties of narrower obligation (*obligatio stricta sive perfecta*), i.e. those obligatory acts that are directly determined by the law, e.g. paying one’s debt” (V-MS/Vigil 27.2: 578).¹⁴ We here encounter the formulation “*obligatio stricta sive perfecta*”, which is accompanied in the same lecture by “*officia lata vel imperfecta*” (V-MS/Vigil 27.2: 669).

I am therefore convinced that there is no escaping to the “equivalence interpretation”, as Fabiola Rivera has called it (2006, 79). That is why I also believe that Kant’s texts cannot be reconciled consistently without admitting that they contain inaccuracies and a certain degree of clumsiness. This is probably due, among other things, to the many years that Kant spent developing a metaphysics of morals – the period in which the lectures were held and which, in a complicated way, was ultimately reflected in his later work. What may be possible is to offer an interpretation that reduces the tensions in Kant’s text without, as it were, “correcting” too much of his work or thought. The point is to make some degree of sense of Kant’s retention of the perfect/imperfect distinction – even if he was not consistent on this front.

If we start from the “core structure” identified previously, we can go further and ask not just what the pairs of terms mean but also why certain obligations belong to one or the other category and what this means for the clarification of the terms. Of course, the central tension is still the existence of perfect duties towards oneself, which should consequently be perfect ethical duties.¹⁵

Not only is the thesis that all ethical duties prescribe maxims exegetically evident, as we have seen, but it is also justifiable. The point is to clarify the consequences of this.

Each rule – and thus also each practical prescription – determines not a singular action but an action type. The step from the rule to the case is made by the power of judgment. If one should in general keep promises, then in this particular case, I am obliged to act according to what can be subsumed under the term “promise” and to keep my promise. This is the first layer of generality that is actually present *every time a concept is applied*. What is it that gives an ethical rule its special “latitude”, and what makes a rule of action an ethical prescription? What is the significance of the claim that ethics concerns not actions but maxims?

5. Maxims, Actions, and the *Spielraum*

According to Lara Denis, Kant often “portrays perfect duties to oneself as directly forbidding certain external actions” (2010, 174). Suicide is banned as morally impermissible, for instance. However, Kant’s description is misleading with regard to the essential features of duties of virtue.¹⁶ Morality does not dictate that one should keep promises (as a duty to others), for example, or that one should not take one’s own life (as a duty to oneself). Rather, it dictates that, out of immediate respect for the moral law (out of duty), the subject should adopt as a maxim determining his will the principle of respecting the “humanity in his person” rather than disparaging it, that is, the principle of not disposing “of oneself as a mere means to some discretionary end . . . debasing humanity in one’s person” (MS 6: 423). In fact, it is not so much an act that is prohibited as “a dishonouring of the worth of humanity in one’s own person” (V-Mo/Collins, 27: 342).¹⁷ Or, even better: what is prohibited is an inner disposition (*Gesinnung*) that allows for the dishonoring of human dignity. Thus perfect duties towards oneself as a *moral* being do not immediately forbid actions that express lying, avarice or false humility;¹⁸ on the contrary, what is commanded is a maxim that obliges granting the agent access to morality, that is, one that obliges the agent not to block this access. What is prohibited is depriving oneself of “inner freedom” or making it “one’s basic principle to have no basic principle and hence no character” (MS 6: 420).¹⁹ These vices – which as such are not types of action but moral attitudes – “adopt principles that are directly contrary to man’s character as a moral being (in terms of its very form), that is, to inner freedom, the innate dignity of man” (MS 6: 420).

In this respect, *ethical duties to oneself cannot be defined otherwise than as wide and imperfect*.²⁰ The rule should be directly applied not to a case but rather to a maxim, which includes a number of possible actions. As a result, there seems to be some *Spielraum* in this case as well. The relevant space is not that which lies between our purpose and the possible actions that can

realize or promote it. Rather, it is the space between the maxim – which commands, from a double perspective (of humans as animal beings and of humans solely as moral beings), that we refrain from impairing human dignity – and specific forms of life conduct (rather than actions per se) that either conform to the maxim or contradict it.

The fact that Kant “fills” this free space with more concrete, specific cases to derive specific, individual duties does not change the structure of the moral action discussed here. On the other hand, the determination of more specific obligations, such as the prohibition against suicide, avarice and so on, is an expression of the fact that duties to oneself, which are referred to as *perfect* duties, are, qua ethical duties, *wide* in terms of obligation and can be divided into several forms depending on the different ways in which their purposes can be realized.²¹ Despite appearances, the determination of action in so-called negative duties is not as clear as is often assumed. Of course, I can only abstain from suicide by not taking my own life (which is quite sharply defined). But since Kant takes into account a continuum of physically self-destructive actions – acts of “disembodiment”²² – what counts as observance or violation of the maxim of respect for humanity in one’s own person is not always so clearly delineated. Where, exactly, does “defiling oneself by lust” (MS 6: 424) or self-numbness through “drunkenness” and “gluttony” (MS 6: 427) begin? With the first, second or third glass?

I am not repeating the argument that goes back to Chisholm and which Kersting rightly rejects as “artificial, unreliable, [and] untrustworthy” (1982, 214, note 78). According to Chisholm, “no duties are perfect” because I can, for example, pay a debt in cash, by check, by bank transfer and so on (1963)²³ This argument applies to any concept that is general by definition and thus allows for a variety of possible specifications (although from a legal point of view, what counts is the specific difference that defines the action, not further specifications that are not mentioned in the law itself). The question does not concern the fact that an action can be carried out in different ways; the point is rather that no action alone implements the maxim, precisely because maxims and not acts are prescribed. In this respect, I do not find the distinction between maxims of actions and maxims of ends applicable.²⁴

This view²⁵ is confirmed by the fact that Kant includes “casuistic questions” in his treatment of perfect duties. According to Kant’s (also somewhat misleading) presentation in *The Metaphysics of Morals*, ethics “inevitably leads” to these questions “because of the latitude that it allows in its imperfect duties”.²⁶ Although Kant refers to casuistry in this passage only in relation to imperfect duties, its presence in the treatment of perfect duties shows that there is *Spielraum* in their case as well:²⁷ without it, casuistic questions would not be conceivable at all (as in the *Doctrine of Right*). Importantly, at least in some cases (such as suicide), the casuistic question refers not to a “simple” application problem, that is, to the question of whether something

is the case of a rule, but specifically *to the question of possible exceptions*: this “wobble room” is not simply the space that always exists between the rule and the case but rather the space in which alternative maxims may be examined, compared and assessed according to their ranking.

6. The Meaning of the *Spielraum*

Given the previous, it would seem that there is no overriding logical coherence in Kant’s use of these terms. After reformulating the traditional concept of perfect/imperfect duties and introducing the new narrow/wide distinction, Kant apparently has difficulty defining the idea of perfect ethical duties – a notion which was already present in the *Groundwork* (in the duty not to commit suicide and the duty to keep one’s promises). Nevertheless, this idea may have a systematic core sense that can be reconstructed and which could explain why Kant does not simply drop these terms. Kant may have thought that the distinction could still play a role in the new context offered by the fundamental narrow/wide distinction, perhaps trying to identify a hidden justification for a classification that had become a matter of course.

I think it is useful to return to the notion of *Spielraum*, which plays a crucial role in dealing with the perfect/imperfect distinction. In *The Metaphysics of Morals*, Kant emphasizes that

a wide duty is not to be taken as a permission to make exceptions to the maxim of actions, but only as permission to limit one maxim of duty by another (e.g., love of one’s neighbor in general by love of one’s parents).

(MS 6: 390)

So this latitude is not just a space of indeterminacy (which in any case exists between the rule and the case): it is rather a space that opens up at the level of the maxims, through which the duty of virtue (based on the moral law) is confronted with a variety of possibly competing maxims. Every duty of virtue moves in this space. *Imperfect* duties prescribe a very general purpose, which in fact cannot be achieved in principle (in this respect always remaining distant), such as one’s own perfection and the happiness of others; the agent can as a consequence deal with situations in which there are many ethically relevant considerations, which may lead to a collision of duties. However, Kant also describes imperfect duties in such a way that the indeterminacy of this *Spielraum* seems to consist primarily of the fact that the *measure* or *degree* of the obligation is left open. As the Vigilantius notes say, “whether, when, how much need to be done” (V-MS/Vigil 27: 536.23–24) may remain indefinite. But notice: what is left open here is not merely *when* and *how much* but also *whether* something is to be done. If one excludes “supererogatory” actions, which are entirely arbitrary and do

not coincide with Kant's "meritorious" actions, then its being open as to *whether* something should be done simply equates to the restriction "of a maxim of duty by another", which Kant speaks of in *The Metaphysics of Morals*. It is strange, however, that Kant discusses this possibility especially in the case of duties that are considered perfect in *The Metaphysics of Morals*, such as in the context of suicide.

Thus perfect duties are to be viewed not as not leaving room for doubt and for the consideration of diverging maxims but as duties that guide the deliberative process *differently than in the case of imperfect duties*. Because of the particular way in which they are justified, they have a fundamental priority that can and should be recognized in the process carried out by the practical power of judgment. In the Vigilantius notes, this is expressed as follows:

lex fortior vincit; regulae si collidunt, a minori fit exceptio. So imperfect duties always succumb to perfect ones, just as several imperfect duties outweigh a single one; for example, the distress of another, were it even to be mortal, could not compel me to contract debts or be grateful, when my parents would starve.

(V-MS/Vigil 27: 537)

The space of possibilities that interests Kant is not so much that in which we can choose ethically indifferent means for given purposes but the horizon of ethically relevant decisions between actions that can have moral meaning (a maxim) and different moral weight. On this horizon, which is common to all moral action, perfect duties take precedence not because they relate directly to actions or because they are negative duties of omission (and therefore more clearly defined) but by virtue of the specific way in which they are justified.

This particular mode of justification is that which Kant – as we saw at the beginning of this chapter – describes in the *Groundwork*. It is apparently not used afterwards but comes back into play in the late Vigilantius ethics lectures.²⁸ After discussing the example of the imperative "You shall absolutely speak the truth", Kant claims that the opposite maxim would lead to a contradictory consequence:

the question arises, whether this maxim could stand as a universal law. We would then have to presuppose that nobody will tell the truth to his disadvantage, and in that case nobody would continue to have any trust; the liar could thus never succeed in deceiving anyone by lying, and the law would therefore automatically destroy itself.

(V-MS/Vigil 27: 496)

In other words – in the words of the *Groundwork* – the maxim of the action "cannot even be *thought* without contradiction as a universal law of

nature” (GMS 4: 424). The text of the lecture continues: “So it is with all perfect duties; if the opposite were to occur, it would so determine the action as to bring about a contradiction with itself, which could never become a universal law” (V-MS/Vigil 27: 496).

The lecture thus expresses the idea, also present in the *Mrongovius II* lecture notes,²⁹ that in the case of *imperfect* duties, it is not the impossibility of thinking but the impossibility of *wanting* the maxim that determines the duty or the prohibition: “With so-called imperfect duties, the situation is quite different. . . . Here the action does not straightway abolish itself by the law that contradicts it” (V-MS/Vigil 27: 496).

The primacy of perfect duties when it comes to the wiggle room of moral deliberation is based on the nature of their foundation: the conceivability of the maxim is a precondition that limits the field of what is morally permissible (first excluding what is “morally impossible”, V-Mo/Mron II 29: 609). Viewed from the opposite perspective, an action that is the result of a maxim that is inconceivable without contradiction undermines the moral value of the *Gesinnung*. In other words: what *directly* hurts humanity in our person has – as a (negative) precondition of morality – priority in a process of deliberation, where different circumstances and possibilities must be taken into account. It is in fact able to orient the decision. Therefore, as we read in the *Vigilantius* lectures, “imperfect duties always succumb to perfect ones” (MS/Vigil 27: 537). What is at stake is thus not whether there is latitude but what role perfect and imperfect duties can play in this space.

The violation of a perfect duty affects the rights of humanity, albeit not in a strictly legal sense,³⁰ and is therefore the first aspect that should be considered. It can work as a preliminary decision criterion, whereas the same does not apply to violations of imperfect obligation, which can never be reliably identified as such and therefore cannot serve as a decision criterion. The greater indeterminacy of actions that violate imperfect obligations (but also those that fulfill them) is a consequence of the nature of the justification (based on the impossibility of willing a certain maxim and not on inconceivability): it does not depend on whether the duties relate to actions or to maxims. An action can always be defined as such with sufficient precision from a juridical point of view: “A good act could be better, but a right act not righter” (19: 232).³¹

It follows that the distinction between perfect and imperfect duties can have meaning within ethics in the context of the deliberation process and the possibility of a collision of duties.³² This distinction indeed played a role in pre-Kantian philosophy, also in this problematic context. This function is completely independent of any distinction between actions and maxims, however – which makes no sense within ethics – and thus of the newly introduced “narrow/wide” concept pair. Kant’s justification does not coincide with the justification of this pair of terms or with the traditional justification based on the presence or absence of enforceability. For example, the

prohibition against self-disembodiment is a perfect duty not because the duty relates to actions or omissions rather than to purposes or maxims (not because it is narrow) and not because it is “owed to” to someone but because the opposite maxim directly contradicts the right of mankind and is therefore unthinkable – and this inconceivability gives it priority over obligations based on the impossibility of willing a certain maxim. Perfect obligations are not without ethical *Spielraum*; rather, they are those obligations that have stronger (and perhaps more immediately comprehensible) grounds of obligation within that ethical *Spielraum*.³³

Notes

- 1 The German terms *eng/wid* are sometimes translated (for example, in the Cambridge translation of *The Metaphysics of Morals*) as “narrow/wide” and at other times as “strict/wide” (see Gregor 1963; Rivera 2006). I follow the first translation, in which the “spatial” meaning prevails, reserving “strict” for *streng*.
- 2 Kant’s works are cited according to volume and page number of the *Akademie-Ausgabe* (Kant 1902–), using the standard German abbreviations for individual titles. I use the following English translations: Kant (2011) for GMS; Kant (1991) for MS; Kant (1997) for V-Mo/Collins, V-Mo/Mron II, V-MS/Vigil. All translations contain references to the German text as well. The translations of VAMS and V-Mo/Kaehler(Stark) are my own.
- 3 According to Ross (1954), p. 54, this formulation leads to an “inadmissible distinction”. See MS, 6: 390: “a wide duty is not to be taken as permission to make exceptions to the maxim of actions”.
- 4 It should be noted that this interpretation changes the literal meaning of Kant’s statement. Rivera (2006) also provides a sympathetic interpretation, relating Kant’s claim to the meritorious/not meritorious distinction (p. 85). Kersting’s interpretation (1982, p. 203) leaves open the possibility that the formulation is not contradictory; in his opinion, Kant is making use of the “argument of the difference in determination”, which ran through the entire 18th-century theoretical debate on duties. According to Kersting, the concrete level of action is not reached in the case of imperfect duties, and thus the influence of inclinations in the fulfillment of imperfect obligations cannot be excluded. Kersting (1993) stresses Kant’s revision of his earlier view in *The Metaphysics of Morals* (MS 6: 390). This reading is also shared by Rivera (2006), pp. 85–86. On the other hand, see V-Mo/Mron II, 29: 633, from the time of the *Groundwork*, where the possibility of exceptions is linked to the distinction between *stricte obligantes* and *late obligantes*. Concerning *The Metaphysics of Morals*, it is also important to note that Kant says that “exactitude cannot be expected in the doctrine of virtue, which cannot refuse some room for exceptions (*latitudinem*)” (MS 6: 233), thus bringing the concept of exception back into play.
- 5 From a moral point of view, we never judge *an action* but always a maxim. In this respect, Allison is right to emphasize that the formulation of the categorical imperative in relation to the natural law “cannot rule out a course of action tout court, as would be required to ground a negative or perfect duty” (2011, p. 179). In my opinion, however, it is an open question whether there are certain types of action that simply cannot be brought under morally permissible maxims. This may be the case, for example, with maxims that cannot be thought of as general laws without contradiction insofar as the concept of action that they use (the

- characterization of the action itself) is contradictory in itself. A concept such as a “non-repayable loan” cannot be “saved” by any maxim and cannot be positively accepted in any moral context. In this sense, a categorical imperative (although it refers to maxims and regulates them alone) can prohibit an action the concept of which is contradictory in itself (and thus makes the maxim contradictory). With this said, it must be noted that refraining from performing the prohibited action does not guarantee the morality of the agent.
- 6 However, it should also be noted that the division principle used in the *Groundwork* to distinguish narrow duties from wide ones, namely the difference between maxims that cannot be *thought* and those that cannot be *wanted*, is used again in the *Vigilantius* lectures for the distinction between perfect and imperfect duties (V-MS/Vigil 27: 496; English translation in Kant 1997) – which appears to entail the correspondence, if not the identity, of both pairs of terms. We will return to this in the following.
 - 7 The CUP translation (which in § 5 omits a sentence) uses “killing oneself” for this expression, which is not exact. Denis (2010) uses “self-disembodiment”.
 - 8 See Gregor (1963), p. 135; Hill (1971); Rivera (2006), among others.
 - 9 So Kersting (1993), p. 192.
 - 10 See the formulation in MS 6: 390: “if the law can prescribe only the maxim of actions, not actions themselves, this is a sign that it leaves a latitude (*latitudo*) for free choice in following (complying with) the law, that is, that the law cannot specify precisely in what way one is to act and how much one is to do by the action for an end that is also a duty”. (In Kant 1991, Mary Gregor translates *Spielraum* as “latitude”; in Kant 1996, however, she uses “playroom”).
 - 11 Of course, each rule of action only prescribes a type of action, and thus this is also what a law in the legal sense does. From a legal point of view, all that matters is that the action is performed. From a moral point of view, however, the maxim counts as well, and this is always a maxim of purposes. In this respect, the idea that only imperfect ethical duties prescribe purposes (see subsequently) is misleading.
 - 12 In addition to the connection between the pairs of terms, the text of the *Vorarbeiten* also clearly expresses the assignment of narrow/perfect duties to right and wide/imperfect duties to ethics. See VAMS 23: 384: “Pflichten sind entweder strikt= oder late= determinirend; jene stehen unter dem Gesetz der Handlungen unmittelbar diese unter dem Gesetz der Maximen der Handlungen (da diese also einen Spielraum für die Willkühr lassen). Jene sind Vollkommene (Rechts=pflichten) diese Unvollkommene d. i. Tugendpflichten”.
 - 13 V-Mo/Kaehler(Stark), 50: “die obligationes externae sind aber perfectae; denn da kommt noch ausser der innern Verbindlichkeit die äussere Nöthigung dazu”.
 - 14 “Pflichten von engerer Verbindlichkeit (*obligatio stricta sive perfecta*) d.i. diejenige Pflichthandlungen, welche unmittelbar durch das Gesetz bestimmt werden, z.E. Bezahlung seiner Schuld” (V-MS/Vigil 27.2: 578). In the same lecture, we encounter “*officia lata vel imperfecta*” (27.2: 669).
 - 15 Mary Gregor reverses the line of argumentation that derives the contradictory character of perfect duties to oneself from the wide/imperfect nature of all ethical duties. According to Gregor (1963, p. 115 ff.), the perfect nature of duties to oneself leads to the conclusion that they are actually “legal duties”: “ethical duties of omission remain something of an anomaly, participating in certain characteristics of both juridical duties and duties of virtue, yet inclining, apparently, toward duties of virtue” (p. 126).
 - 16 This is, on the contrary, Mary Gregor’s assumption and the reason why she designates these duties “juridical duties”: “If, by ‘juridical duty’ we mean . . . rather a duty which consists merely in an action, then we can have juridical duties to

- ourselves as well as to other men” (1963, p. 116). Gregor bases this view on a passage from the *Vorarbeiten*, where Kant speaks of “*officia iuris interni* (erga seipsum)” (VAMS 23: 395). Nevertheless – as Gregor herself recalls – Kant distinguishes the latter from the “*officia iuris externi sive juridical*”, and thus it would be strange to use the term “juridical” for the *officia iuris interni*. Kant actually uses a concept of right in the *Vorarbeiten* that is more general than the legal concept. See 23: 394: “the moral concept of right [*Recht*] (*rectum*) or wrong (*minus rectum*)” expresses “the *minimum* of the action through which a law (rule) can be followed”. However, this term establishes the distinction between a duty owed to someone (*schuldige Pflicht*) (“based on the right of the subject”) and a meritorious duty (“based on the *purpose* of the subject”) and thus a certain understanding of the distinction between perfect and imperfect duties.
- 17 Kant generally relates this expression to “the vices against oneself that are called *Crimina corporis*”. See also V-Mo/Kaehler(Stark), 173–174 (Kant 2004), where the “violation of human dignity in his own person” is mentioned.
 - 18 Untruthfulness (rather than lying), avarice and false humility are obviously not acts but vices, that is, types of thinking that determine how one conducts one’s life as a whole. But even individual acts (such as suicide, mutilation) are expressions of an attitude that makes it possible to deprive “oneself (permanently or temporarily) of one’s capacity for the natural (and so indirectly for the moral) use of one’s powers” (MS 6: 421).
 - 19 It should be noted that Kant speaks in the singular of “man’s duty to himself as an animal being” (6: 421) and of “man’s duty to himself as a moral being *only*” (6: 420); the latter consists “in what is *formal*, in the consistency of the maxims of his will with the *dignity* of humanity in his person”. Bacin (2013) refers to Reflection 8096 (19: 640–641) in this regard, which describes lying as “formal evil”.
 - 20 I therefore agree fully with Rivera (2006) when she claims that “that there cannot be strict duties that proceed from ethical legislation” and that “that there is no good interpretation of Kant’s text that will accommodate his claim that the duty not to commit suicide is strict” (p. 96). To be precise, Kant includes the prohibition against suicide among *perfect* duties to oneself and speaks of these duties as being *streng*, not *eng* (MS 6: 422).
 - 21 This connection is brought up clearly by Kant: “But since ethical obligation to ends, of which there can be several, is only wide obligation – because it involves a law only for the maxims of actions, and an end is the matter (object) of choice – there are many different duties, corresponding to the different ends prescribed by the law, which are called duties of virtue (*officia honestatis*) just because they are subject only to free self-constraint” (MS 6: 395).
 - 22 Kant even views suicide, which is of course a discrete and clearly defined act, as the extreme end of a continuum, which he defines as “self-disembodiment” (which can be partial or total) (6: 421). Thus, at the end of § 6 of the *Doctrine of Virtue* (MS, 6: 423), Kant refers to “partial suicide”, which is realized in actions such as selling one’s teeth or castration.
 - 23 See also O’Neill (2013), pp. 118–119, note 9.
 - 24 See O’Neill (2013), p. 116. Rivera (2006), p. 90, also uses the distinction, albeit with reservations (“the distinction is not clear and sharp”). Rivera claims that some narrow duties can be “underdetermined” (p. 94); in her view, the point is that “in the case of a strict duty it is crucial to fix the facts in order to determine the duty”. I believe that the underdetermination is always there and that what is essential is rather the logic of this underdetermination, that is, what is possible in the *Spielraum*.
 - 25 Rivera’s interpretation moves in a similar direction, emphasizing that one should take Kant at his word when he says that all ethical duties are narrow in

- obligation; the essence of a virtue lies in adopting an attitude rather than carrying out (or omitting) acts (p. 94). However, I do not share her view that the distinction between perfect and imperfect duties consists of the fact that the former are “owed to someone in particular” and therefore are not meritorious (p. 93).
- 26 MS, 6:411: “ethics, because of the latitude that it allows in its imperfect duties, inevitably 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”.
- 27 Schüssler (2012) offers a precise definition of casuistry according to which the casuistic questions are not actually casuistics. Unfortunately, I cannot deal with this here. For the question of casuistry, see La Rocca (2020) and the other contributions in Di Giulio and Frigo (2020).
- 28 In this respect, it is not correct to say that the distinction between the wide criterion of the possibility of thinking (*Denkenkönnen*) and the narrower criterion of the possibility of willing (*Wollenkönnen*) can only be found in the *Groundwork* (Kersting 1983, 406).
- 29 See V-Mo/Mron II, 29: 608–9. “With perfect duties, I ask whether their maxims can hold good as a universal law. But with imperfect ones, I ask whether I could also will that such a maxim should become a universal law” (29: 609).
- 30 See V-Mo/Mron II, 29: 620, where the *leges* are first defined as legal laws, and Kant then says: “In selfregarding actions there are also strict laws [*stricte Gesetze*], i.e., *leges*, but they are not juridical”.
- 31 The broad – not strictly juridical – sense in which Kant refers to the “right of humanity” and its connection with the concept of perfect duty is clear in V-MS/Vigil 27: 603, where we read that “duties to oneself refer . . . always to the right of humanity in one’s person [*das Recht der Menschheit in seiner Person*]” and that “each and every duty is either perfect (or a duty of right) or a duty of love [*entweder vollkommen oder Rechts- oder Liebespflichten*], and thus the duties to oneself are also of this double nature, depending, that is, on whether they refer to the right of humanity or to the end of humanity in a person” (transl. modified). So, according to V-MS/Vigil 27: 604, any transgression of perfect duties “is thus a violation of the right of humanity in our own person; we thereby make ourselves unworthy of the possession of our person that is entrusted to us, and become worthless, since the preservation of our own worth consists solely in observing the rights of our humanity”.
- 32 Strictly speaking, according to Kant, it is not possible to speak of a collision of obligations but only of a collision of obligating reasons or grounds of obligation (*Verpflichtungsgründe*). On this question, see Timmermann (2013). (“Obligating reasons” is the translation preferred by O’Neill 2002, cf. p. 341 f.)
- 33 I thank Carolyn Benson for her very careful linguistic revision of this chapter.

References

- Allison, H.E. 2011. *Kant’s Groundwork for the Metaphysics of Morals. A Commentary*. Oxford: Blackwell.
- Bacin, S. 2013. “The Perfect Duty to Oneself Merely as a Moral Being (TL 6: 428–437).” In A. Trampota, O. Sensen, and J. Timmermann (eds.), *Kant’s Tugendlehre: A Comprehensive Commentary*. Berlin; Boston: de Gruyter, 245–268.
- Chisholm, R. 1963. “Supererogation and Offense.” *Ratio* 5: 1–14.

- Denis, L. 2010. "Freedom, Primacy and Perfect Duties to Oneself." In L. Denis (ed.), *Kant's Metaphysics of Morals: A Critical Guide*. Cambridge: Cambridge University Press, 170–191.
- Di Giulio, S. and Frigo, A. (eds.). 2020. *Kasuistik und Theorie des Gewissens. Von Pascal bis Kant*. Berlin: de Gruyter.
- Gregor, M. 1963. *Laws of Freedom*. Oxford: Oxford University Press.
- Hill, T.E. 1971. "Kant on Imperfect Duty and Supererogation." *Kant-Studien* 62: 55–76.
- Höpfner, L.J.F. 1780. *Warum sind die Menschenpflichten entweder vollkommene oder unvollkommene? und welche Pflichten gehören zu den ersten, welchen zu den letzten Gattung?*, in L.J.F. Höpfner, *Naturrecht des einzelnen Menschen, der Gesellschaft und der Völker*. Gießen: Krieger, 201–240.
- Kant, I. 1991. *The Metaphysics of Morals*. Introduction, Trans. and notes by Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. 1996. *Practical Philosophy*. Edited by and Trans. Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. 1997. *Lectures on Ethics*. Trans. Peter Heath. Cambridge: Cambridge University Press.
- Kant, I. 2004. *Vorlesungen zur Moralphilosophie*. Hrsg. von Werner Stark. Berlin-New York: de Gruyter.
- Kant, I. 2011. *Groundwork of the Metaphysics of Morals*. Edited by J. Timmermann. Trans. M. Gregor. Cambridge: Cambridge University Press.
- Kersting, W. 1982. "Das starke Gesetz der Schuldigkeit und das schwächere der Gütigkeit. Kant und die Pflichtenlehre des 18. Jahrhunderts." *Studia Leibnitiana* 14(2): 184–220.
- Kersting, W. 1983. "Der kategorische Imperativ, die vollkommenen und die unvollkommenen Pflichten." *Zeitschrift für philosophische Forschung* 37: 404–421.
- Kersting, W. 1993. *Wohlgeordnete Freiheit. Immanuel Kants Rechts- und Staatsphilosophie*. Frankfurt a.M.: Suhrkamp.
- La Rocca, C. 2020. *Gewissen und Urteilskraft in Kants Auffassung der Kasuistik*. Di Giulio: Frigo, 177–192.
- O'Neill, O. 2002. "Instituting Principles: Between Duty and Action." In M. Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*. Oxford: Oxford University Press, 331–347.
- O'Neill, O. 2013. *Acting on Principles: An Essay on Kantian Ethics*, second edition. Cambridge: Cambridge University Press.
- Rivera, F. 2006. "Kantian Ethical Duties." *Kantian Review* 11: 78–101.
- Ross, D. 1954. *Kant's Ethical Theory: A Commentary on the Grundlegung zur Metaphysik der Sitten*. New York: Clarendon Press.
- Schüssler, R. 2012 "Kant und die Kasuistik: Fragen zur Tugendlehre." *Kant-Studien* 103: 70–95.
- Sulzer, J.G. 1773. *Versuch einen festen Grundsatz zu finden, um die Pflichten der Sittenlehre und des Naturrechts zu unterscheiden*. In J.G. Sulzer. *Vermischte philosophische Schriften*. Leipzig: Weidmanns Erbe und Reich, 389–398.
- Timmermann, J. 2007. *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge: Cambridge University Press.
- Timmermann, J. 2013. "Kantian Dilemmas? Moral Conflict in Kant's Ethical Theory." *Archiv für Geschichte der Philosophie* 95(1): 36–64.
- Wolff, C. 1763. *Institutiones juris naturae et gentium, in quibus ex ipsa homini natura continuo nexu omnes obligationes et iurfa omina deducuntur*. Halle: Officina Rengeriana.

INDEX

- Abbass, H. 135–136
Adams, R. M. 57, 68–69
Albertzart, M. 52, 69
Allais, L. 20, 66, 69
Allison, H. E. 1, 4, 40, 153, 156
Almeida, M. 66, 69
Ameriks, K. 38, 46
*Anthropologie in pragmatischer
Hinsicht* (Anth) 1, 7, 8, 9, 11, 20, 23,
26–28, 30–31, 41, 44, 123
anthropology 1, 5, 9, 19, 21, 23,
26–28, 30–31, 33–35, 125
*Anthropology from a Pragmatic
Point of View* see *Anthropologie in
pragmatischer Hinsicht*
Aportone, A. 2–4
Appiah, K. A. 51, 69
autonomy 1, 4, 95–96, 102, 104–106,
109, 112–114, 116, 121, 125, 127,
131–133, 135–136
- Bacin, S. 2–3, 106–107
Barber, A. 83–84, 90
Barnhill, A. 66, 68–71
Basedow, J. B. 23–24, 32–35
Battaglia, F. 135–136
belief(s) 6, 12, 33, 51, 53–65, 67–70,
72, 77, 79, 80, 88–89
Bernstein, M. 66, 69
Berstler, S. 81–82, 89–90
Bilge, S. 20–21
Blackburn, S. 3, 37, 45–46
Blake, W. 32, 34
Boey, G. 51, 70
Bok, S. 88, 90
Borges, M. 1–3
Braidotti, B. 135–136
- Bramble, B. 52, 70
Budolfson, M. B. 48–49, 66, 69–71
- Cakic, V. 135–136
Callanan, J. 66, 69
Camus, A. 67, 70
Caranti, L. 1
Carnevale, A. 135–137
Carson, T. L. 79, 84–85, 88–90
Chance, B. 48, 50–60, 63, 66–68, 70
Chignell, A. 2–3, 57
Chisholm, R. 149, 156
Chodowiecki, D. 23
Clarke, S. 135–136
Coady, T. 136
Cokely, E. T. 135–136
Collins, P. Hill 20–21, 23, 28, 119,
148, 153
Corradetti, C. 119, 125
Critique of Practical Reason 1, 7, 9, 14,
16, 26, 39, 41, 43–45, 95–98, 100,
102–105, 114, 119–120, 122–123,
132–133, 135
Critique of Pure Reason 2, 25–26, 30,
32, 38, 56–57, 87, 96, 102–103,
105, 114, 119, 121–123
Critique of the Faculty of Judgment 1,
2, 7, 9, 14, 16, 21–22, 25–26, 28,
32, 38, 39, 41–45, 51–52, 57, 68,
71, 110, 125, 132–133, 135
Crutzen, P. 119
Cuneo, T. 51, 67, 70–71, 90
- Damasio, A. 45
deceiving 45, 76–81, 83–84, 151
decision theory 49, 61–62, 64–66, 114,
127–131, 136, 152

- desire 29, 42, 60, 68, 97, 100–101, 103, 106–107, 120, 128, 132
 despair 51–53, 55, 58–60, 63, 65, 68, 70, 72
 DeWitt, J. 41
Die Metaphysik der Sitten
 see Metaphysics of Morals
 difference-making 47, 50–52, 59, 64, 69
 Di Giulio, S. 156–157
 dignity (human) 1, 6, 31, 94, 120, 125–126, 131–132, 148–149, 155
 dishonesty 3, 73, 75, 81, 84–85, 90
Doctrine of Right 8, 10–12, 14, 17, 20, 101, 105, 117–118, 120, 149, 156
Doctrine of Virtue 1, 4, 42–43, 45, 75, 77–78, 80–81, 88, 94, 100–101, 110–112, 117, 120, 135, 139, 142–144, 153, 155
 Domingues, I. 135–136
 Dutch laws 14
 duties: ethical 143–144, 147–150, 154–155, 157; imperfect 4, 139–143, 146, 148–156; moral i, 2; narrow 139–142, 144, 154–155; perfect 4, 139–143, 146–152, 155–156; wide 4, 139–142, 144, 154

 Ebels-Duggan, K. 68, 70
 Elster, J. 68, 69, 70
 emotions 2, 4, 37, 41–44, 46
 enhancement: external 134; human 126–128, 135–137; internal 126–127; moral i, 2, 4, 127–129, 131–134; pharmacological 129, 131–134
 enlightenment 21–24, 32, 35, 112–113, 123
 equality 6, 8, 9, 124, 135

 faith 67, 70–71
 Fallis, D. 75, 81, 86, 88–90
 Faulmüller, N. 136
 feeling (moral) 2, 3, 7, 24, 33, 37–45, 109, 120
 Fimiani, M. 135–136
 Fink, H. 33, 35
 Flach, W. 123, 125
 Fleischacker, S. 32, 35
 Frey, R. G. 48, 70
 Frierson, P. 40, 41, 45–46
 Fugate, C. 68, 70

 gender 2, 14, 16, 18, 20, 124
 Gigliotti, G. 122, 125
 Giubilini, A. 136
 God 34, 51, 53–54, 56–58, 60, 61, 68, 70–71
 Gouges, Olympe de 14–15
 Graves, F.P. 33, 35
 Gregor, M. 153–155, 157
Groundwork of the Metaphysics of Morals 1, 7, 15, 25–26, 28–29, 31–32, 38–45, 85–86, 88, 103, 108, 121, 127, 133–135, 139–143, 150–154, 156–157
Grundlegung der Metaphysik der Sitten
 see Groundwork of the Metaphysics of Morals
 Guyer, P. 1, 5, 33, 38, 40–41, 45–46

 Halteman, M. 52, 66, 68–71
 Hansson, M. G. 124–125
 happiness 4, 5, 29, 68, 93, 99, 102–104, 115–117, 120, 124, 150
 Harman, E. 52, 68, 71
 Hay, C. 20–21
 Henrich, D. 124–125
 Henry, B. 135–136
 Henson, R. 39, 46
 heteronomy 96–97, 103, 115, 132–133, 135
 Homer 30–31, 34–35
 hope 3, 4, 7, 11, 22–23, 50–52, 54–65, 68, 70–72, 78
 Höpfner, L. J. F. 142, 157
 humanity 3, 8, 15, 17, 23, 29, 30, 33–35, 43, 77, 81, 83, 92–96, 106, 108–111, 113–116, 118, 120, 134, 148, 149, 152, 155–156
 Hume, D. 2, 22, 24, 33, 35, 38, 97, 125

 imperative(s) 1, 4, 12–13, 16–18, 25–26, 29, 42, 63, 67, 99–100, 105, 108, 113, 116, 121, 126, 128–129, 132–133, 135–136, 140–141, 151, 153, 154

 Jacobs, A. 14, 15, 35

 Kagan, S. 66, 71
 Kersting, W. 141, 145, 149, 153–154, 156–157
 Kingston, E. 48, 67, 69, 71
 Kleingeld, P. 2

- Kohler, G. 121, 125
Kritik der praktischen Vernunft
 see *Critique of Practical Reason*
Kritik der reinen Vernunft see *Critique of Pure Reason*
Kritik der Urteilskraft see *Critique of the Faculty of Judgment*
- Lackey, J. 79, 86, 89–90
 Landschoof, J. A. 24, 35
 language: gender-neutral 12, 14, 17, 21;
 inclusive 2, 7, 16–18; race-neutral 12
 La Rocca, C. 2, 4, 112, 114
 Lavazza, A. 135–136
 law(s): general 104, 140, 142, 144, 153;
 moral 1, 2, 25, 29, 39–43, 45–46,
 52, 54–55, 67, 87, 93, 95–108, 111,
 113, 116, 120–121, 123, 126–129,
 132, 148, 150; practical 105, 107,
 116–117, 120, 123; universal 17–18,
 63, 95, 100, 117, 151–152, 156
 lawgiving 97, 117, 119, 120, 132
 Lawless, W. F. 135–136
 Loudon, R. B. 1, 2, 13, 25
 Lütge, C. 34, 36
 lying 2, 3, 73–91, 143, 148, 151, 155
- Macarthur, D. 25, 34–35
 Mahon, J. E. 76, 87–91
 Martin, A. 51, 71
 Martin, M. 48, 71
 Maslen, H. 135–136
 maxim(s) 18, 30, 44, 63, 92, 94–117,
 119–124, 132–133, 140, 142,
 144–146, 148–156
 McCabe, D. 13, 16, 21
 McGeer, V. 68–69, 71
 McKaughan, D. 67, 69, 72
 McPherson, T. 51, 71
 Mecacci, G. 135, 137
 Meibauer, J. 88–91
Metaphysics of Morals 7–8, 11, 17,
 18, 26–27, 29, 42–44, 88, 123, 125,
 132–133, 135, 139, 141–144, 146–157
 Michaelson, E. 48–49, 71
 Mikkola, M. 13, 20–21
 Mills, C. 12–13, 15, 19, 21
 moral psychology 3, 51, 54
 Morrison, I. 40–41, 45–46
 Mozart, Wolfgang Amadeus 24, 33, 36
 Munzel, G. F. 1, 5, 122, 125
 naturalism 22–23, 25, 29, 32–35
 needy 55, 117
 Nefsky, J. 67, 71
 Niesen, P. 19, 21
 Nobis, N. 52, 71
 Norcross, A. 66, 71
 normativity 25–26, 28–29, 33, 35
 Nozick, R. 68, 71
- obligation 7, 42, 71, 83, 86–87, 93,
 95, 111, 115, 117–118, 120, 132,
 140–143, 145–147, 149–150,
 152–153, 155–156
 O'Neill, O. 1, 5, 29, 36, 147,
 155–157
 opportunism 47, 49–50, 52, 66, 70
 Oppy 47–50, 52, 66–67
 Orth, E. W. 123, 125
- Pace, M. 67, 72
 Parcell, J. 66, 72
 Parfit, D. 52, 72
 Parker, C. 23, 36
 Pasternack, L. 67–68, 70
 Paulsen, F. 33, 36
 Petronzio, S. 123, 125
 Pettit, P. 68–69, 72
 Pierce, V. 66, 72
 Pinzani, A. 135
 Pirni, A. 2, 4
 pleasure 37–46, 49, 120
 poverty 53, 116–118, 124
 Prinz, J. 1, 3, 37, 46
 Putnam, H. 28, 34
- Quick, R.H. 23, 36, 62
- race 2, 6, 11–16, 18–21, 28, 69
 racism i, v, 2, 6–7, 11–21, 69
 Raz, J. 66, 72
 Reath, A. 40–41
 Recki, B. 122, 125
 religion 33, 53–54, 57, 68, 70, 72,
 98–99, 101, 123, 125
 respect 6, 32, 39–43, 45–46,
 48, 50, 52, 53, 57–62, 64–65,
 74–75, 80–81, 86, 88, 92, 94–98,
 103–108, 111–112, 115–117,
 119, 128, 138, 141, 144, 148–150,
 153–154, 156
 Ricci, G. 135, 137

INDEX

- right 2, 3, 5, 8, 10, 12, 14, 17–20, 23,
37, 38, 50, 55, 60, 69–70, 76–77,
81, 87–91, 108, 111, 113–115,
117–118, 124, 133–134, 143–144,
149, 152–156
- Rivera, F. 147, 153–155, 157
- Rorty, R. 20–21, 26, 36
- Ross, D. 89, 153, 157
- Rousseau, J. 23–24, 33, 36, 110
- Santoni de Sio, F. 135, 137
- Sanyal, S. 136
- Saul, J. M. 82–83, 90–91
- Savulescu, J. 136
- Schüssler, R. 156–157
- Sensen, O. 90, 120, 125, 156
- sensibility 38–39, 42–45, 96
- sexism i, v, 2, 6–7, 11–13, 16–19
- Sorensen, K. D. 42, 46, 89–91
- Sorensen, R. 42, 46, 89–91
- Spielraum* 138, 142, 144, 145,
148–150, 153–155
- Spinoza 52–53, 59
- Stokke, A. 84, 89, 91
- Stradella, E. 135, 137
- Sulzer, J. G. 145, 157
- Supposed Right to Lie (On a)* 77, 79,
81, 84, 86, 88–91, 109
- sustainability 4, 115–116, 119
- sympathy 38–39
- Timmermann, J. 89–91, 106, 108, 122,
125, 141, 156–157
- Tommasi, F. V. 30, 35–36
- Totaro, F. 135, 137
- transcendental 2, 3, 22, 29–32, 34–35,
56, 102, 106, 121
- truth 56, 79, 80, 91, 151
- Tugendlehre* see *Doctrine of Virtue*
- Van Impe, S. 20, 68, 72
- Varden, H. 14, 20–21
- Vigilantius (lectures)* 76–78, 80, 147,
150–152, 154
- virtue 2, 7, 8, 17, 20, 26, 29, 43, 49,
51, 55, 71, 92–95, 106, 109, 113,
119–120, 144, 148, 150–151,
153–156
- Weber, M. 30, 36
- Willaschek, M. 68, 72
- Williams, B. 76, 82, 88–89, 91
- Wolfe, C. 34, 36
- Wolff, C. 145, 157
- Wolffian school 122–123
- Wood, A. W. 5, 67–68, 72, 89, 91



Taylor & Francis Group
an informa business



Taylor & Francis eBooks

www.taylorfrancis.com

A single destination for eBooks from Taylor & Francis with increased functionality and an improved user experience to meet the needs of our customers.

90,000+ eBooks of award-winning academic content in Humanities, Social Science, Science, Technology, Engineering, and Medical written by a global network of editors and authors.

TAYLOR & FRANCIS EBOOKS OFFERS:

A streamlined experience for our library customers

A single point of discovery for all of our eBook content

Improved search and discovery of content at both book and chapter level

REQUEST A FREE TRIAL
support@taylorfrancis.com

 **Routledge**
Taylor & Francis Group

 **CRC Press**
Taylor & Francis Group